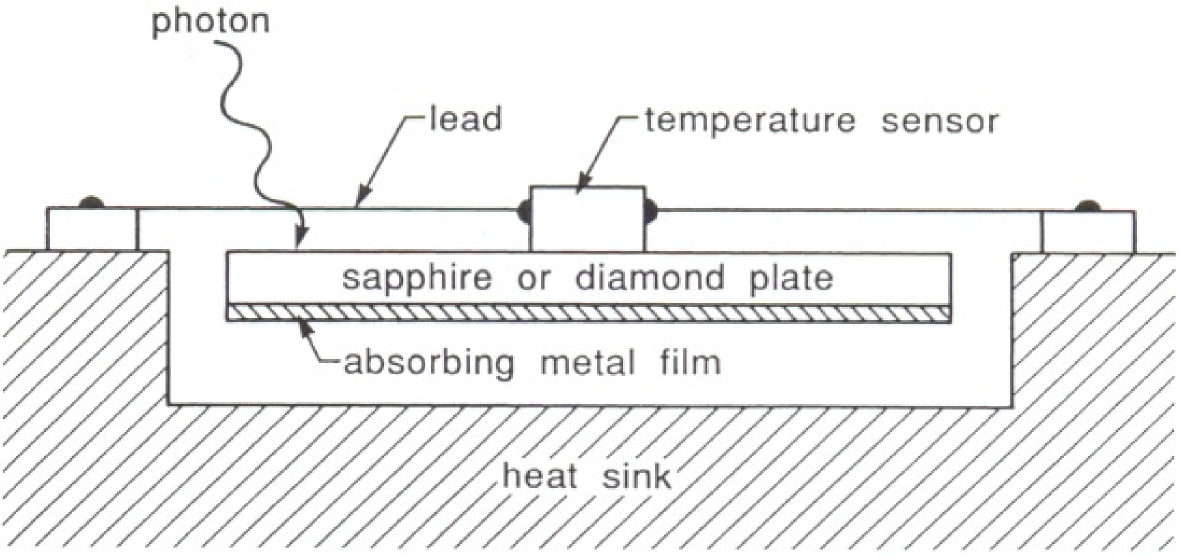
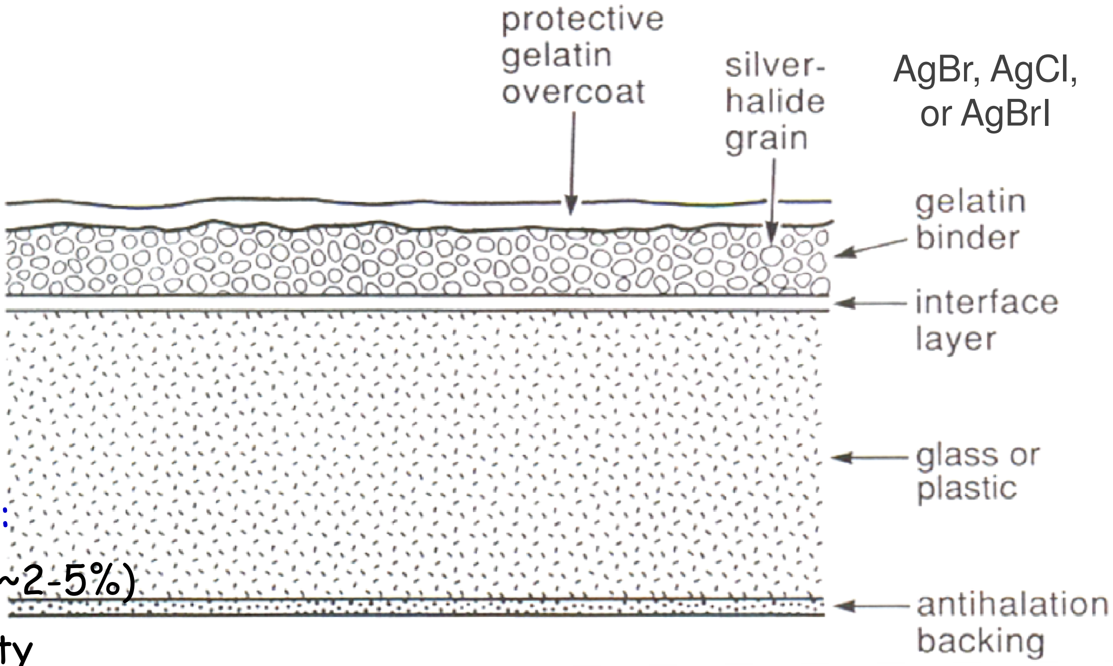
Astronomische Waarneemtechnieken (Astronomical Observing Techniques)

8th/9th Lecture: 10/17 November 2010



Based on "Detection of Light - from the Ultraviolet to the Submillimeter", by George Rieke, 2nd Edition, 2003, Cambridge University Press, ISBN 0-521-01710-6.

The Photographic Plate



Disadvantages:

- Low DQE (~2-5%)
- Non-linearity
- Non-uniformity
- Time resolution
- Wavelength coverage
- Digitization

Advantages:

- 10" plate → 10^{12} grains → ~10^9 pixels
- Inexpensive
- Include own data storage system
- Stable over very long periods of time

Three Basic Types of Detectors

1. **Photon detectors**

   Respond directly to individual photons → releases bound charge carriers. Used from X-ray to infrared.

   Examples: photoconductors, photodiodes, photoemissive detectors

2. **Thermal detectors**

   Absorb photons and thermalize their energy → modulates electrical current. Used mainly in IR and sub-mm detectors.

   Examples: bolometers

3. **Coherent receivers**

   Respond to electrical field strength and preserve phase information (but need a reference phase "local oscillator"). Mainly used in the sub-mm and radio regime.

   Examples: heterodyne receivers

Outline

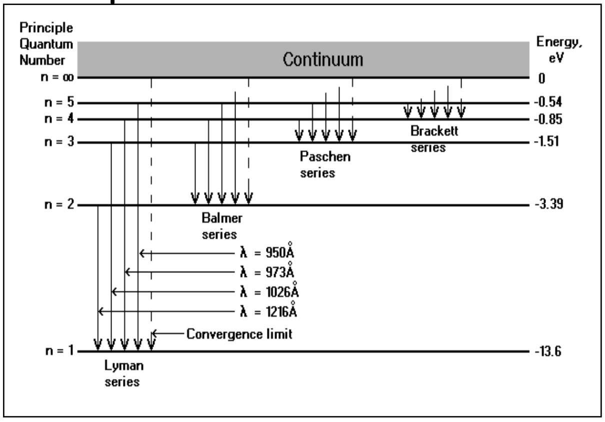
1. Photon Detectors
   a) Principle of (intrinsic) Photoconductors
   b) Variations of Photoconductors
   c) Readout and Operations
   d) Detector Artifacts
2. Thermal Detectors
   a) Principle of Bolometers
   b) Variations of Bolometers
3. Coherent Receivers
   a) Principle of Heterodyne Receivers
   b) Key components of Heterodyne Receivers
   c) Performance characteristics

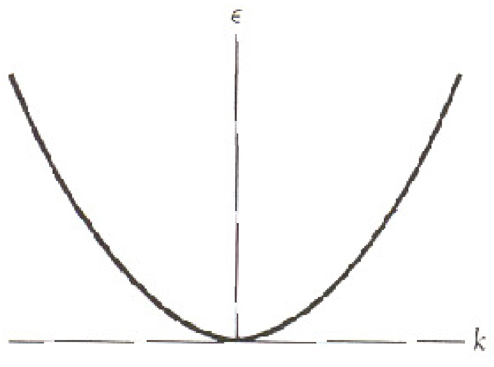
Part I
Photon Detectors
Part II
Thermal Detectors
Part III
Coherent Receivers

Preface: Electronic States and Bands

Single atomic system

Example: H atom



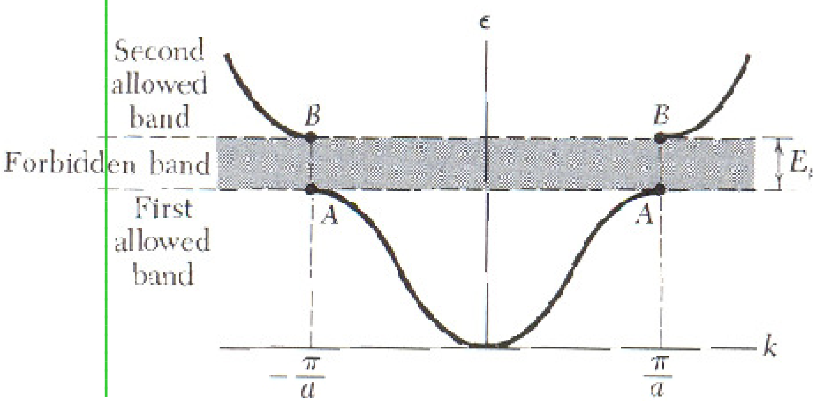


Atomic crystal

Wavefunctions Ψ overlap

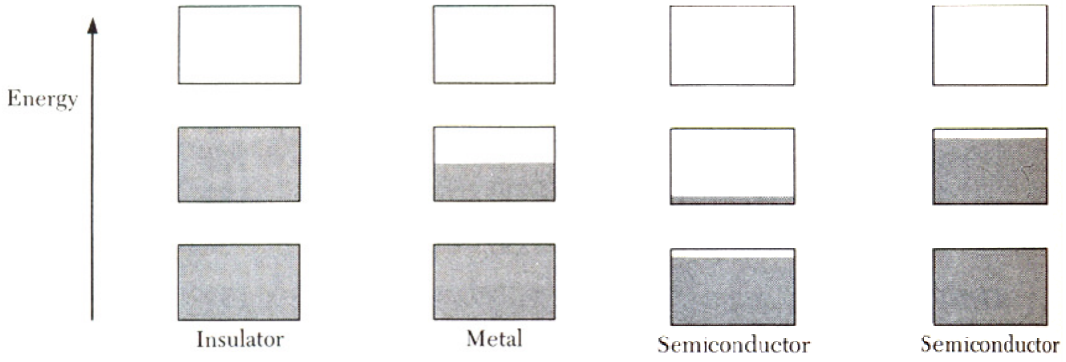
→ Energy levels of individual atoms split due to Pauli principle (avoiding the same quantum states)

→ Multiple splitting ➔ "bands"



