SIGNAL FORMATION AND SIGNAL PROCESSING IN DETECTORS

LECTURES AT THE UNIVERSITY OF FREIBURG

MARCH 9-14, 2020

LECTURE 2

NORBERT WERMES
UNIVERSITY OF BONN

Outline overview

UNIVERSITÄT BONN

1. What is a detector “signal”?

2. Charge transport in gases and solids

3. Induced signals on electrodes
   - Schottky-Ramo Theorem
   - Current, charge or voltage?
   - Applying Ramo to detectors
   - ➡ Structured electrodes
   - (calculation of E_W by “conformal mapping”)

4. Signal fluctuations and (electronic) noise
   - Why bother?
   - Signal fluctuations (Fano noise)
   - Electronic noise

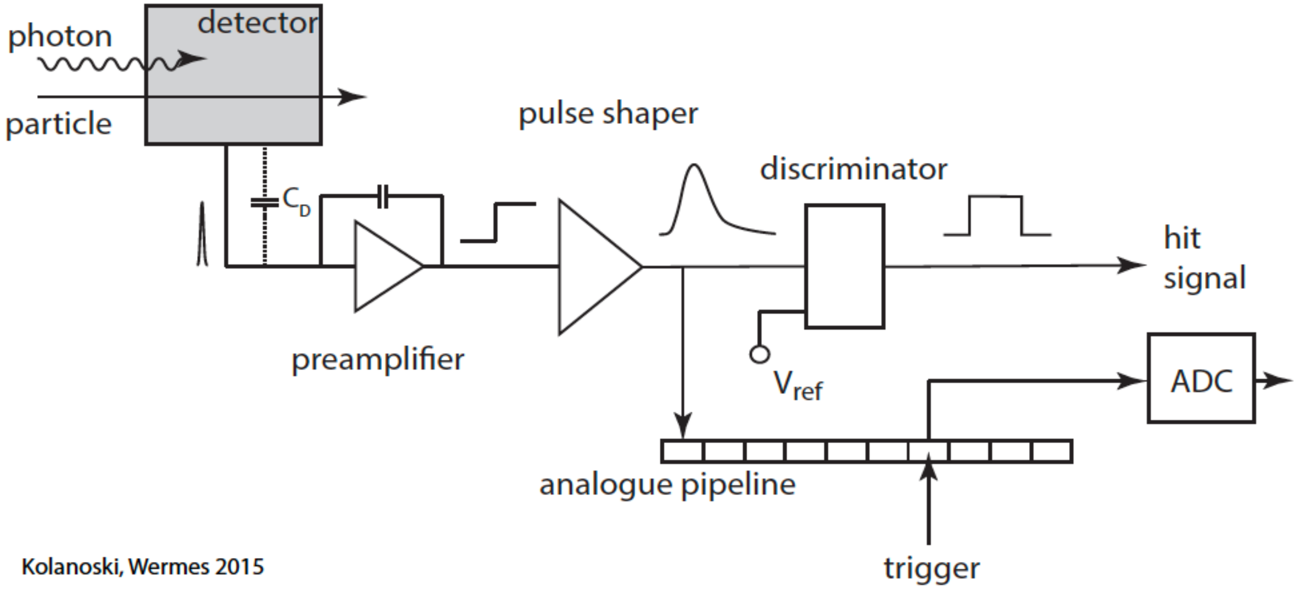
5. Readout of signals
   - Amplification
   - (Excursion: Laplace transform)
   - Filtering
   - Discrimination
   - Digitisation
   - (Example: a readout chip)

6. Signal transmission off detector

7. (Deadtime)

8. Noise of a readout system
   - Explicit calculation of noise
     - ATLAS pixel detector
     - ATLAS strip detector
     - ATLAS Liq. Argon calorimeter

Generic R/O scheme = guide line of what follows



Kolanoski, Wermes 2015

back end electronics

front end electronics

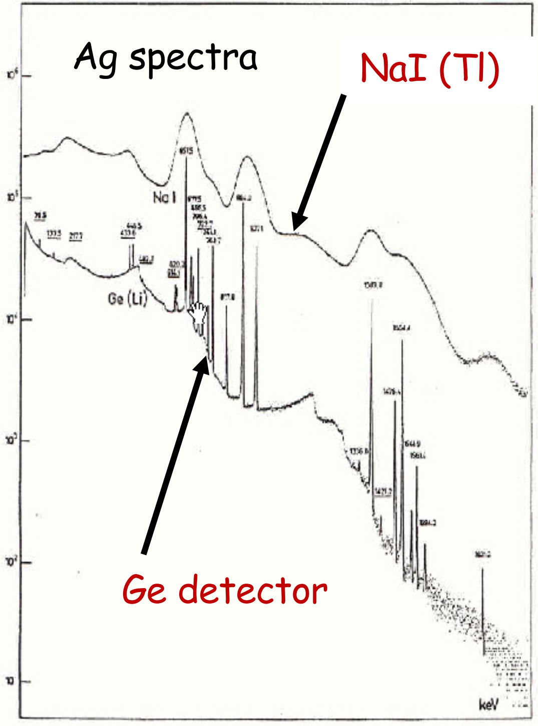
UNIVERSITÄT BONN

SI LAB Silizium Labor Bonn

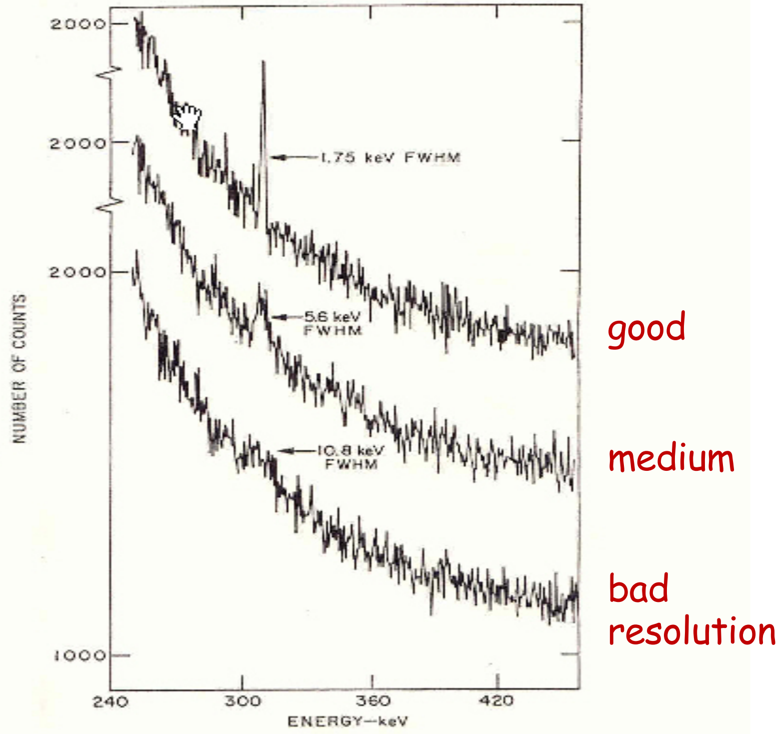
Signal fluctuations and (electronic) noise

Intro – why bother about noise?

Why bother about noise?



(J.Cl. Philippot, IEEE Trans. Nucl. Sci. NS-17/3 (1970) 446)