Preface: Electric Conductivity

Conductivity requires charge carriers in the conduction band

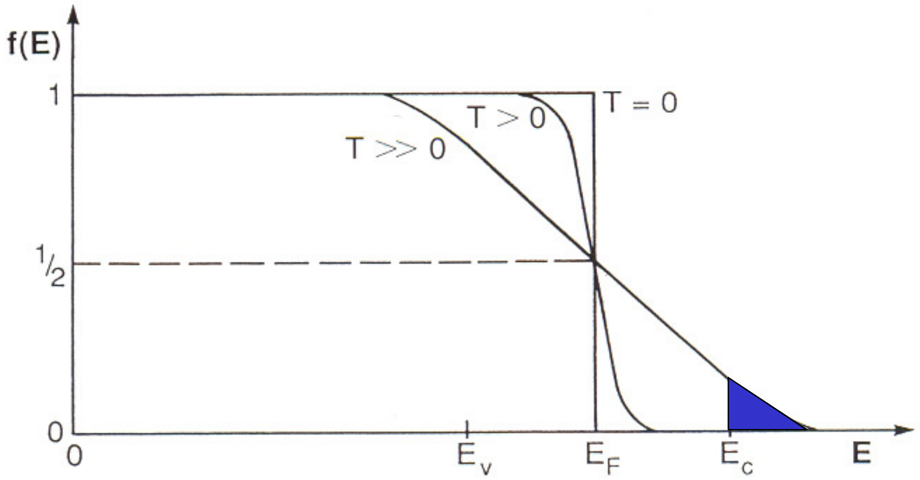


One needs to overcome the bandgap E_g to lift an e^- into the conduction band. This can be done via:

1. external excitation, e.g. via a photon ←photon detector
2. internal excitation due to thermal energy
3. impurities

Preface: The Fermi Energy

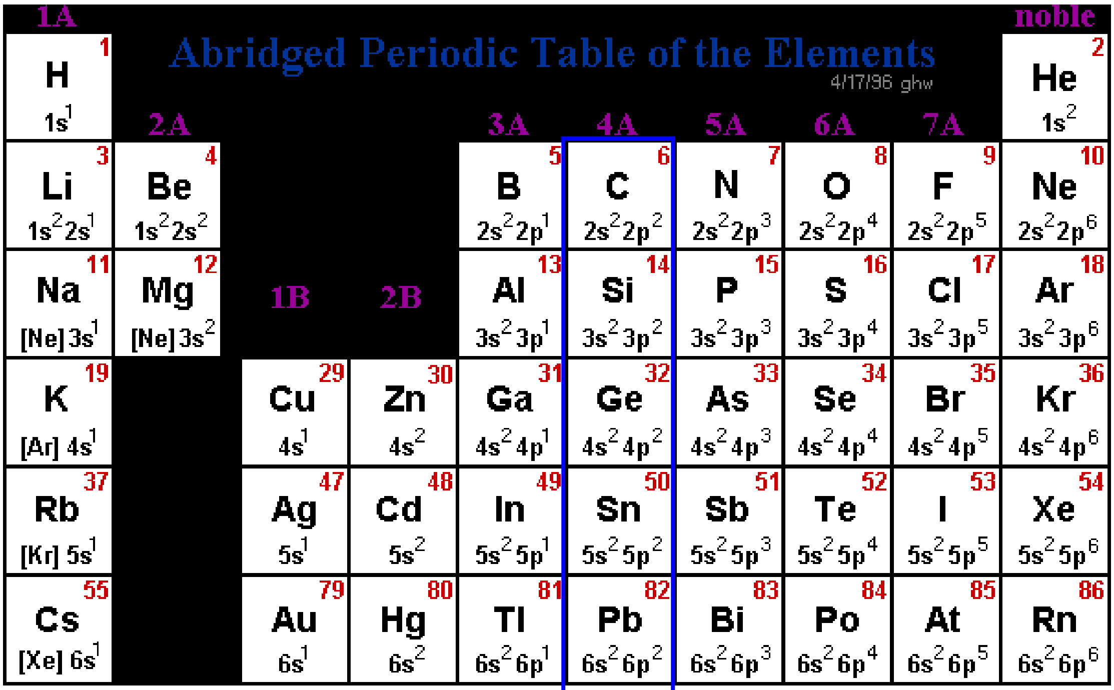
The Fermi energy E_F determines the concentration of thermally excited electrons in the conduction band.



$$n_0 = N_c f(E_c) \quad \text{where} \quad N_c = 2\left(\frac{2\pi m_{eff} kT}{h^2}\right)^{3/2}$$

$$\text{and} \quad f(E_c) = \frac{1}{1+e^{(E_c-E_F)/kT}} \overset{E_c-E_F>>kT}{\approx} e^{-(E_c-E_F)/kT}$$

Preface: The Periodic System of the Elements



Abridged Periodic Table of the Elements

4/17/96 ghw

1A	2A	1B	2B	3A	4A	5A	6A	7A	noble
1 **H** $1s^1$									2 **He** $1s^2$
3 **Li** $1s^2 2s^1$	4 **Be** $1s^2 2s^2$			5 **B** $2s^2 2p^1$	6 **C** $2s^2 2p^2$	7 **N** $2s^2 2p^3$	8 **O** $2s^2 2p^4$	9 **F** $2s^2 2p^5$	10 **Ne** $2s^2 2p^6$
11 **Na** $[Ne] 3s^1$	12 **Mg** $[Ne] 3s^2$			13 **Al** $3s^2 3p^1$	14 **Si** $3s^2 3p^2$	15 **P** $3s^2 3p^3$	16 **S** $3s^2 3p^4$	17 **Cl** $3s^2 3p^5$	18 **Ar** $3s^2 3p^6$
19 **K** $[Ar]\ 4s^1$		29 **Cu** $4s^1$	30 **Zn** $4s^2$	31 **Ga** $4s^2 4p^1$	32 **Ge** $4s^2 4p^2$	33 **As** $4s^2 4p^3$	34 **Se** $4s^2 4p^4$	35 **Br** $4s^2 4p^5$	36 **Kr** $4s^2 4p^6$
37 **Rb** $[Kr]\ 5s^1$		47 **Ag** $5s^1$	48 **Cd** $5s^2$	49 **In** $5s^2 5p^1$	50 **Sn** $5s^2 5p^2$	51 **Sb** $5s^2 5p^3$	52 **Te** $5s^2 5p^4$	53 **I** $5s^2 5p^5$	54 **Xe** $5s^2 5p^6$
55 **Cs** $[Xe]\ 6s^1$		79 **Au** $6s^1$	80 **Hg** $6s^2$	81 **Tl** $6s^2 6p^1$	82 **Pb** $6s^2 6p^2$	83 **Bi** $6s^2 6p^3$	84 **Po** $6s^2 6p^4$	85 **At** $6s^2 6p^5$	86 **Rn** $6s^2 6p^6$

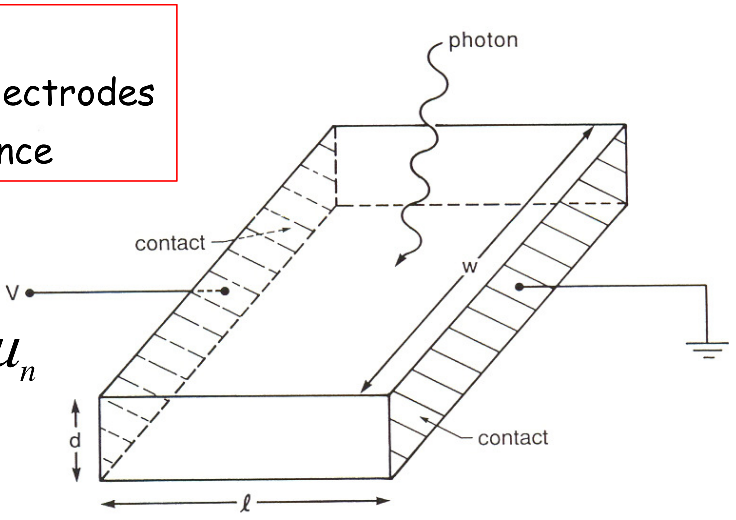
"classical" semiconductors: 4 e^- in valence state (outer shell)

Elements with 4 e^- form crystals with diamond lattice structure (each atom bonds to four neighbors)

Principle of Photoconductors

Basic Principle of intrinsic Photoconductors

- E_γ lifts e^- into conduction band
- electric field $\bar{E}$ drives charges to electrodes
- few charge carriers → high resistance



Conductivity: $\sigma = \frac{1}{R_d}\frac{l}{wd} = qn_0\mu_n$

where:
R_d = resistance
w,d,l = geometric dimensions
q = electric charge
n_0 = number of charge carriers $n = \frac{\varphi\eta\tau}{wdl}$
φ = photon flux
τ = mean lifetime before recombination
μ_n = electron mobility ~ mean time between collisions.

Important Quantities and Definitions

Quantum efficiency η $\equiv \frac{\text{\# absorbed photons}}{\text{\# incoming photons}}$

Responsivity S $\equiv \frac{\text{electrical output signal}}{\text{input photon power}}$

Wavelength cutoff: $\lambda_c = \frac{hc}{E_g} = \frac{1.24\mu m}{E_g[eV]}$

→ Germanium: 1.85μm
→ GaAs: 0.87μm

Photo-current: $I_{ph} = q\varphi\eta G$

Photoconductive gain G: $G = \frac{I_{ph}}{q\varphi\eta} = \frac{\tau}{\tau_t} = \frac{\text{carrier lifetime}}{\text{transit time}}$

The product ηG describes the probability that an incoming photon will produce an electric charge that will penetrate to an electrode.

The main Noise Components

G-R noise

$$\left\langle I_{G-R}^2 \right\rangle = 4q^2 \varphi \eta G^2 \Delta f$$

fundamental statistical noise due to the Poisson statistics of the incoming photon stream → transferred into the statistics of the generated and recombined holes and electrons → two independent statistical processes $(2N)^{1/2}$.

Johnson or kTC noise

$$\left\langle I_J^2 \right\rangle = \frac{4kT}{R} \Delta f$$

fundamental thermodynamic noise due to the thermal motion of the charge carriers. Consider a photoconductor as an RC circuit: fluctuations in $E_{storage}$ are associated with a noise current I_J. Since $\langle Q^2 \rangle = kTC$, the charge noise is also called kTC noise or reset noise.

1/f noise

$$\left\langle I_{1/f}^2 \right\rangle \propto \frac{I^2}{f} \Delta f$$

increased noise at low frequencies, due to bad electrical contacts, temperature fluctuations, surface effects (damage), crystal defects, and JFETs. Physical origin mostly unclear.

Some Comments on the Noise

The total noise in the system is: $\left\langle I_N^2 \right\rangle = \left\langle I_{G-R}^2 \right\rangle + \left\langle I_J^2 \right\rangle + \left\langle I_{1/f}^2 \right\rangle$

The bandwidth $\Delta f = \frac{1}{2\Delta t_{\text{int}}}$; if Poisson-distributed → relative error ~ $1/\sqrt{t}$

Operationally, background-limited performance (BLIP) $\left\langle I_{G-R}^2 \right\rangle >> \left\langle I_J^2 \right\rangle + \left\langle I_{1/f}^2 \right\rangle$ is always preferred*

The noise equivalent power (NEP) is the signal power that yields an RMS S/N of unity in a system of $\Delta f = 1$ Hz:

$$NEP_{G-R} = \frac{2hc}{\lambda} \left(\frac{\varphi}{\eta} \right)^{1/2}$$

In BLIP the NEP can only be improved by increasing the quantum efficiency η.

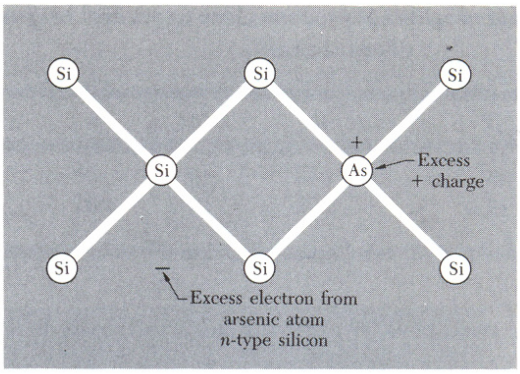
Variations of Photoconductors

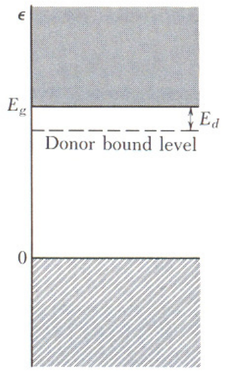
Extrinsic Semiconductors

Limitations of intrinsic photoconductors:

- short wavelength cutoffs
- non-uniformity of material
- problems to make good electrical contacts
- difficult to "keep clean" and minimize Johnson noise

Solution: add impurities at low concentration to provide excess electrons → much reduced bandgap → longer wavelength cutoff





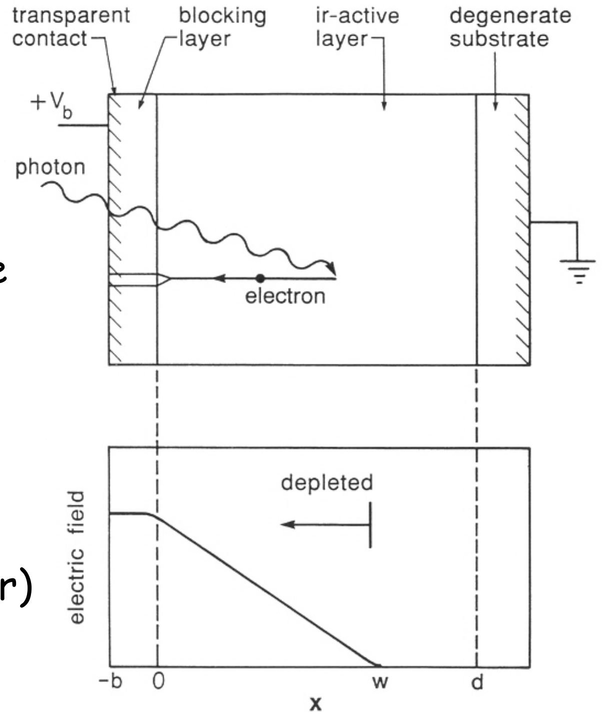
		Ge	Si
Impurity	Type	Cutoff wavelength λ_c (μm)	Cutoff wavelength λ_c (μm)
Al	p		18.5^a
B	p	119^b	28^a
Be	p	52^b	8.3^a
Ga	p	115^b	17.2^a
In	p	111^b	7.9^a
As	n	98^b	23^a
Cu	p	31^b	5.2^a
P	n	103^b	27^a
Sb	n	129^b	29^a

Example: addition of boron to silicon in the ratio 1:100,000 increases its conductivity by a factor 1000!

Blocked Impurity Band (BIB) Detectors

Limitation of extrinsic photoconductors: absorption coefficients are 2 – 3 orders of magnitude less than those for direct absorption in intrinsic photoconductors → low QE → active volumes must be large → conflicts:

- Efficient absorption [$\alpha(\lambda) = \sigma(\lambda) N_I$] requires large N_I → high conductivity
- Noise requirements: high resistance = low conductivity



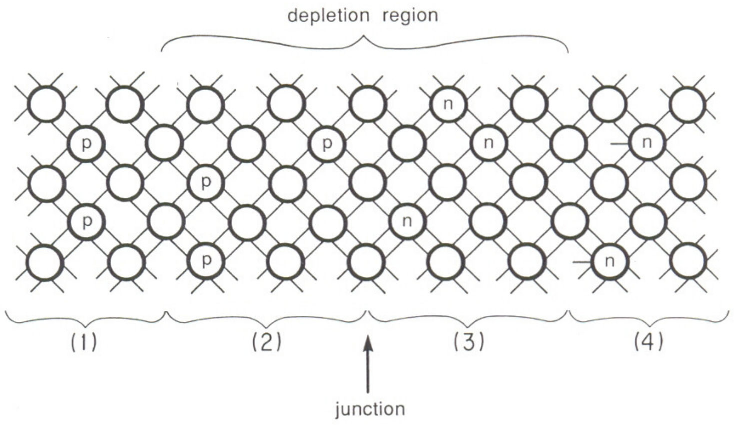
Solution: use separate layers to optimize the optical and electrical properties independently:

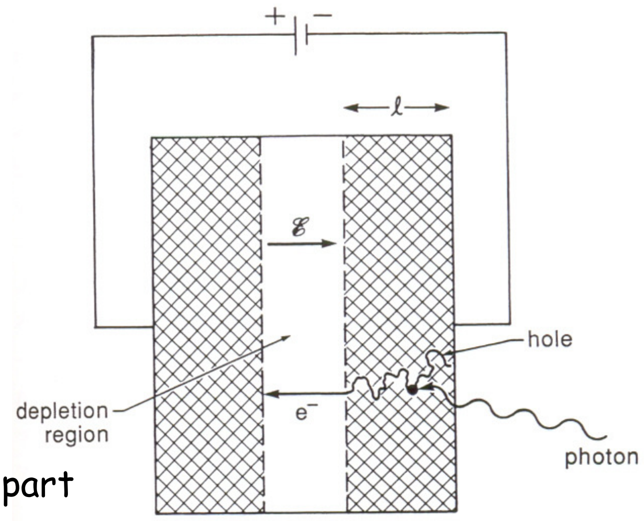
IR-active layer: heavily doped
Blocking layer: thin layer of high purity
(intrinsic photoconductor)

Typical species are Si:As or Si:Sb

Photodiodes

1. Based on junction between *two* oppositely doped zones
2. The two adjacent zones create a depletion region with high impedance

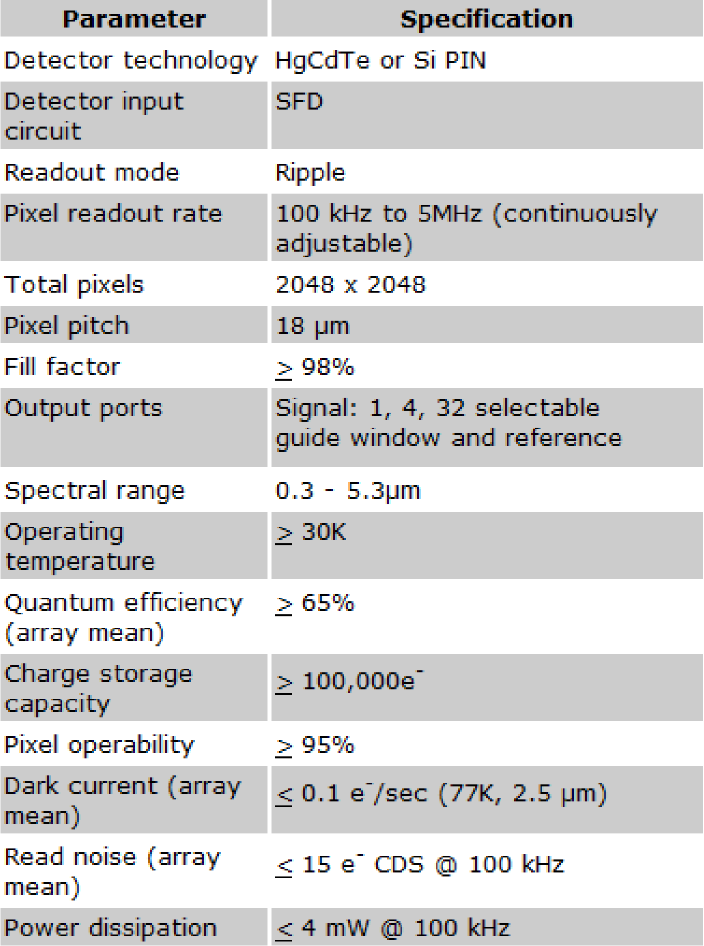




1. Photon gets absorbed e.g. in the p-type part
2. Absorption creates an e^--hole pair
3. The e^- diffuses through the material
4. Voltage drives the e^- across the depletion region → photo-current

Same if absorbed in the n-type part but then the hole migrates through the junction.