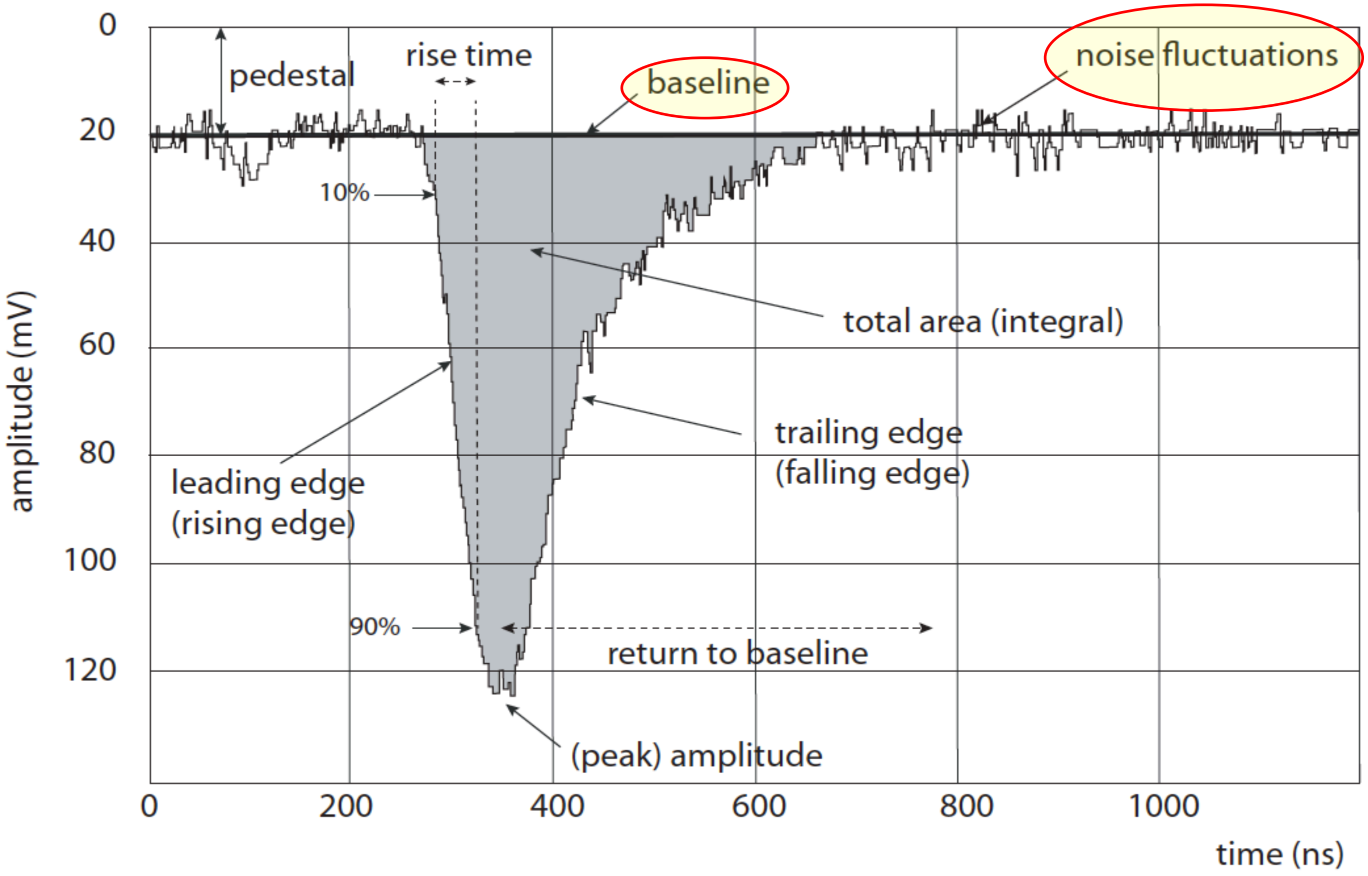


G.A. Armantrout *et al.*, IEEE Trans. Nucl. Sci. NS-19/1 (1972) 107

Low noise improves the resolution and the ability to distinguish (signal) structures.

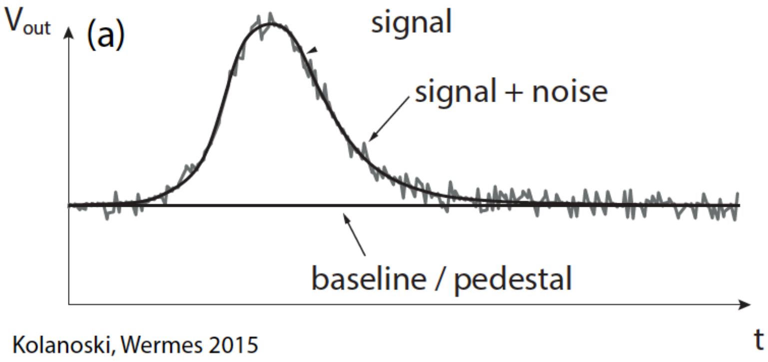
Low noise improves the signal to back-ground ratio (signal counts are in fewer bins and thus compete with fewer background counts).

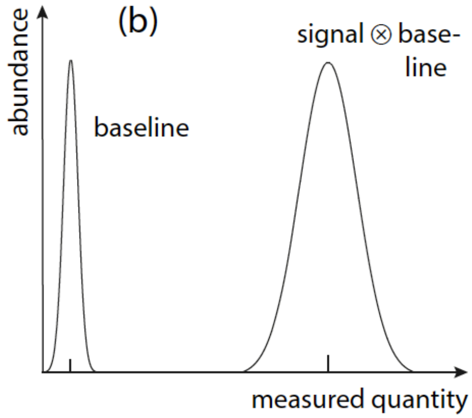
Nomenclature, a typical detector pulse



Kolanoski, Wermes 2017

Signal fluctuations and electronic noise





$\sigma_S \gg \sigma_B$

$\sigma_S \ll \sigma_B$

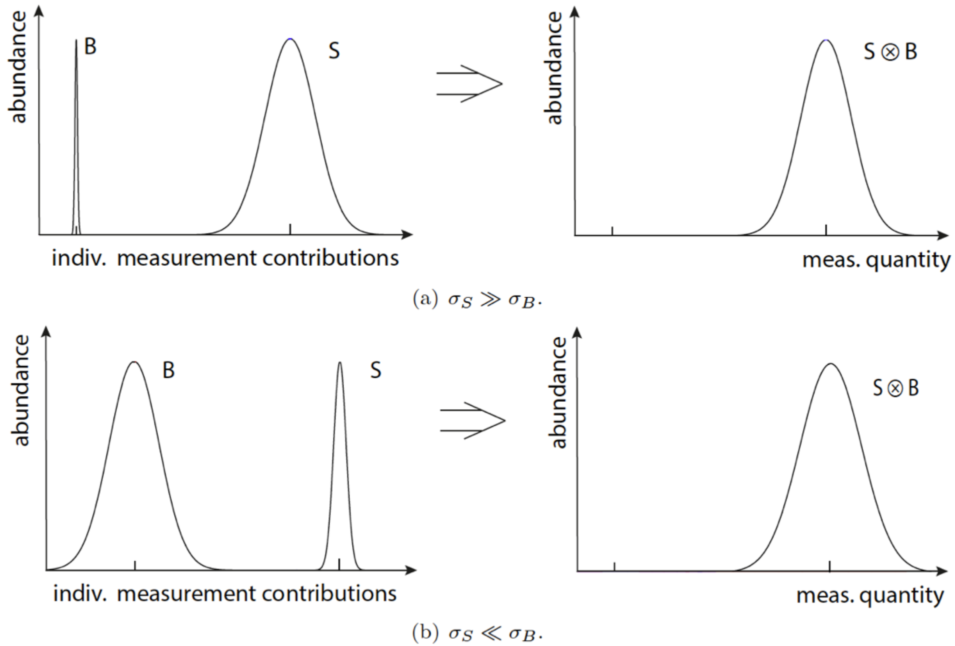


Fig. 17.57 Fluctuation contributions to a measurement spectrum for the cases (a) if $\sigma_S \gg \sigma_B$, as for example typical for NaI(Tl) scintillation detectors with PMT readout, and (b) if $\sigma_S \ll \sigma_B$, being typical for example for silicon microstrip detectors (with relatively large detector capacitance) with electronic readout. For both, on the left hand side the individual fluctuation contributions are shown and on the right hand side the measurements containing both contributions. The resolution is given by quadratic addition $S \otimes B$ of the widths of baseline (B) and signal (S) distributions, respectively. In (a) the resolution is dominated by the fluctuations of the signal process, whereas in (b) the electronic noise fluctuations are large in comparison to the signal fluctuations and hence contribute much to the total resolution

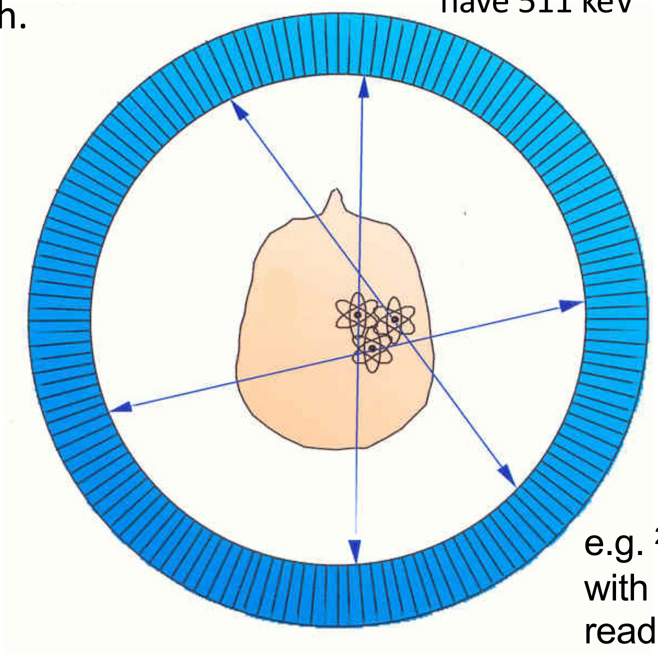
EXAMPLE I: Positron Emission Tomography (PET)

What is Position Emission Tomography (PET)?