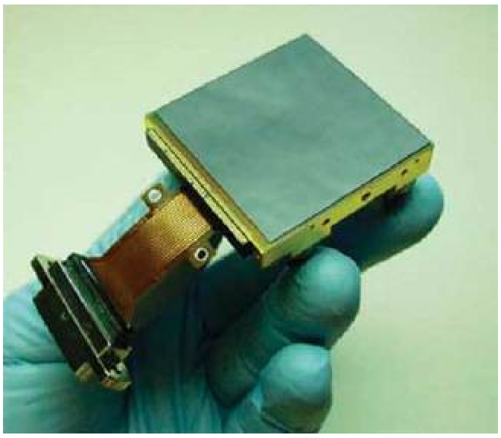
Example: The Teledyne HAWAII-2RG



Parameter	Specification
Detector technology	HgCdTe or Si PIN
Detector input circuit	SFD
Readout mode	Ripple
Pixel readout rate	100 kHz to 5MHz (continuously adjustable)
Total pixels	2048 x 2048
Pixel pitch	18 μm
Fill factor	$\geq$ 98%
Output ports	Signal: 1, 4, 32 selectable guide window and reference
Spectral range	0.3 - 5.3μm
Operating temperature	$\geq$ 30K
Quantum efficiency (array mean)	$\geq$ 65%
Charge storage capacity	$\geq$ 100,000e^-
Pixel operability	$\geq$ 95%
Dark current (array mean)	$\leq$ 0.1 e^-/sec (77K, 2.5 μm)
Read noise (array mean)	$\leq$ 15 e^- CDS @ 100 kHz
Power dissipation	$\leq$ 4 mW @ 100 kHz

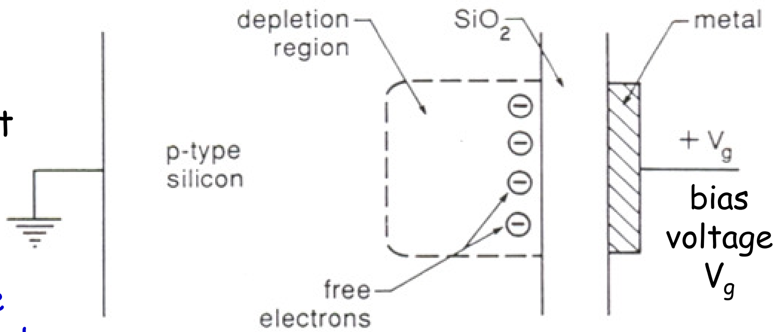
See `http://www.rsc.rockwell.com/imaging/hawaii2rg.html` for more info

Can also be combined to a 2x2 mosaic



Charge Coupled Devices (CCDs)

CCDs = array of integrating capacitors.
Pixel structure: metal "gate" evaporated onto SiO_2 (isolator) on silicon = MOS



1. photons create free e^- in the photoconductor
2. e^- drift toward the electrode but cannot penetrate the SiO_2 layer
3. e^- accumulate at the Si—SiO_2 interface
4. the total charge collected at the interface is a measure of the number of photons during the exposure
5. → read out the number of e^-

Note that there are two types: front-illuminated and back-illuminated CCDs.

Front-illuminated: electrode of heavily doped Si blocks blue/UV photons

Back-illuminated: long distance to depletion region → low QE → thinning

Next: Detector Operations

CCDs and infrared arrays are fundamentally different!

CCDs:

- destructive reads
- charges are physically shifted to the output line
- shutter determines exposure time

IR arrays:

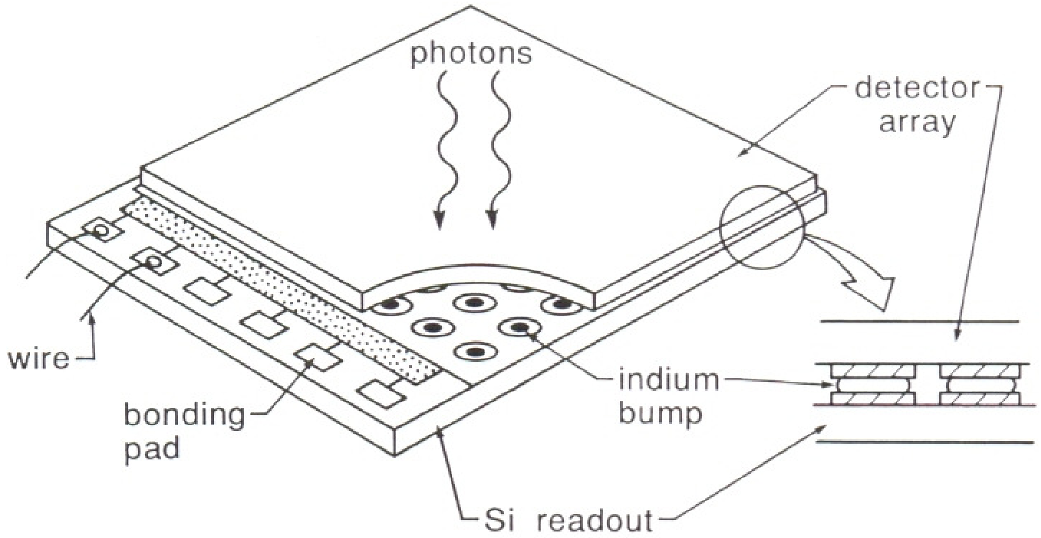
- non-destructive reads
- readout addresses individual pixels directly
- read/reset determines exposure time

Readout and Operations

I. Infrared Arrays

Infrared Arrays – Construction

1. Produce a grid of readout amplifiers
2. Produce a (matching mirror image) of detector pixels
3. Deposit Indium bumps on both sides
4. Squeeze the two planes together ➔ hybrid arrays
5. The Indium will flow and provide electrical contact



Multiplexers

Multiplexing: "Pixel signals → Sequential output lines"

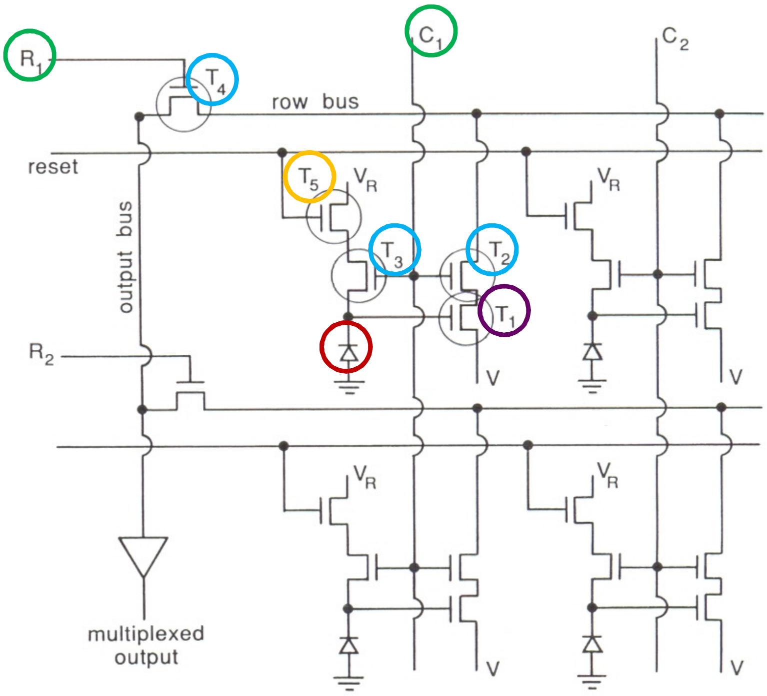
MUX Tasks:
- address a column of pixels by turning on their amplifiers
- pixels in other columns with power off will not contribute a signal

Signal at photodiode → gate T_1

Readout uses row driver R_1 and column driver C_1 to close the switching transistors T_2, T_3, T_4.

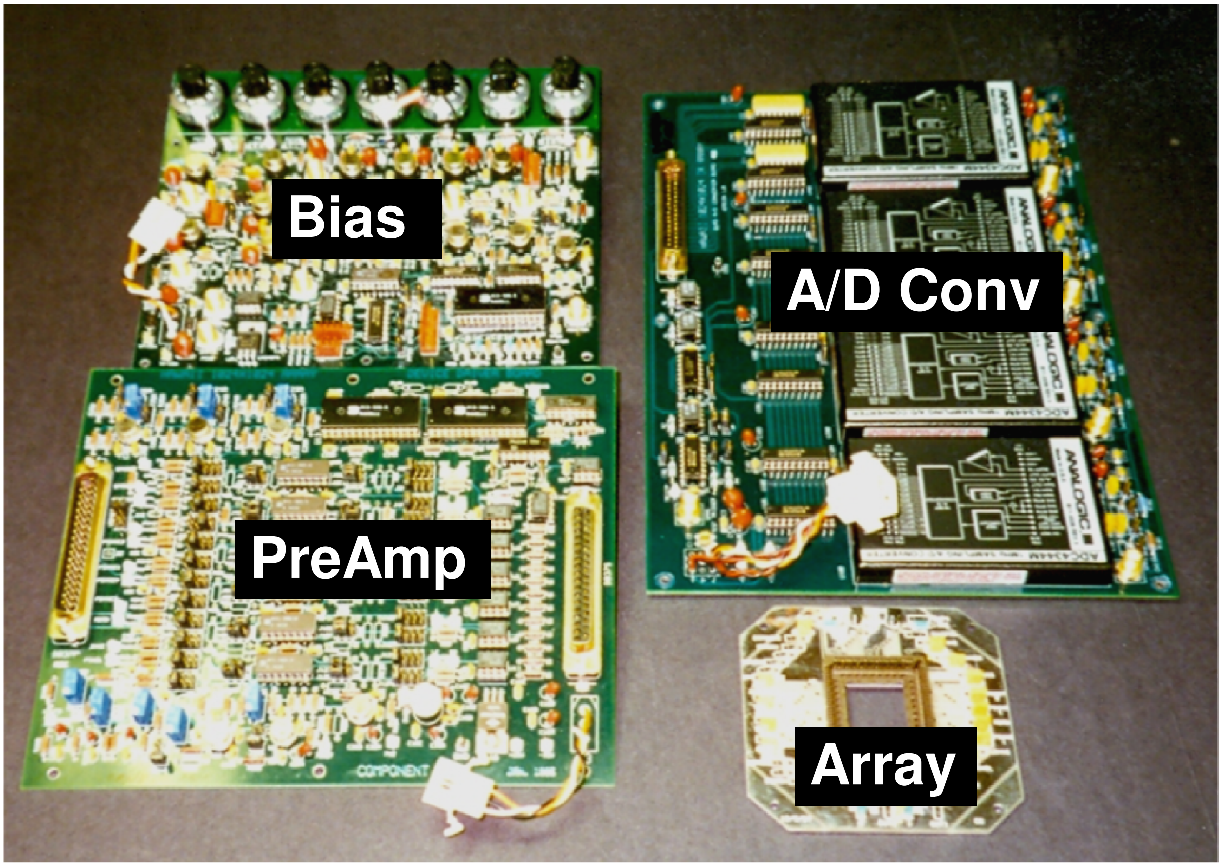
→ Power to T_1 → signal to the output bus

Reset: connect V_R via T_5 and T_3.



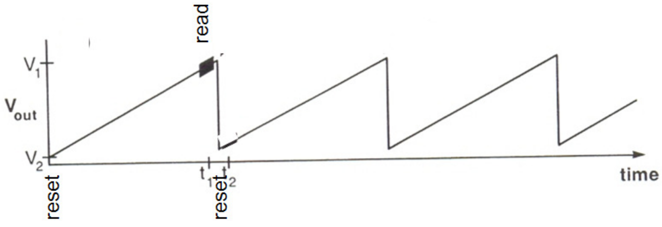
Elements of a Detector Electronics System

Example: PHARO (the Palomar High Angular Resolution Observer)



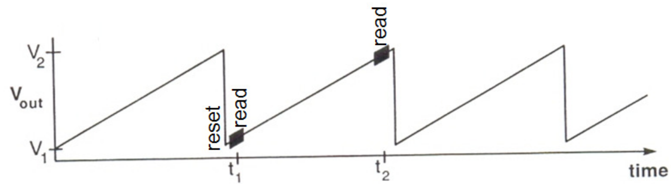
IR Array Read Out Modes

Single Sampling



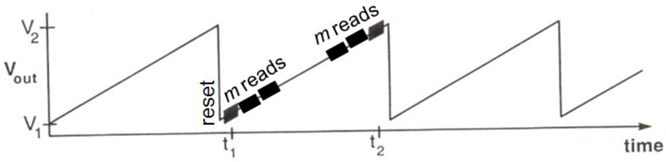
- most simple approach
- does not remove kTC noise
- measures the absolute signal level

Reset-Read-Read



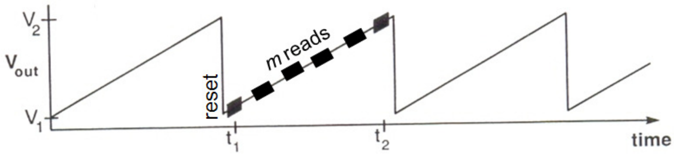
- Resets, reads and reads pixel-by-pixel
- Signal = Read(2) – Read(1)
- best correlation, no reset noise
- but requires frame storage
- reduced dynamical range (saturation!)

(Multiple) Fowler Sampling



- similar to reset-read-read ...
- ... but each read is repeated *m* times
- Signal = mean(read2) – mean(read1)
- Reduces readout noise by √m over RRR

Sample-up-the-ramp Fitting



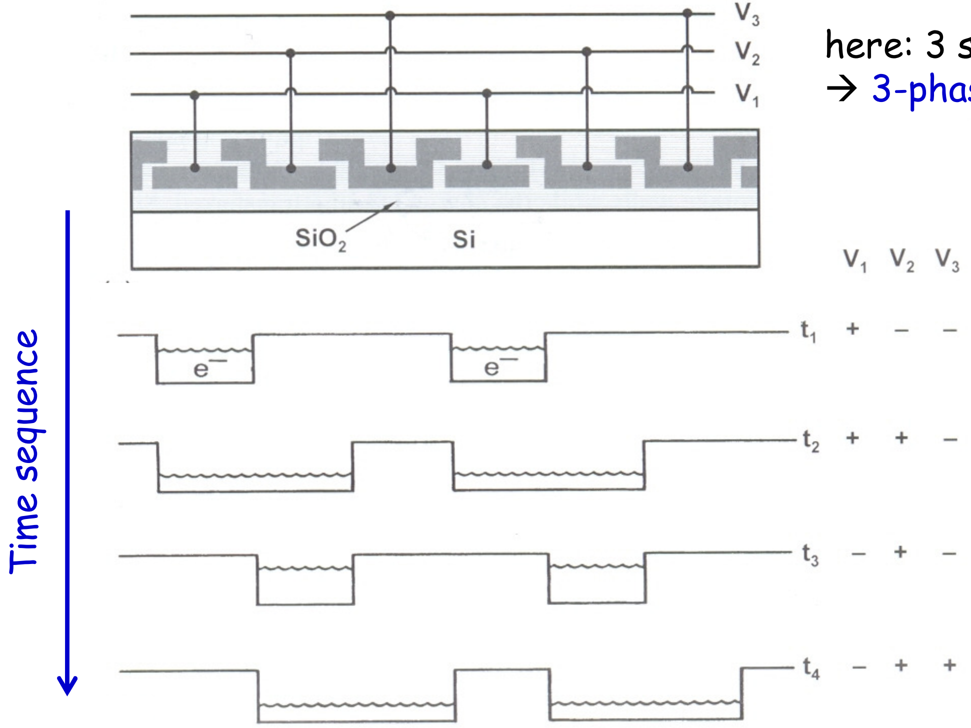
- *m* equidistant reads during integration
- linear fit → "slope"
- reduces readout noise by √m
- particularly useful in space (cosmics!)

Readout and Operations

II. CCDs

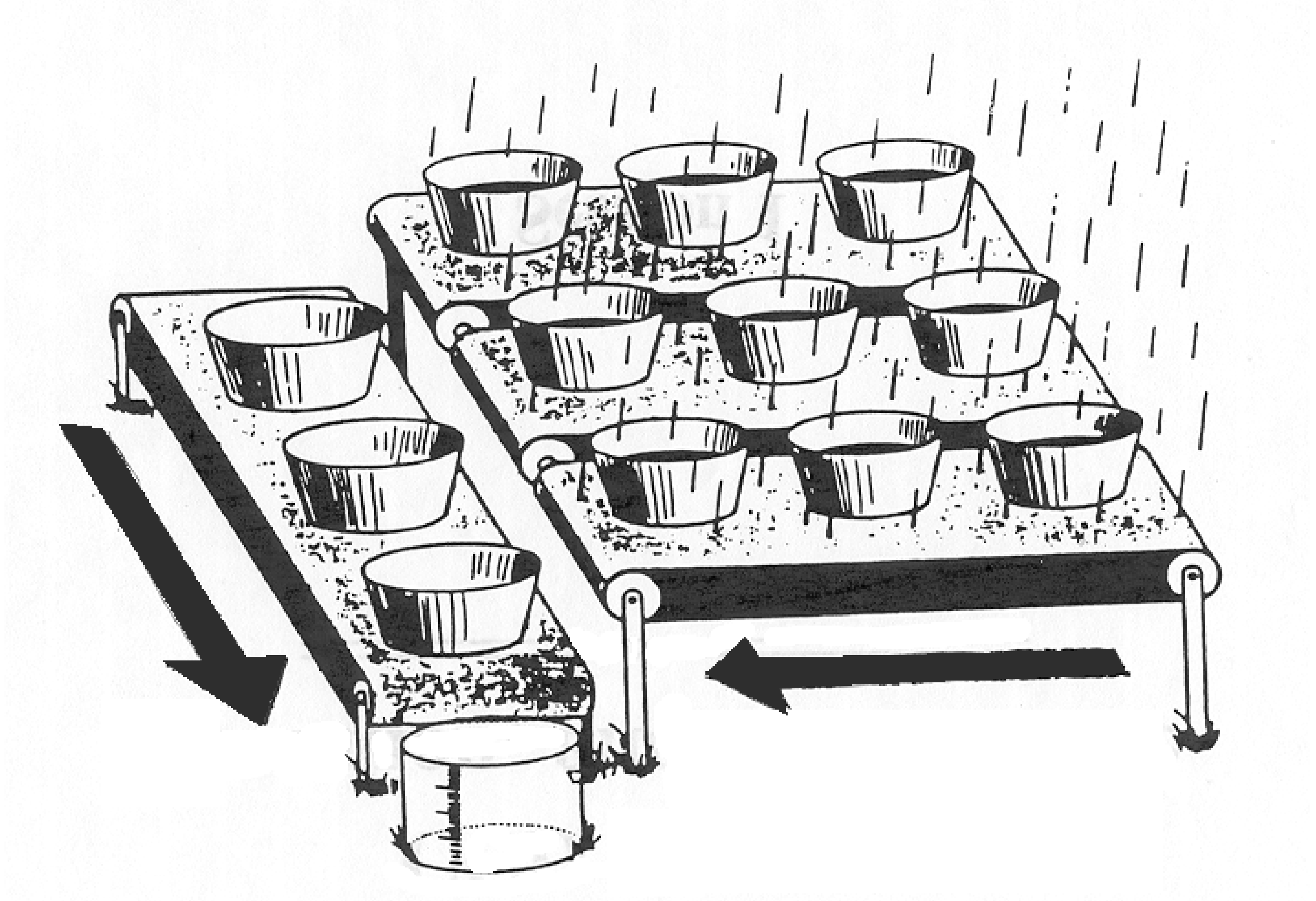
Charge Coupled Readouts

Collected charges are passed along the columns to the edge of the array to the output amplifier.



here: 3 sets of electrodes
→ 3-phase CCD

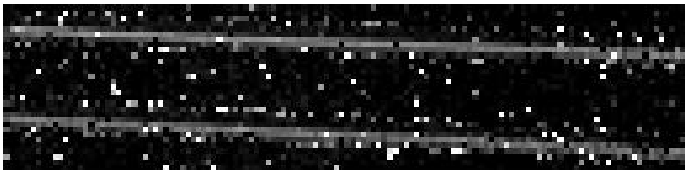
Be aware of charge transfer (in-)efficiencies (CTEs) due to electrostatic repulsion, thermal diffusion and fringing fields.