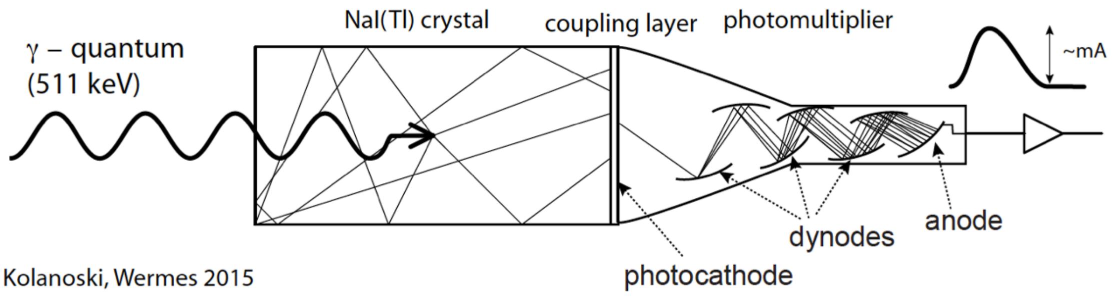
- Patient injected with drug having ß$^+$ emitting isotope (e.g. ^{22}Na)
- Drug localises in patient.
- Isotope decays, emitting e$^+$
- e$^+$ annihilates with e$^-$ from nearby tissue forming bach-to-back photon pair with E_γ = 511 keV each.
- 511 keV photons are detected via time coincidence
- Position of emission lies on line defined by detector pair

Result: a planar image of a “slice” through the patient

What is the dominant noise origin that must be fought when trying to improve the energy resolution?



UNIVERSITÄT BONN



- In a NaI(Tl) crystal about 43 000 scintillation photons are generated per MeV of deposited energy (see table 13.3 on page 500). A γ quantum of 511 keV hence generates $N_\gamma \approx 22\,000$ (optical) photons in a crystal.
- Of these about $N_\gamma \approx 13\,000$ (60%) reach the photocathode of the photomultiplier.
- Assuming a quantum efficiency of 20% we then have about $N_e \approx 2\,500$ photoelectrons that exit the photocathode of the PMT.
- For a typical PMT gain of 10^6 about $N_e = 2.5 \times 10^9$ electrons reach the anode leading to a current signal in the mA range assuming a typical pulse duration of NaI(Tl) detectors.

Where is the dominant noise source?

Example cont'd

The achievable energy resolution σ_E/E is determined by the smallest number in this chain, that is the one contributing with the largest statistical uncertainty. The other quantities contribute much less to the total uncertainty due to quadratic error propagation. In this case it is the number of photons at the first dynode that are dominating the error:

$$\sigma_{1.\,\mathrm{dyn}} = \sqrt{N_{1.\,\mathrm{dyn}}} = \sqrt{2500}\,e^- = 50\,e^- = 2\%\ N_{1.\,\mathrm{dyn}}$$

$$\Rightarrow \qquad \sigma_{\mathrm{signal}} = \sigma_{1.\,\mathrm{dyn}} \times (\mathrm{amplification}) = 5 \times 10^7 e^- \,.$$

The statistical fluctuations of the signal (or what is left over after the conversion to photoelectrons) are much larger in this case than the electronic noise responsible for the width of the baseline (B). For PMTs this typically is in the range of $10^{-(4-5)}$ times the signal [472] and is hence negligible here for the resolution of the detection system.

Example II

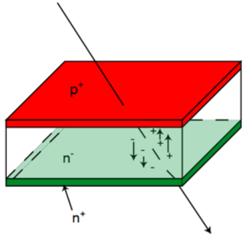
For the inverse case $\sigma(S) \ll \sigma(B)$ we take as an example the detection of a 60 keV-γ ray from a ^{241}Am radioactive source detected with a Si semiconductor detector. The number of electron–hole pairs is

$$N_{e/h} = \frac{E_\gamma}{\omega_i} = \frac{60\,000\,\mathrm{eV}}{3.65\,\mathrm{eV}/(\mathrm{e/h})} \approx 16\,500\ \ \mathrm{e/h\ pairs}\,.$$

The signal fluctuations are computed as:

$$\sigma_{\mathrm{signal}} = \sqrt{N_{e/h} \cdot F}$$
$$= \sqrt{N_{e/h} \cdot 0.1} = 40\, e^- \,.$$

F = fano factor



Relative to signal $\dfrac{\sigma_{e/h}}{N_{e/h}} = 0.2\%$ compare to 2% for NaI(Tl)

Noise in Si detectors ($\propto C_D$) is of order 100 e- (pixels) or 1000 e- (strips) i.e. larger than the signal fluctuation.
They thus determine the resolution of the system.

Signal fluctuations and Fano Factor

Fano Factor I

The exact general computation of the Fano factor is complicated. We consider here as a simple system a silicon detector for which the discussion can essentially be reduced to two energy loss mechanisms (see also [863]): the generation of electron–hole pairs and lattice excitations (phonon excitations). For the creation of an e/h pair at least the bandgap energy (in silicon $\Delta E_{gap} = 1.1\,\text{eV}$) is needed. For every event, however, the deposited energy is subdivided differently for the generation of e/h pairs or for lattice excitations such that on average the energy of $w_i = 3.65\,\text{eV}$ is needed to create one e/h pair.

Let us assume that N_p phonon excitations are created with a statistical fluctuation of $\sigma_p = \sqrt{N_p}$; $N_{e/h}$ electron–hole pairs are generated in ionisation processes with a fluctuation of $\sigma_{e/h} = \sqrt{N_{e/h}}$. A fixed energy E_0 shall be deposited with every event in the detector (for example the energy of an X-ray or γ quantum from a radioactive source) which is available for the creation of phonons and e/h pairs:

$$E_0 = E_i \cdot N_{e/h} + E_x \cdot N_p \qquad (17.82)$$

where E_i is the energy necessary for one individual ionisation and E_x the energy necessary for one individual phonon excitation.

simplified approach

only two energy loss mechanisms

1) e/h ionisation
2) phonon exc.

assume also Gaussian statistics

Fano Factor II

The energy E_0 can split arbitrarily between ionisations or lattice excitations. Since E_0 is fixed, however, when complete absorption holds, and is hence the same for every absorbed quantum, then in every absorption a fluctuation of a larger E_0 portion $(E_0 \cdot \Delta N_p)$ in phonon excitation must be compensated by a correspondingly smaller E_0 portion $(E_0 \cdot (-\Delta N_{e/h}))$ for ionisation:

$$E_x \cdot \Delta N_p - E_i \cdot \Delta N_{e/h} = 0\,.$$

Averaged over many absoptions of the energy E_0 therefore yields:

$$E_x \cdot \sigma_p = E_i \cdot \sigma_{e/h}$$

$$\Rightarrow \quad \sigma_{e/h} = \sigma_p \frac{E_x}{E_i} = \sqrt{N_p}\frac{E_x}{E_i}\,.$$

$$N_p = \frac{E_0 - E_i N_{e/h}}{E_x}$$

and with the average number of e/h pairs, $N_{e/h} = \frac{E_0}{\omega_i}$

$$\sigma_{e/h} = \frac{E_x}{E_i}\sqrt{\frac{E_0}{E_x} - \frac{E_i}{E_x}\frac{E_0}{\omega_i}} = \sqrt{\frac{E_0}{\omega_i}} \cdot \sqrt{\underbrace{\frac{E_x}{E_i}\left(\frac{\omega_i}{E_i} - 1\right)}_{=F\ \text{(Fano factor)}}} = \sqrt{\frac{E_0}{w_i}} \cdot F$$

$$\Longrightarrow \sigma_{e/h} = \sqrt{N_{e/h} \cdot F}\,.$$

Note: E_0 is fixed

Result:
Energy resolution $\sigma_{e/h}$ is in fact smaller than $\sqrt{N_{e/h}}$, since F often (e.g. in Si) < 1

Material	Si	Ge	GaAs	CdTe	diamond	Ar	liq. Ar
Fano factor	0.115	0.13	0.10	0.10	0.08	0.20	0.107–0.116

For detectors with more and also more complex signal generation processes, as for example scintillators, for which exciton processes also play a role (see section 13.3 on page 515), Fano factors larger than one ($F > 1$) can even occur [360].

UNIVERSITÄT BONN

Electronic noise

Noise

Stromrauschen

shot noise

Funkel Rauschen

series noise

Widerstandsrauschen

white noise

switching noise

Schrotrauschen

popcorn noise

flicker noise

Nyquist Noise

parallel noise

Johnson Noise

kt/C noise

Thermal noise

RTS noise

1/f noise

noise current origins

a signal current
$$i = \frac{Nev}{d}$$

can fluctuate in number and in velocity

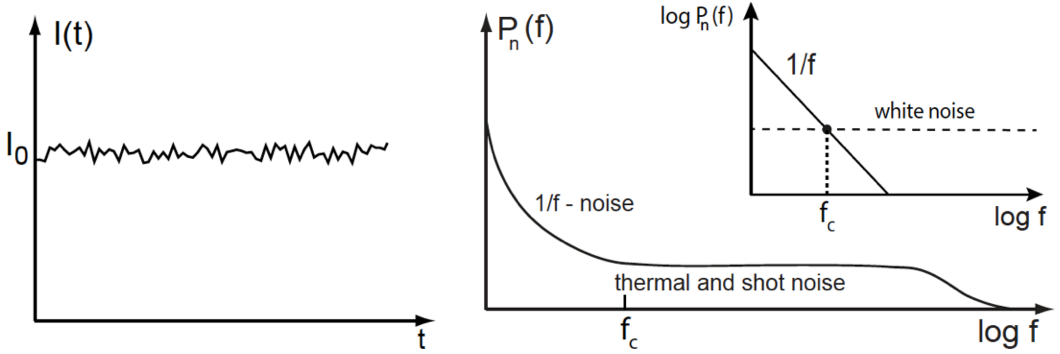
$$(di)^2 = \left(\frac{ev}{d} dN\right)^2 + \left(\frac{eN}{d} dv\right)^2$$

Summary: noise physical origins

- $\langle i^2 \rangle = 2q \langle i \rangle \, df$ shot noise
  number fluctuation
- $\langle i^2 \rangle = 4kT / R \, df$ thermal noise
  velocity fluctuation
- $\langle i^2 \rangle = \text{const.} \; 1/f^{\alpha} \, df$ 1/f noise
  number fluctuation

plus perhaps … popcorn noise, RTS – noise … all $1/f^{\alpha}$ …

Measuring Noise



(a) Current noise as a function of time.