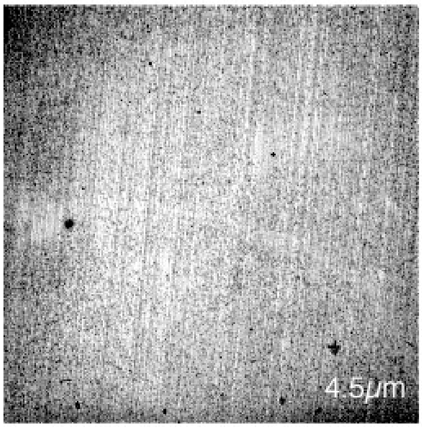
http://solar.physics.montana.edu/nuggets/2000/001201/ccd.png

Detector Artefacts

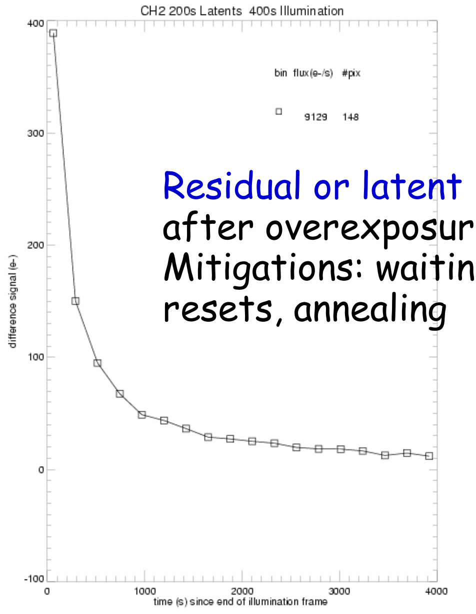
Detector Artefacts (1)



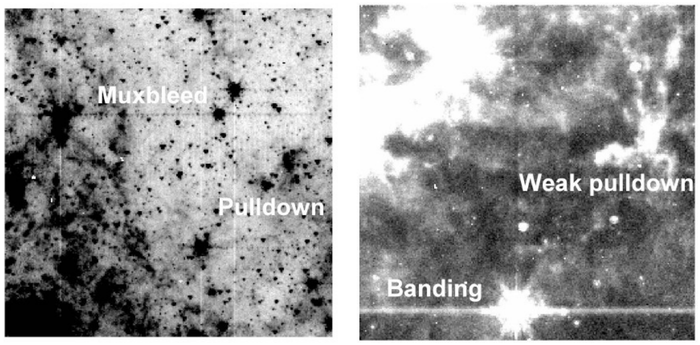
Dead, hot and rogue pixels.
Mitigation: subtract off-source image and/or reduce bias voltage



Fixed pattern noise.
Mitigation: "flat-fielding"



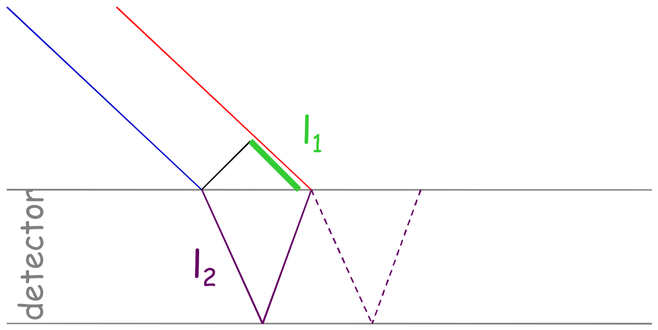
Residual or latent images after overexposure.
Mitigations: waiting, frequent resets, annealing

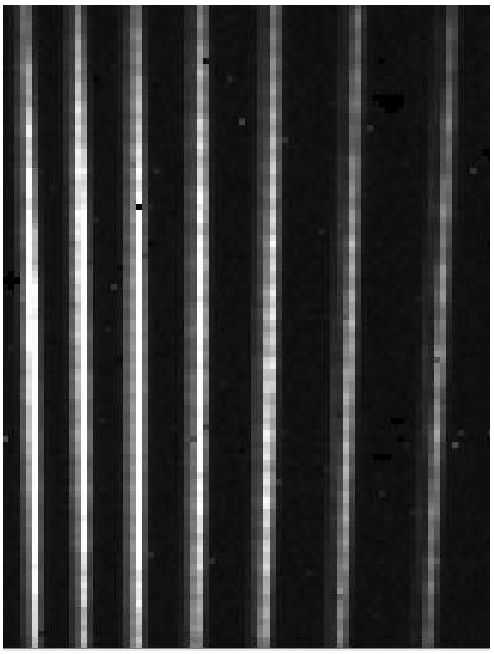


Muxbleed, pulldown and banding.
Mitigation: avoid bright sources, short exposures.

Detector Artefacts (2): Fringing

In spectrographs: photons reflect off the back of the detector and interfere with the incoming light.





If the phase difference between l_1 and $n \cdot l_2$ is an even multiple of π constructive interference occurs. If an odd multiple destructive interference occurs → fringes = wave pattern.

Part I

Photon Detectors

Part II

Thermal Detectors

Part III

Coherent Receivers

Principle of Bolometers

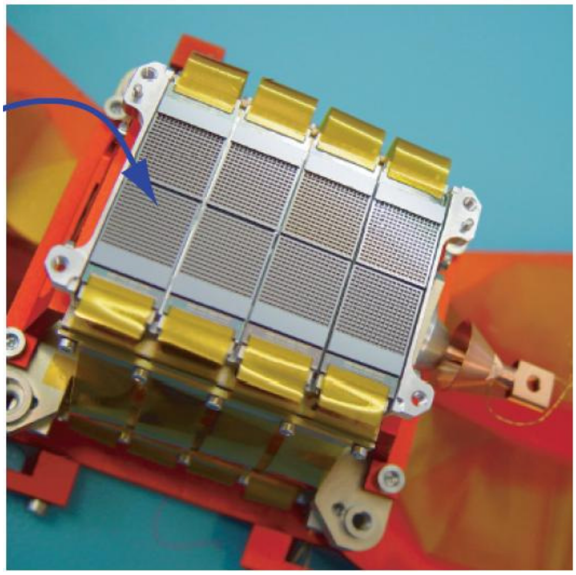
Bolometers

So far: photon detection via direct excitation of charge carriers

Now: photon absorption and conversion into energy (heat)

- Absorber is decoupled from the detection process
- Especially for low light levels
- Especially for the far-IR & sub-millimeter wavelength range

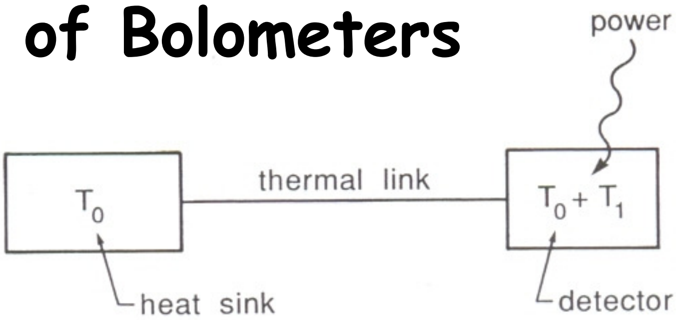
Invention of the Ge:Ga bolometer in 1961 by Frank Low



Herschel / PACS bolometer: a cut-out of the 64x32 pixel bolometer array assembly.

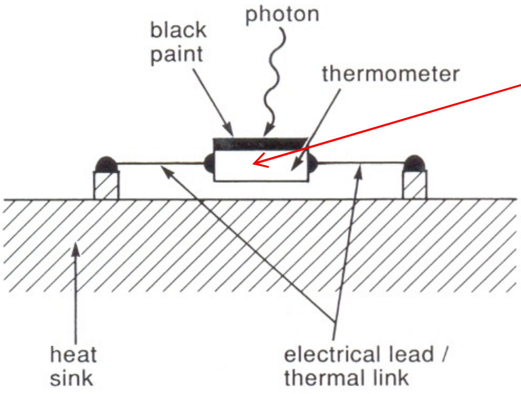
Basic Operation of Bolometers

Principle: the detector is connected via a thermal link with thermal conductance G to a heat sink of temperature T_0.



The total power absorbed by the detector is: $P_T(t) = GT_1 + C\frac{dT_1}{dt}$

- A high input impedance amplifier measures the voltage
- the voltage depends on resistance
- the resistance depends on temperature



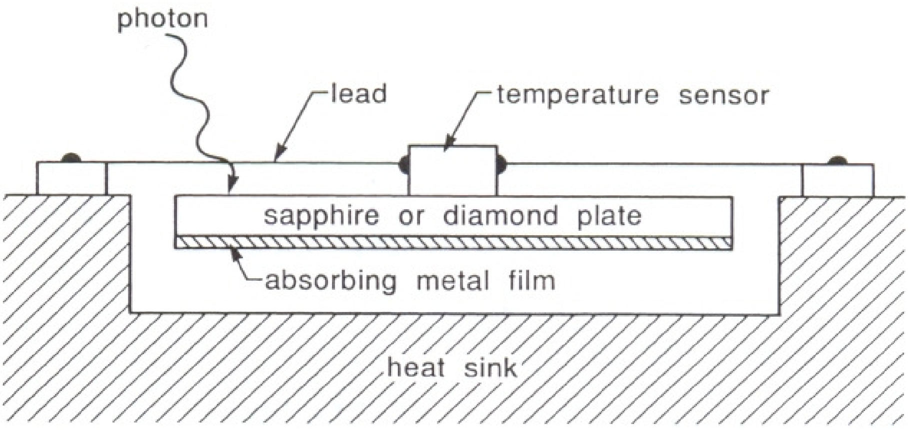
Chip of doped silicon or germanium

Variations of Bolometers

Composite Bolometers

Problem: low QE.
Partial solution: enhance absorption with black paint – but this will increase the heat capacity.
Better solution: composite bolometers, where:

- absorber has low heat capacity C and high QE
- sensor has low C but is well coupled to absorber

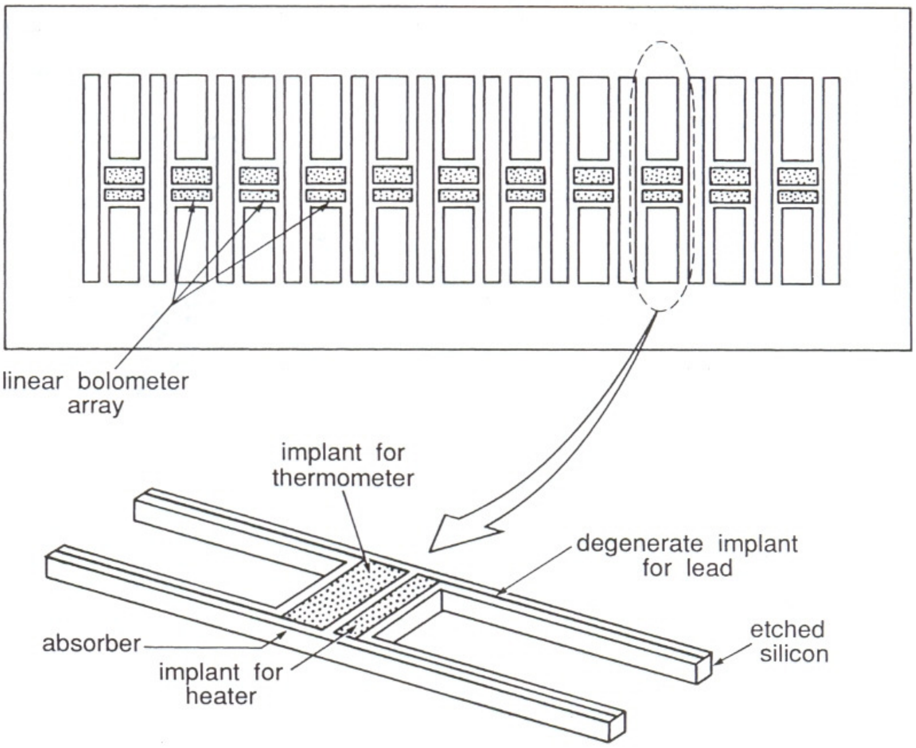


Here the heat capacity of the blackened sapphire plate is only 2% of that of Ge.

Etched Bolometers

Precision etching techniques in Si minimize the size of the structures →

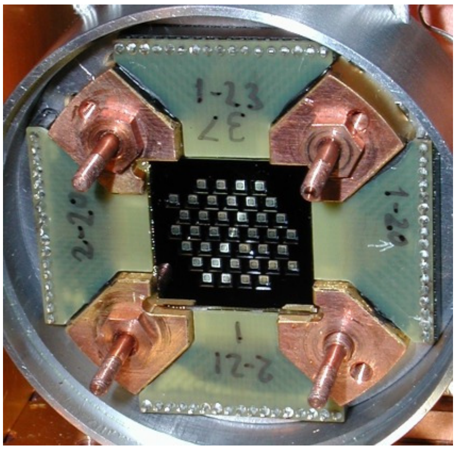
- low heat capacity C
- short thermal time response ~ C/G
- multiplexing advantage ("arrays")

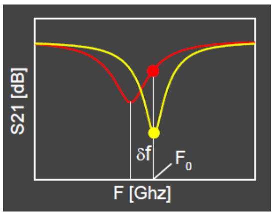


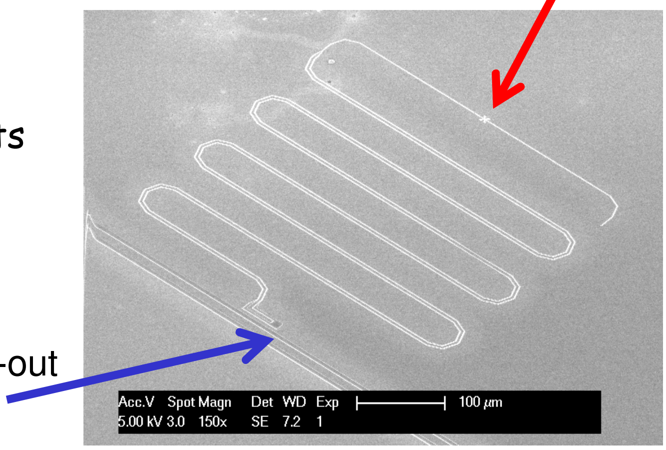
...and many more Types of Bolometers

Most of them utilize the properties of superconductivity. E.g.,

- Transition edge sensors (TES)
- Microwave Kinetic Inductance detectors (MKIDs) [Photons are absorbed in a superconductor, producing quasi-particle excitations, which change its kinetic inductance]







Part I
Photon Detectors
Part II
Thermal Detectors
Part III
Coherent Receivers

Principle of Coherent/ Heterodyne Receivers

Basic Idea

Problems:

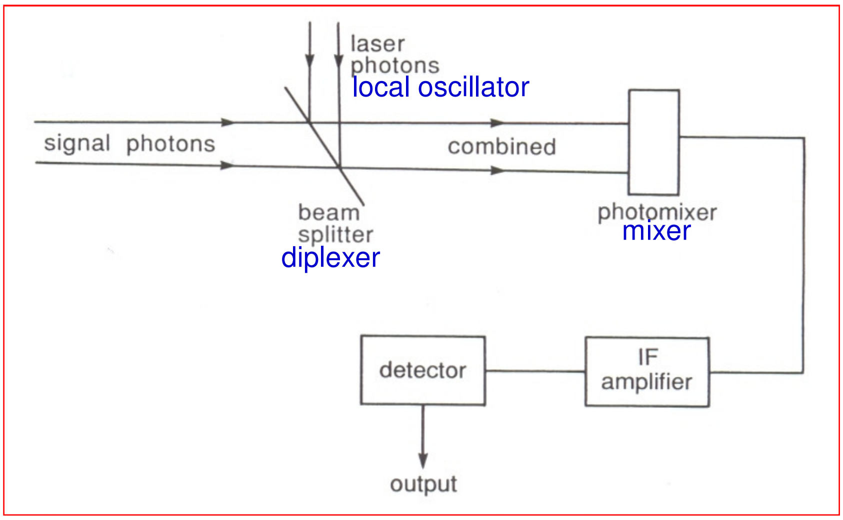
- very low photon energies → wave character of light is dominant
- often very weak signal → amplification essential

Solution:

- mix signal with reference wave

Advantages:

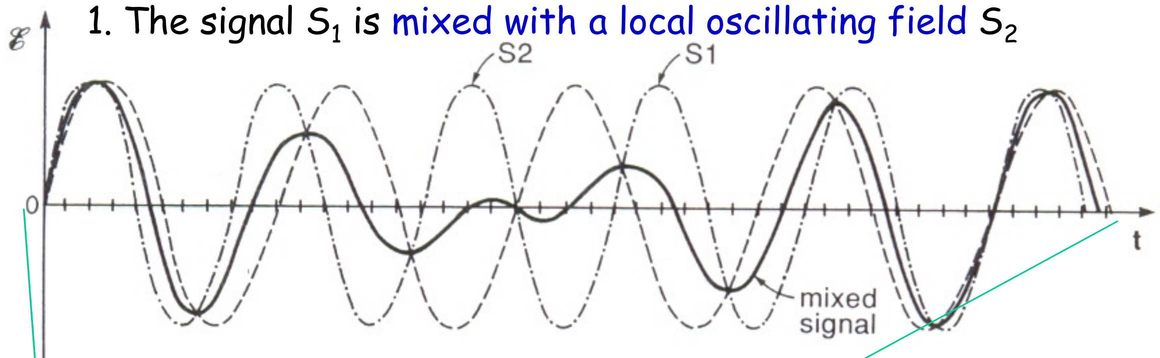
- encodes signal over a wide wavelength range → ideal for spectroscopy
- typically, power(ω_{LO}) » power(ω_S) → amplification by oscillator signal
- down-conversion to frequencies where low-noise electronics exist.



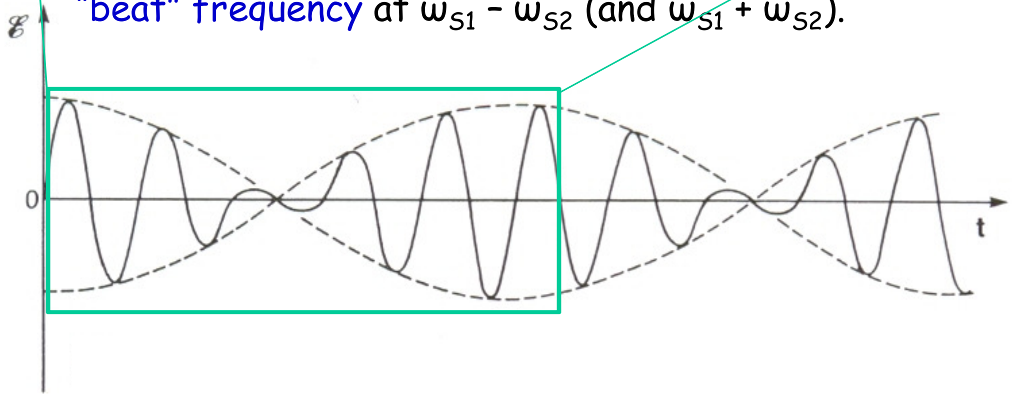
Step 1: Down-convert the Frequency

Note: $\lambda = 1\ \mu m$ ⇔ $\nu = 300$ THz ⇔ $\Delta t = 3.3 \cdot 10^{-15}$ s
$\lambda = 100\ \mu m$ ⇔ $\nu = 3$ THz ⇔ $\Delta t = 3.3 \cdot 10^{-13}$ s
$\lambda = 1$ cm ⇔ $\nu = 30$ GHz ⇔ $\Delta t = 33$ ps

1. The signal S_1 is mixed with a local oscillating field S_2



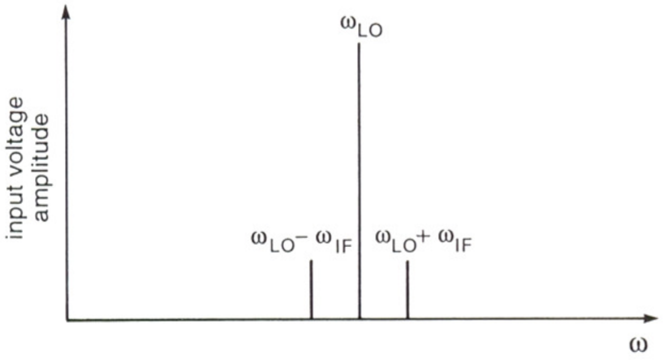
2. The mix produces a down-converted difference, intermediate, or "beat" frequency at $\omega_{S1} - \omega_{S2}$ (and $\omega_{S1} + \omega_{S2}$).



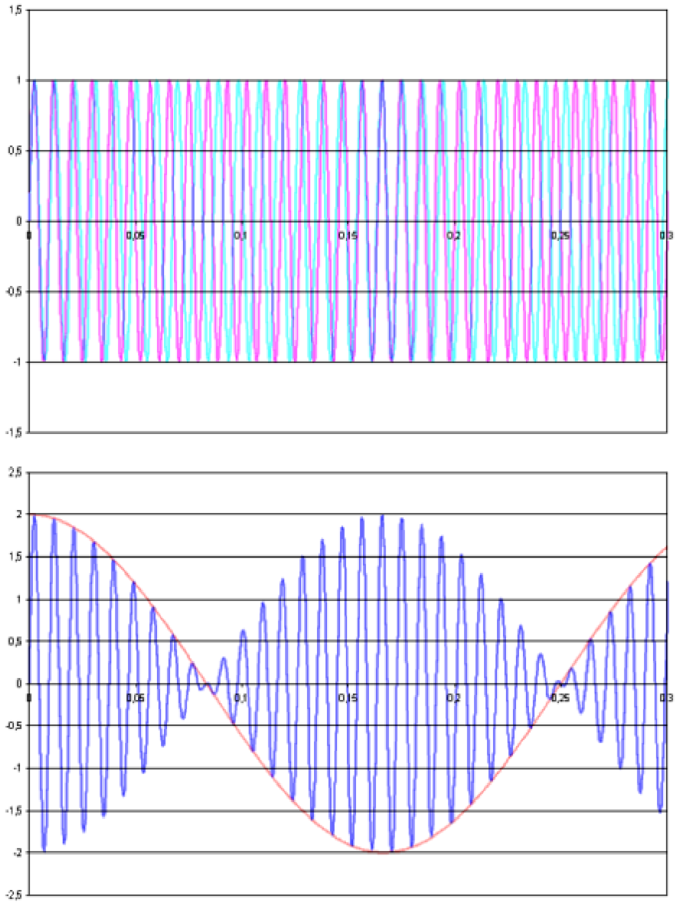
The Intermediate Frequency (IF)

The IF is the "beat frequency", the difference frequency between local oscillator and signal frequency.