(b) Spectral noise density (schematic) as a function of frequency.

$$\sigma^2 = \lim_{T\to\infty} \frac{1}{T} \int_0^T (I(t) - I_0)^2 dt$$

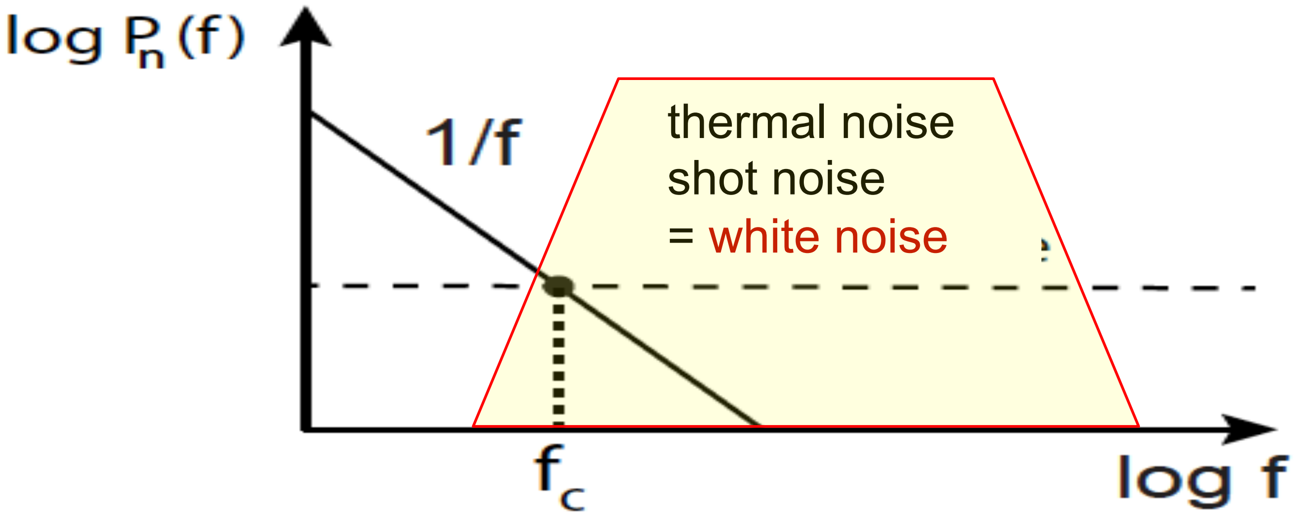
$$dP_n/df = \frac{1}{R} d\langle v^2\rangle/df = R\, d\langle i^2\rangle/df \quad \Rightarrow \quad P_n = \int_0^\infty \frac{dP_n}{df} df$$

spectral noise power density

unit of “voltage noise power density“ = $[\sqrt{v^2/df}] = V/\sqrt{Hz}$

unit of “current noise power density” = $[\sqrt{i^2/df}] = A/\sqrt{Hz}$

Notice …



filtering, i.e. limiting the BW by high- (CR) and low-pass (RC) filters

- reduces the noise
- but: makes the response more slow

thermal noise I

derivation from first principles

1. thermal velocity distribution of carriers
   => time (or frequency) dependence of induced current → a difficult derivation
2. application of Planck's law for black body radiation ("hides" a bit the physics behind
   a general result of statistical mechanics)

   => gives the spectral density of the radiated power
   i.e. the power that can be extracted in thermal equilibrium

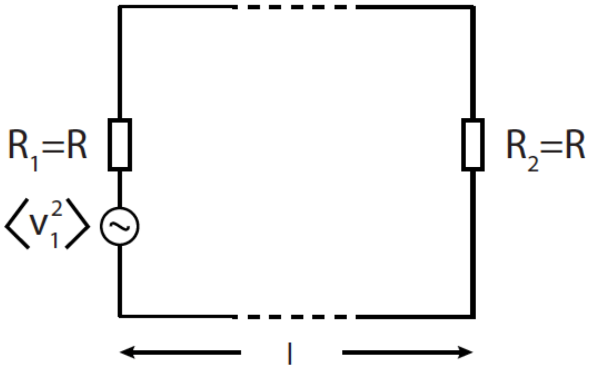
$$\frac{dP}{d\nu} = \frac{h\nu}{exp\left(\frac{h\nu}{kT}\right) - 1} \quad \rightarrow \; (\text{for } h\nu \ll kT) \quad = \frac{h\nu}{1 + \frac{h\nu}{kT} - 1} = kT$$

i.e. at sufficiently low frequencies (< THz)
is P independent of ν and is always the same amount in a bandwidth interval $\Delta\nu$

$P = kT\,\Delta\nu \;\rightarrow\; kT\;\Delta f$ (*)

thermal noise II

consider now an open resistor R_1 which generates a (quadratic) noise voltage $<v_1^2>$



The noise voltage $\langle v_1^2 \rangle$ over R_1 yields a noise power in R_2 when both resistors are short-circuited, where v is the voltage over R_2 caused by $\langle v_1^2 \rangle$.

$$P_{1\to 2} = \frac{v^2}{R_2} = \frac{\langle v_1^2 \rangle}{R_2}\left(\frac{R_2}{R_1+R_2}\right)^2 = \frac{\langle v_1^2 \rangle}{4R}$$

In thermal equilibrium R_2 transfers the same noise power to R_1

$$P_{1\to 2} = P_{2\to 1}$$

for every frequency portion of the noise fluctuation. The power spectrum hence is a function of f, of R, and of the temperature T.

thermal noise III

with (*) P = kT df …. we get

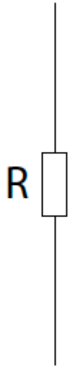
$$d\langle v_n^2 \rangle = 4kTR\,df$$

and

$$d\langle i_n^2 \rangle = d\frac{\langle v_n^2 \rangle}{R^2} = \frac{4kT}{R} df$$

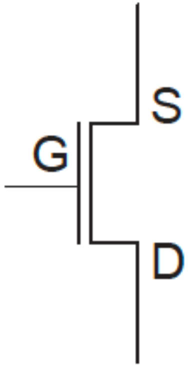
Ohm's law usually relates ⟨i^2⟩ and ⟨v^2⟩

Note: Thermal noise is always there (if T>0). It does not need power.



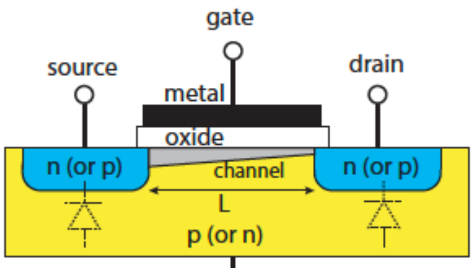
$$d\langle v_n^2\rangle = 4kTR\,df$$

$$d\left\langle i_n^2\right\rangle = \frac{4kT}{R}df$$



$$d\left\langle v_n^2\right\rangle = 4kT\frac{1}{\frac{2}{3}g_m}df$$

$$d\left\langle i_n^2\right\rangle = 4kT\frac{3}{2}g_m df$$

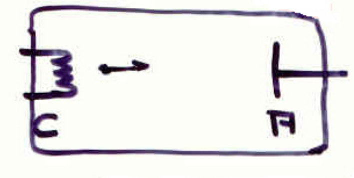


$g_m = dI_D/dV_{GS}$,
transconductance

adjustment

origin: excess electron injection into a device (over a barrier, i.e. NOT in a resistor)

e.g.



in a semiconductor: e/h in depletion zone induce current pulse until recombination

- the current pulses can be regarded as δ - functions, i.e. all frequencies contribute => white noise

$$\int_{-\infty}^{\infty} i_e(t)dt = e$$ ➔ $di_e/df = e \cdot 2$ (convention: $0 < f < \infty$ → $-\infty < f < \infty$)