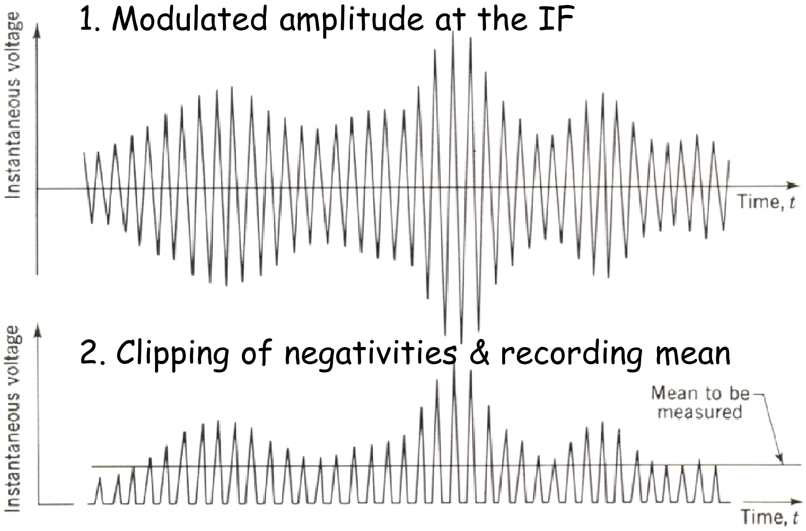
Example: Measure a signal at 1.5 GHz → use an oscillator at 1.55 GHz → down-convert the signal to a 50 MHz carrier.



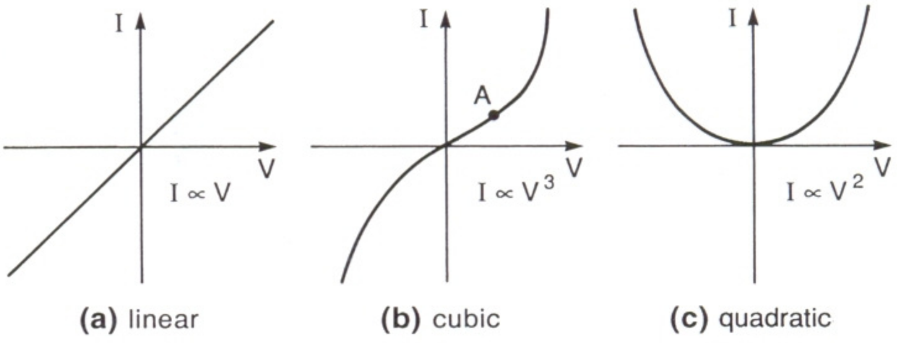
From the measured IF signal we cannot tell whether the signal frequency ω_S was lower or higher than ω_{LO} → assume two components of equal strength, one at $(\omega_{LO}+\omega_{IF})$ and one at $(\omega_{LO}-\omega_{IF})$



Step 2: Measure with a non-linear Device


1. Modulated amplitude at the IF

2. Clipping of negativities & recording mean



A linear device (a) yields no output power at any frequency.
➔Need a non-linear device (b,c) that converts power from the original frequencies to the beat frequency (= mixer)

Key Components of Heterodyne Receivers

Local Oscillators (LO)

The ALMA Central LO in the AOS Technical Building at 5000m. Photo: W. Grammer.

At low frequencies (radio/sub-mm) one uses an electronic LO (in combination with wire antenna or waveguide)

+ easily tuneable in frequency

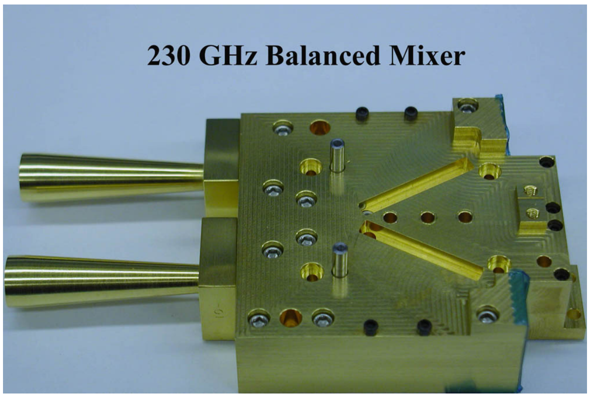
- low power at high frequencies (sub-mm)

At high frequencies (IR) the LO may be a continuous wave (CW) laser

+ high power

- discrete set of frequencies

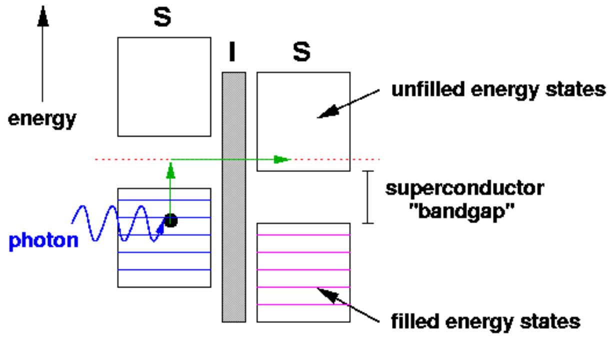
Mixer Technology


230 GHz Balanced Mixer

Problems:
- "pixel" size > λ for efficient absorption, but frequency response ∝ 1/size.
- good & fast photon detectors do not exist for λ > 40μm.

Material	Recombination time
Si	100 μs
Ge	10000 μs ←
PbS	20 μs
InSb	0.1 μs
GaAs	1 μs
InP	~1 μs

Solutions:
- SIS junctions →
- Hot electron bolometers
- Schottky diodes



Performance Characteristics of Heterodyne Receivers

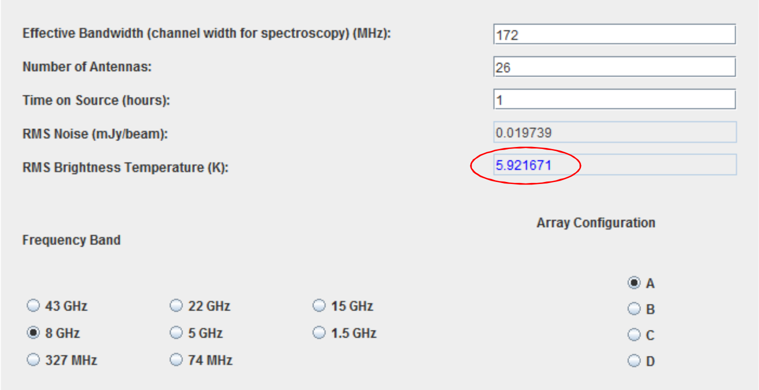
Performance Estimators

To describe the performance of heterodyne detectors we introduce:

1. The noise temperature T_N:
   It is defined such that a matched blackbody at the receiver input at a temperature T_N produces a S/N = 1.
   Obviously, the lower T_N the better the S/N.

2. The antenna temperature T_S:
   It is defined, analogous to the noise temperature, as the strength of the source flux (S/N=1).
   In the Rayleigh-Jeans approximation, the antenna temperature is linearly related to the input flux density: $P_S \sim T_S$

Example: VLA exposure time calculator:
`http://www.vla.nrao.edu/astro/guides/exposure/calc.html`



Effective Bandwidth (channel width for spectroscopy) (MHz): 172
Number of Antennas: 26
Time on Source (hours): 1
RMS Noise (mJy/beam): 0.019739
RMS Brightness Temperature (K): 5.921671

Frequency Band: ○ 43 GHz ○ 22 GHz ○ 15 GHz ● 8 GHz ○ 5 GHz ○ 1.5 GHz ○ 327 MHz ○ 74 MHz

Array Configuration: ● A ○ B ○ C ○ D

Performance Comparison

Case 1: Bolometer operating at BLIP and heterodyne receiver operating in the thermal limit (hν«kT)

➔ the bolometer will perform better

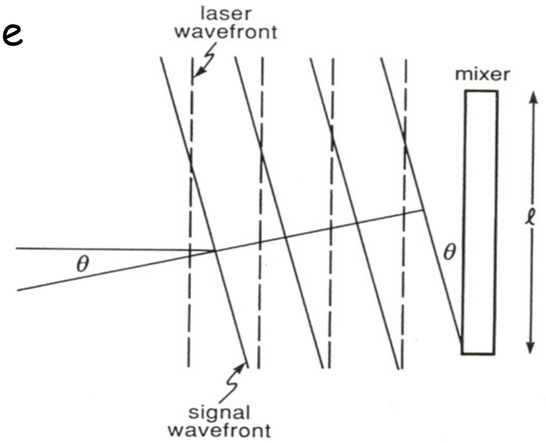
This is always true, except for measurements at high spectral resolution, much higher than the IF bandwidth.

Case 2: detector noise-limited bolometer and a heterodyne receiver operating at the quantum limit (hν»kT).

➔ the heterodyne receiver will outperform the bolometer.

In the case of narrow bandwidth and high spectral resolution the heterodyne system will always win.
If the spectral resolution $\nu/\Delta\nu$ is kept constant over the observed range the figure of merit goes as $1/\nu^2$.

Throughput of a Heterodyne System



Several factors limit the throughput of the system:

1. Only components of the signal electric field vector parallel to the reference field can interfere. Full cancellation occurs when offset ~ λ

$$l \sin\theta_{\max} = \lambda \approx l\theta_{\max}$$

2. The Etendue $A\times\Omega \sim \lambda^2$ sets a constraint on the coherent beam that can be accepted by a telescope of diameter D ➔ a coherent receiver should operate at the diffraction limit of the telescope.

3. Since the laser field is polarized only one polarization component of the source can interfere and produce a signal → heterodyne receivers = single-mode detectors