- for infinitely narrow df the spectral component contributing is one sine wave with mean = 0 and rms = 1/√2

$$=> \frac{di_{e,k}}{df} = \frac{2e}{\sqrt{2}} = \sqrt{2}e$$

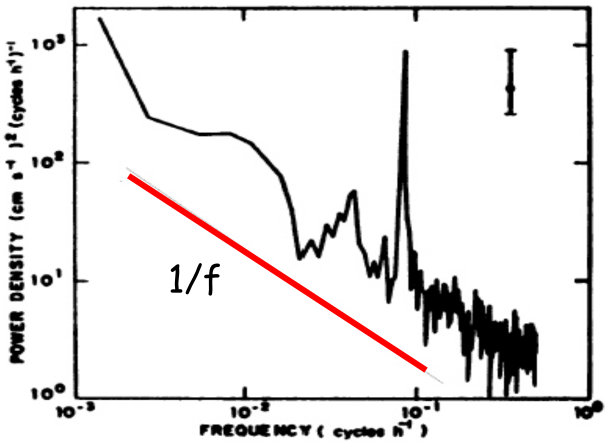
- for N electrons of total average current I = Ne/t = Ne Δf we get

$$\boxed{\langle i^2 \rangle} = \sum_{k=1}^{N} \left(\frac{d_{i,k}}{df} \right)^2 (df)^2 = 2Ne^2(df)^2 = 2e \underbrace{(Nedf)}_{\langle i \rangle} df = \boxed{2e\langle i \rangle df}$$

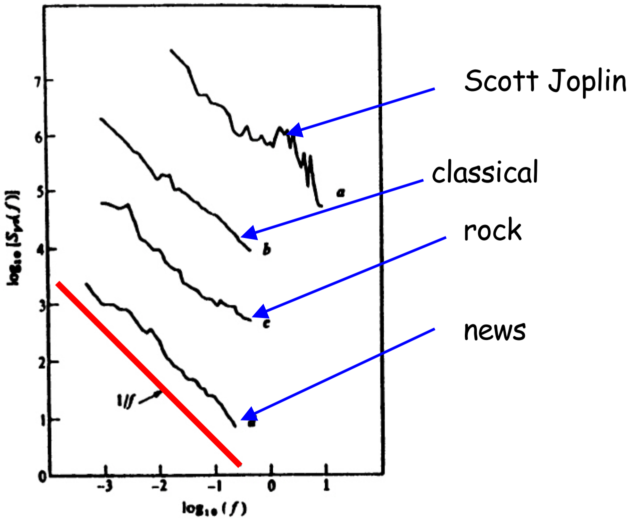
1/f noise - I

physics origin:

- superposition of relaxation processes with different time constants
- appears in many systems (ocean current velocity, music, broad casting, earthquake frequency spectra)
- many papers in literature (all you ever wanted to know) http://www.nslij-genetics.org/wli/1fnoise/



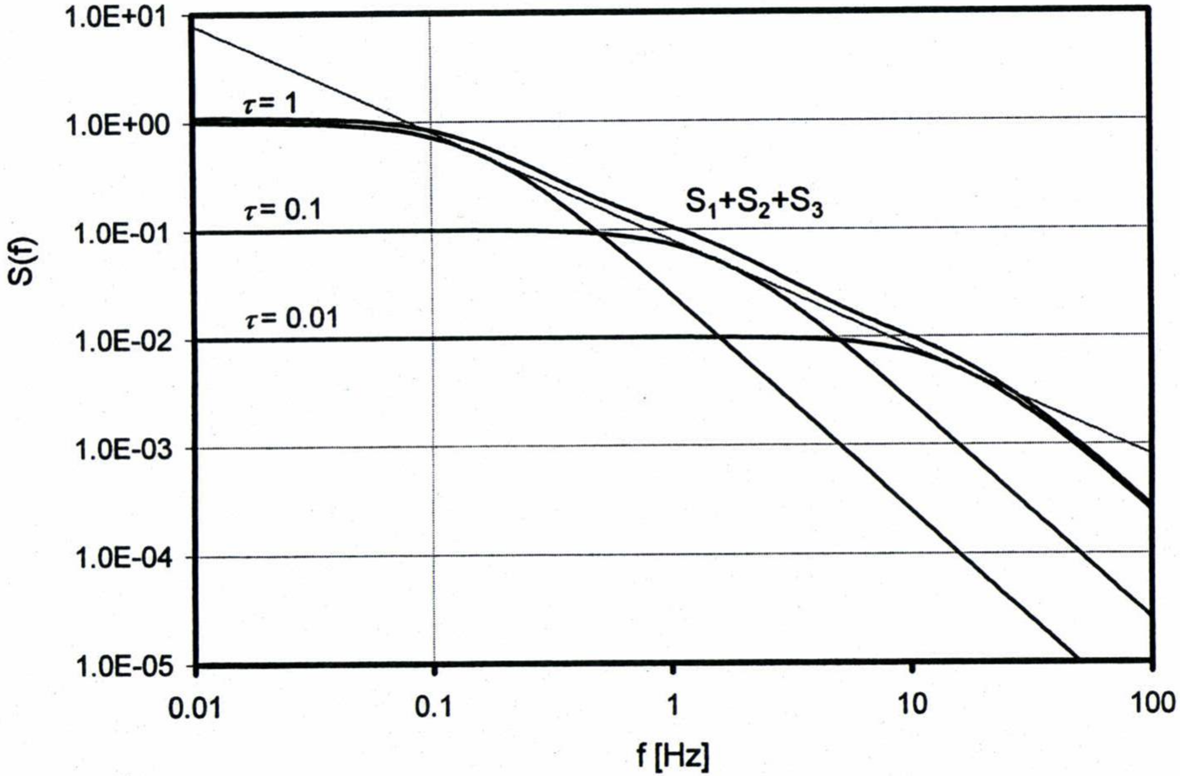
east-west component of ocean current velocity



loudness fluctuations spectra of radio broadcasting

1/f noise - II

superposition of $1/f^2$ spectra with 3 time constants



Assume a trapping site with relaxation time constant τ which releases electrons according to

$$N(t) = N_0\, e^{-\frac{t}{\tau}} \quad \text{for} \quad t \geq 0$$
$$N(t) = 0 \quad \text{for} \quad t < 0$$

Fourier transforming this into the frequency domain yields

$$F(\omega) = \int_{-\infty}^{\infty} N(t) e^{i\omega t} dt = N_0 \int_0^{\infty} e^{(\frac{1}{\tau}+i\omega)t} dt = N_0 \frac{1}{\frac{1}{\tau}+i\omega}$$

For a whole sequence of such relaxation processes occurring at different times t_k

$$N(t, t_k) = N_0 e^{-\frac{t-t_k}{\tau}} \; ; \quad N(t, t_k) = 0$$

but still with the same trapping time constants τ, one gets

$$F(\omega) = \int_{-\infty}^{\infty} \sum_k N(t, t_k) e^{i\omega t} dt = N_0 \sum_k e^{i\omega t_k} \int_0^{\infty} e^{(\frac{1}{\tau}+i\omega)t} dt = \frac{N_0}{\frac{1}{\tau}+i\omega} \sum_k e^{i\omega t_k}$$

1/f noise - IV

The power spectrum then is

$$P(\omega) \;=\; \lim_{T\to\infty}\frac{1}{T}\langle|F(\omega)|^2\rangle = \frac{N_0^2}{(\frac{1}{\tau})^2+\omega^2}\lim_{T\to\infty}\frac{1}{T}\langle|\sum_k e^{i\omega t_k}|^2\rangle = \frac{N_0^2}{(\frac{1}{\tau})^2+\omega^2}\cdot n$$

where n is the average rate of trapping/relaxation processes

If we in addition assume that the relaxation time constants are different
i.e. $\boldsymbol{\tau} \rightarrow \boldsymbol{\tau_i}$ and we integrate/sum over uniformly distributed $\boldsymbol{\tau_1 < \tau_i < \tau_2}$, we find

$$P(\omega) \;=\; \frac{1}{\frac{1}{\tau_1}-\frac{1}{\tau_2}}\int_{\frac{1}{\tau_2}}^{\frac{1}{\tau_1}}\frac{N_0^2\,n}{(\frac{1}{\tau})^2+\omega^2}d(1/\tau) = \frac{N_0^2\,n}{\omega(\frac{1}{\tau_1}-\frac{1}{\tau_2})}[arctan\frac{1}{\omega\tau_1} - arctan\frac{1}{\omega\tau_2}]$$

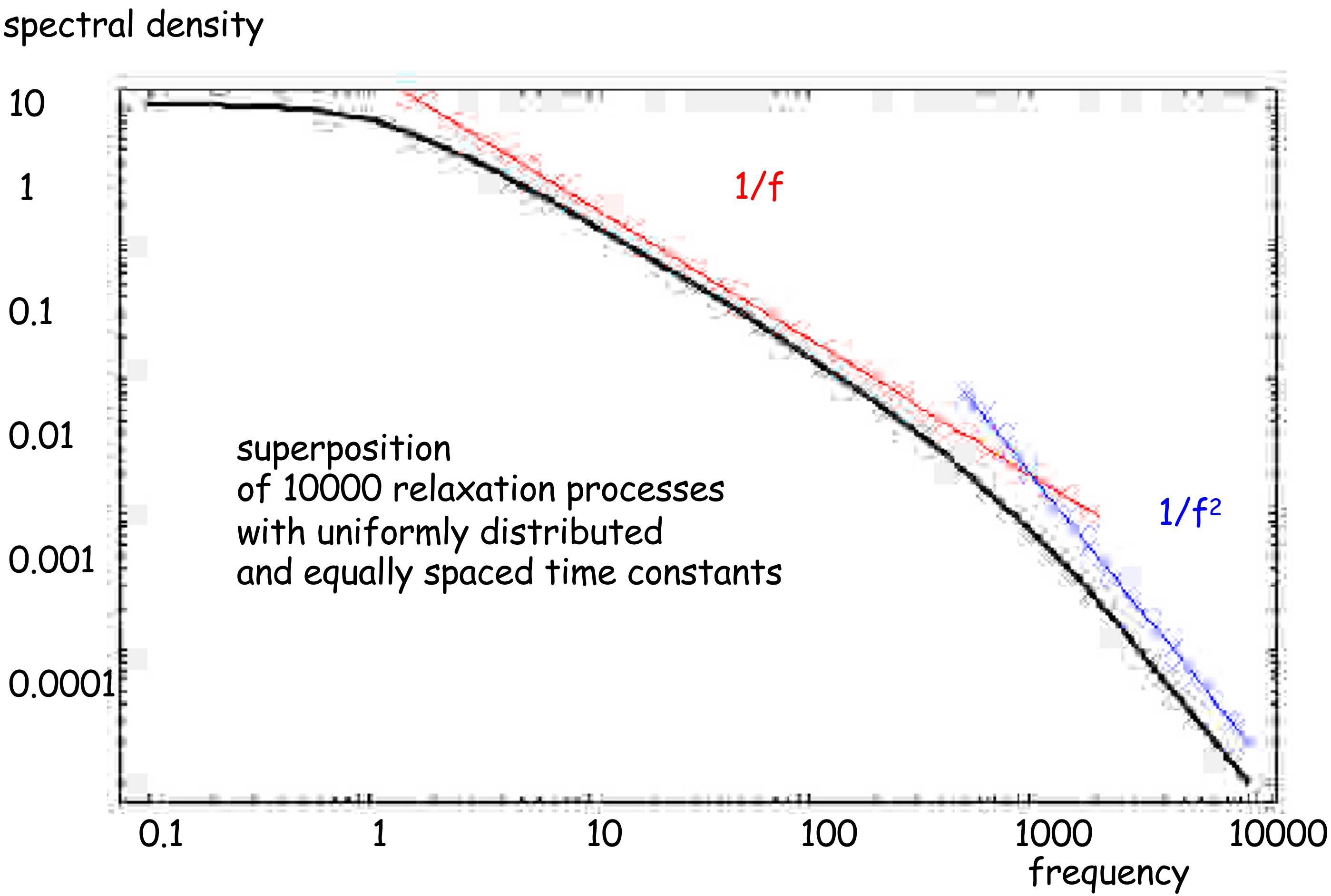
$$\approx \begin{cases} N_0^2 n & \text{if} \quad 0<\omega\ll\frac{1}{\tau_1},\frac{1}{\tau_2} \\ \frac{N_0^2 n\pi}{2\omega(\frac{1}{\tau_1}-\frac{1}{\tau_2})} & \text{if} \quad \frac{1}{\tau_2}\ll\omega\ll\frac{1}{\tau_1} \\ \frac{N_0^2 n}{\omega^2} & \text{if} \quad \frac{1}{\tau_1},\frac{1}{\tau_2}\ll\omega \end{cases}$$

const

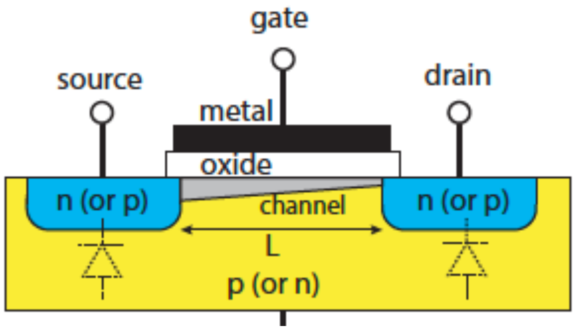
1/f

$1/f^2$

1/f noise - V


spectral density
10
1
0.1
0.01
0.001
0.0001
1/f
1/f²
superposition
of 10000 relaxation processes
with uniformly distributed
and equally spaced time constants
0.1
1
10
100
1000
10000
frequency

1/f noise in a MOSFET



origin

- trapping and release of channel charges in gate oxide
- depends on gate area A = W × L

$$\frac{d\left\langle v_{1/f}^2\right\rangle}{df} = K_f \frac{1}{C'_{ox} W L} \frac{1}{f}$$

empirical parametrisation (e.g. PSPICE)

$$C'_{ox} = \frac{3}{2}\frac{C_{GS}}{WL} \approx \epsilon_0\epsilon/d$$

$K_f^{NMOS} \approx 30 \times 10^{-25}$ J, $K_f^{PMOS} \approx 0.05\text{-}0.1 \times K_f^{NMOS}$

Fazit: Only three physical noise origins to consider …

- $\langle i^2 \rangle = 4kT / R \, df$ — thermal noise

  thermal fluctuations (Brownian motion)
  velocity fluctuation

  (in resistors, transistor channels)

- $\langle i^2 \rangle = 2q \langle i \rangle \, df$ — shot noise

  fluctuations in hopping over a barrier (shot)
  number fluctuation

  (where currents due to barrier crossings appear, e.g. in diodes, NOT in resistors)

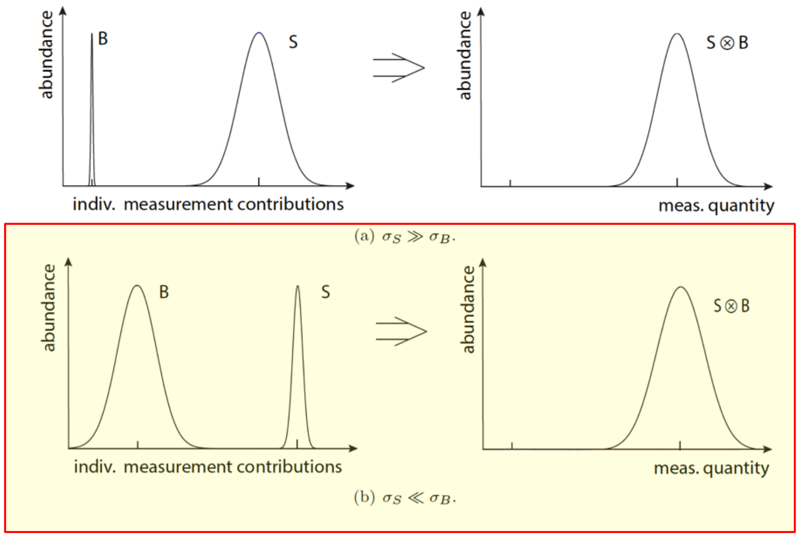
- $\langle i^2 \rangle = \text{const. } 1/f^{\alpha} \, df$ — 1/f noise

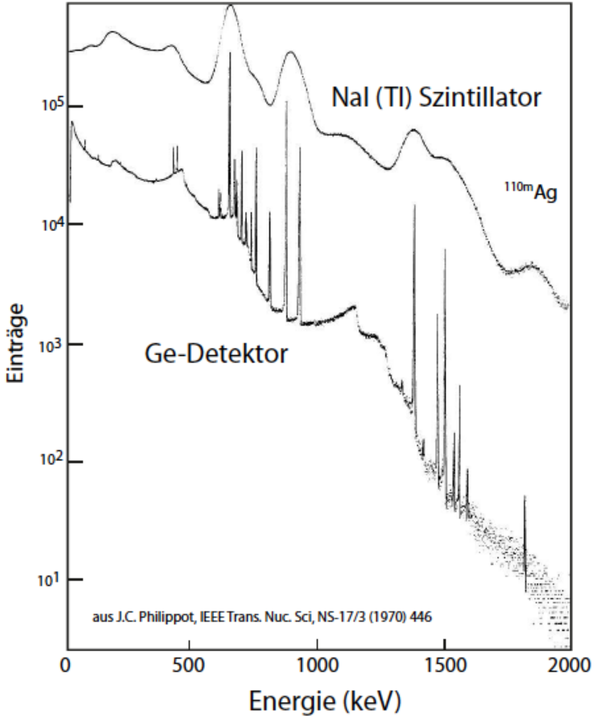
  trap/release fluctuations of carriers
  number fluctuation

  (whenever trapping occurs, e.g. in (MOS) transistor channels)

When to care about noise …

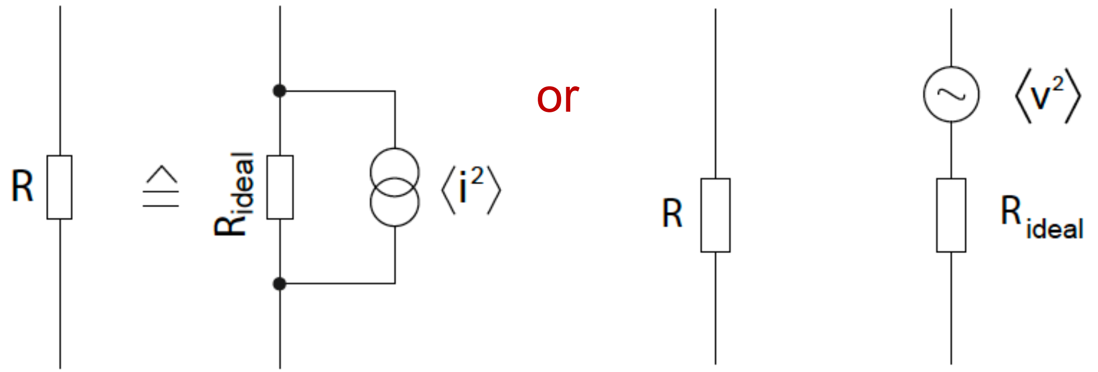
- always …
- but particularly, when situation is like this





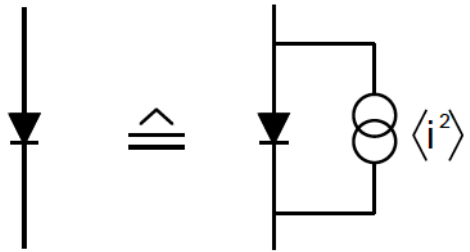
- even if you are not interested in an energy measurement, remember … thresholds

Noisy circuit elements

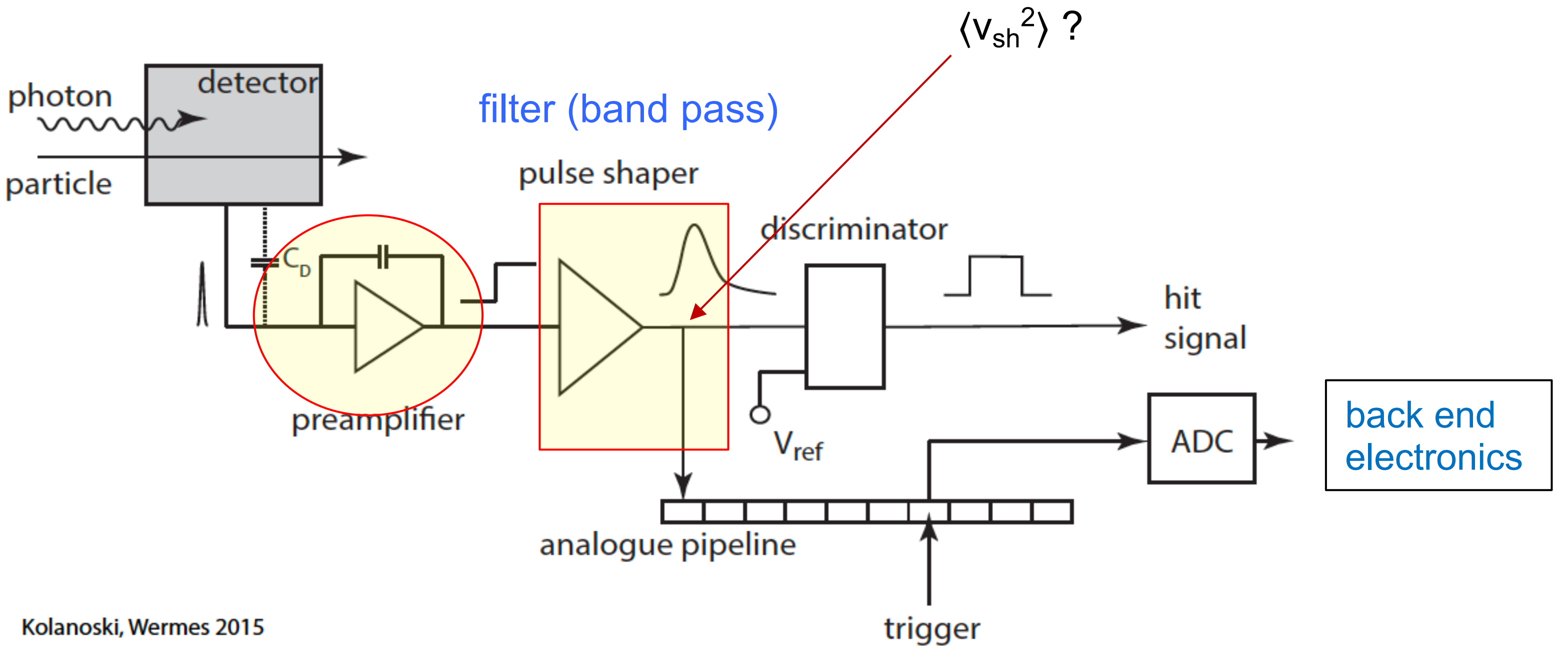


(a) Replacement circuit with parallel current noise source.

(b) Replacement circuit with serial voltage noise source.

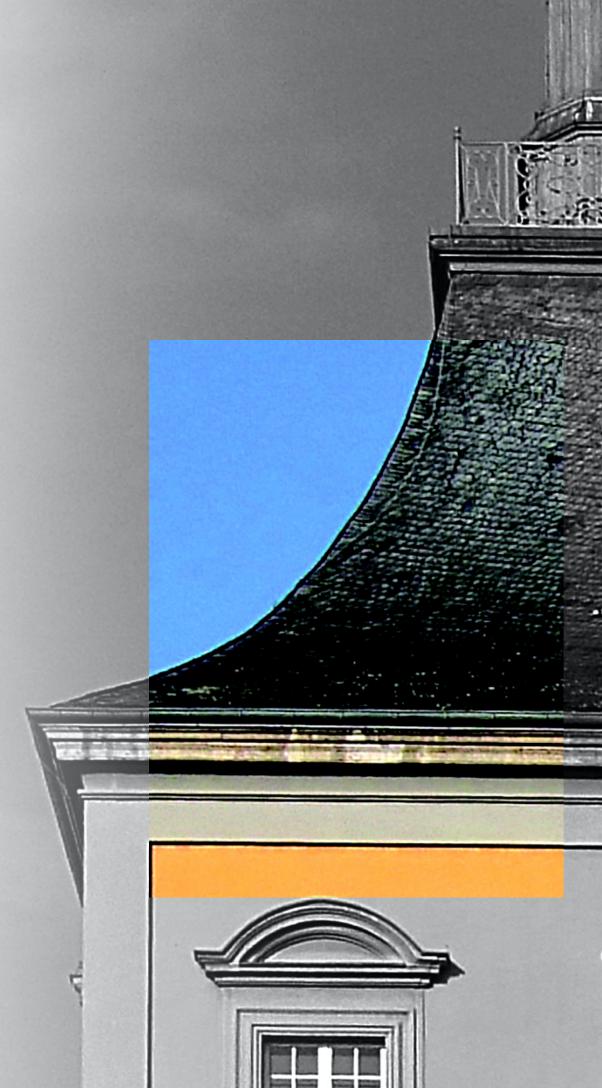


Generic R/O scheme: the dominant noise components



front end electronics

Readout of signals
- a typical readout chain



The amplifier

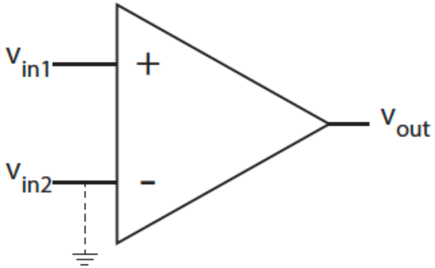
(often more specific:
“first amplifier” or “preamplifier”)

OpAmps

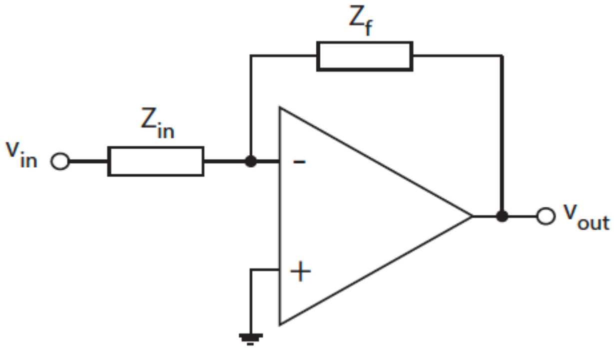
Operational amplifiers (OpAmps) = amplifiers with high internal gain a_0 whose behaviour is determined to first order by external circuit elements, in particular by impedances fed back to the input. For this reason the amplifier itself can be treated as a generic circuit element represented by a triangle. In the ideal case the (internal) 'open loop gain' a_0 of the OpAmp is ∞. In praxis, a_0 is smaller than ∞ (typically ~10^5) and frequency dependent.

OpAmp Golden Rules

(1) The output attempts to do whatever is necessary to make the voltage difference of the inputs vanish
(2) The inputs draw (almost) no current.



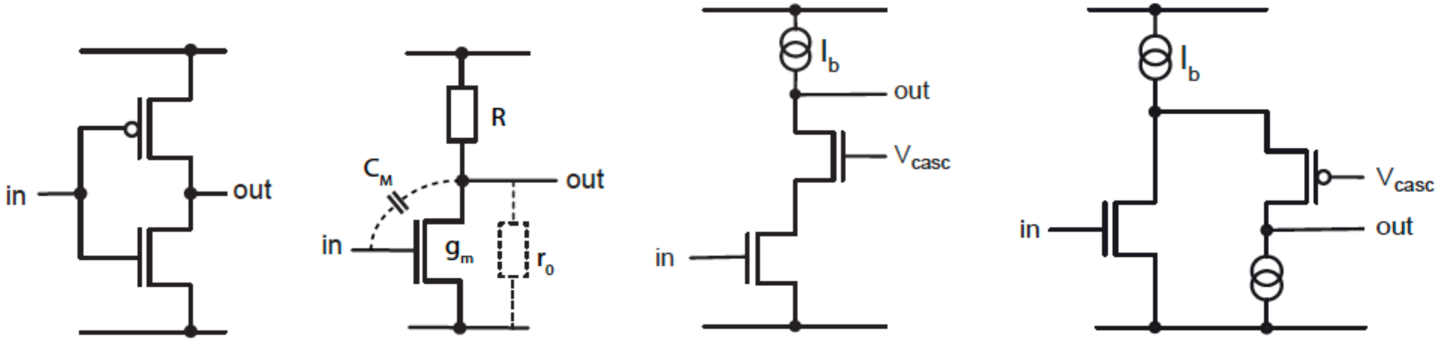
(a) OpAmp wiring symbol.



(b) OpAmp with feedback.

$$\frac{v_{out}}{v_{in}} = -\frac{Z_f}{Z_{in}}$$

If one input is grounded the other is said to be on virtual ground due to golden rule 1 .



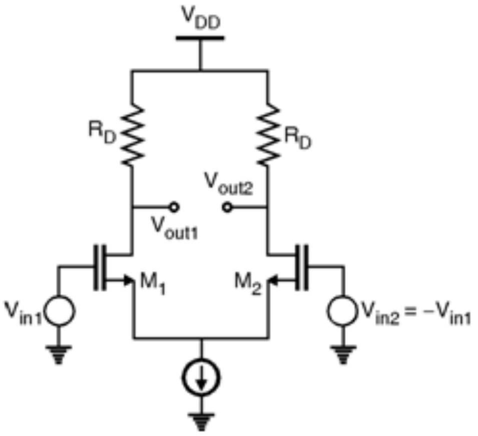
(a) CMOS inverter. (b) Simple transistor amplifier. (c) Cascoded transistor amplifier. (d) Folded-cascode stage.

$$|a_0| = g_m R || r_0$$

$$g_m = di_D / dv_{gs}$$

r_0 is the dynamic output resistance

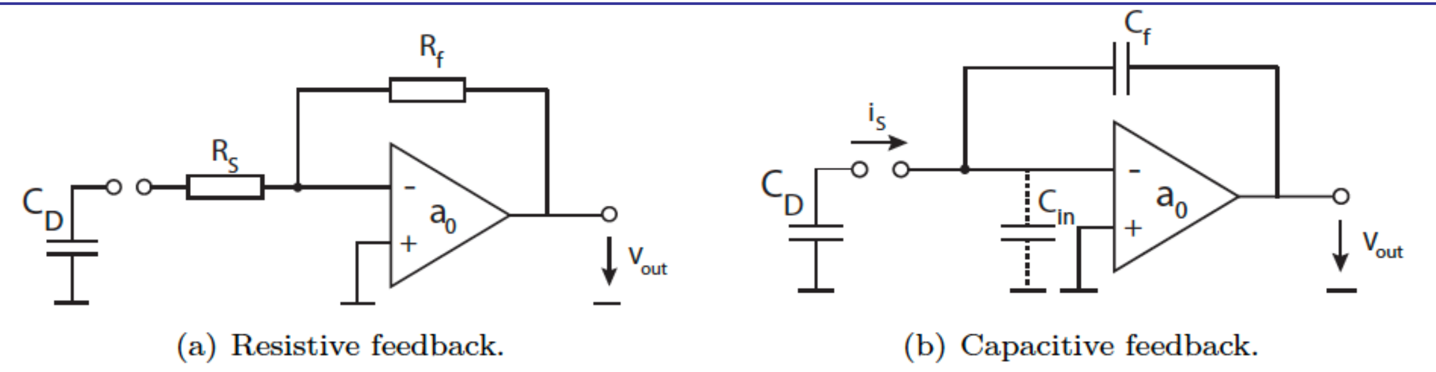
typically differential



Symmetric differential amplifier

amplifiers can be

voltage amplifier: $V \rightarrow V$,
current amplifier: $I \rightarrow I$,
transconductance amplifier: $V \rightarrow I$,
transimpedance amplifier: $I \rightarrow V$,
charge amplifier: $Q \rightarrow V$ (or I).



current or voltage amplifier

charge (sensitive) amplifier (CSA)
(= current integrator)

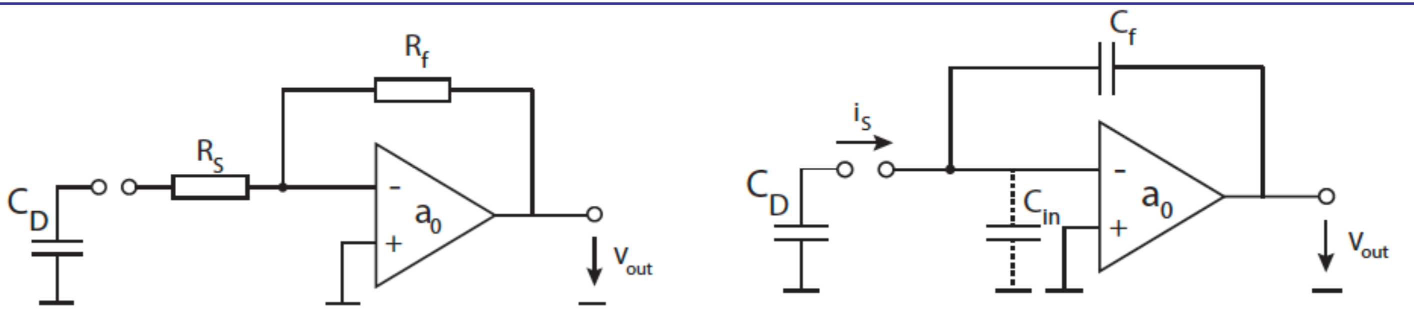
detector ≙ "capacitance to be discharged"

if $R_S C_D \gg \Delta t_{signal}$ (=> "detector integrates signal on C_D")
$\Rightarrow$ have voltage V_D at amplifier input: $v_{in}(t) = V_D \exp(-t/R_S C_D)$

$$\Rightarrow v_{out}(t) = -\frac{R_f}{R_S} v_{in}(t) = -\frac{R_f V_D}{R_S} \exp(-t/R_S C_D)$$

voltage amplifier
used when voltage input signals are sufficiently large and fast rise times are aimed for

if r_{out} = large ≙ (current source) => V → I transconductance. ampl.



(a) Resistive feedback. (b) Capacitive feedback.

current or voltage amplifier

charge (sensitive) amplifier (CSA)
(= current integrator)

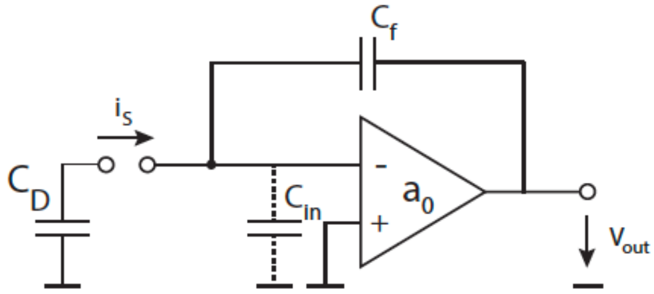
detector ≙ "capacitance to be discharged"

if $R_S C_D \ll \Delta t_{signal}$ (=> "detector delivers signal immediately to the amplifier")

$$\Rightarrow \quad v_{out}(t) = -\frac{R_f}{R_S} v_{in}(t) = -R_f i_S(t)$$

current amplifier
used when voltage input signals are sufficiently large and fast rise times are aimed for

if r_{out} = small ≙ (voltage source) => I → V transimpedance ampl.



(b) Capacitive feedback.

charge (sensitive) amplifier (CSA)

(= current integrator)

The signal current is integrated on C_f

$$v_{out}(t) = -a_0\, v_{in}(t) = -\frac{1}{C_f}\int_0^t i_S\, dt' = -\frac{Q_S(t)}{C_f}$$

over C_f we have
$$v_f = v_{in} - v_{out} = v_{in}\,(a_0 + 1) = \frac{Q_f}{C_f}$$

Golden Rule 2 =>
$$Q_S = Q_f = C_f(a_0+1)\,v_{in}$$

=>
$$C_{in} = \frac{Q_S}{v_{in}} = C_f\,(a_0+1)$$

dynamic input capacitance

(should be very large, else C_D incompletely discharged => unwanted x-talk possible)

for a_0 large => $C_{in} \gg C_D$
$$A_Q = \left|\frac{v_{out}}{Q_S}\right| = \frac{a_0 v_{in}}{(C_D + C_{in})\,v_{in}} \approx \frac{a_0}{C_{in}} = \frac{a_0}{a_0+1}\,\frac{1}{C_f} \approx \boxed{\frac{1}{C_f}}$$

gain