

#### From analog to digital

#### and back again...





## Analog to Digital Converter

An ADC converts a continuously variable signal, a voltage or a current, into a sequence of numbers, represented by logic levels on a group of wires.



From analog to digital





- Quantization replaces a range of continuous values by a set of discrete ones.
- Usually the number of levels is a power of 2.
- The difference between the original signal and the discrete representation is the <u>quantization error</u>



Power in original signal:

$$P_{s} = \frac{A^{2}}{T} \int_{0}^{T} \sin^{2} \omega t \, dt = \frac{A^{2}}{2}$$

Quantized to *n* bits, one quantum is:

$$q = \frac{2A}{2^n} = A \cdot 2^{-(n-1)}$$

Maximum quantization error:

$$\varepsilon = \frac{\pm q}{2} = \pm A \cdot 2^{-n} (with \ p(\varepsilon) = \frac{1}{q})$$

Power in quantization error:

$$P_{\varepsilon} = \int_{-q/2}^{q/2} p(\varepsilon) \varepsilon^2 d\varepsilon \approx \frac{A^2 \cdot 2^{-2n}}{3}$$





Multiplication of the signal by a train of impulses w(t) with period  $T_s$  (=  $1/F_s$ ):

$$g(t) = u(t) \cdot w(t)$$
  

$$g(t) = u(t) \sum_{n = -\infty}^{\infty} \delta(t - nT_s)$$
  

$$g(t) = \sum_{n = -\infty}^{\infty} u(nT_s) \delta(t - nT_s)$$

Fourier transform of 
$$w(t)$$
:

$$W(f) = \int_{-\infty}^{\infty} w(t) e^{-j2\pi f t} dt$$

$$W(f) = \sum_{n = -\infty}^{\infty} e^{-j2\pi n f T_s} = 1 + 2\sum_{n=1}^{\infty} \cos 2\pi n f T_s = \sum_{n = -\infty}^{\infty} \delta(f - \frac{n}{T_s})$$

U(f)



The spectrum of the sampled signal is the convolution of U(f) and W(f)

$$G(f) = U(f) * W(f) = \int_{-\infty}^{\infty} U(\phi) W(f - \phi) d\phi \quad \overset{W(f)}{=}$$

$$G(f) = \int_{-\infty}^{\infty} U(\phi) \sum_{n = -\infty}^{\infty} \delta(f - \frac{n}{T_s} - \phi) d\phi$$

$$G(f) = \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} U(\phi) \delta(f - \frac{n}{T_s} - \phi) d\phi$$

$$G(f) = \sum_{n=-\infty}^{\infty} U(f - \frac{n}{T_s})$$

After sampling, the signal spectrum repeats for all multiples of  $F_s$ 

- Fs -

0

 $\rightarrow$ 



If the sampling rate  $F_s$  is less than twice the signal bandwidth, the spectral images overlap. The avoid this, the following condition must be fulfilled:

 $F_s > 2 \cdot BW$ 

#### This is the Nyquist criterion

One way this condition can be fulfilled is by filtering the analogue signal prior to digitizing it, using what is called an anti-aliasing filter. Since brickwall filters cannot be made, the sampling rate should usually be quite a bit greater than twice the signal bandwidth.





Each of the images in the spectrum of the sampled signal contains all the information needed to reconstruct the original. They are *aliases*.

We might reconstruct the original signal with a filter that rejects everything except the original frequency band. After filtering, the spectrum is exactly that of the original signal, in other words, no information is lost. We have recovered the original signal exactly.

This is Shannon's theorem





Filter the baseband using a rectangular filter H(f). The filter timedomain response is the inverse Fourier transform of its frequencydomain shape:

$$h(t) = \mathscr{F}^{-1} \{ H(f) \} = \int_{-\infty}^{\infty} H(f) e^{j2\pi ft} df = \int_{-F_s/2}^{F_s/2} e^{j2\pi ft} df = F_s \frac{\sin \pi F_s t}{\pi F_s t}$$
  
Convolution of filter with sample stream:  
$$u_r(t) = g(t) * h(t)$$
$$u_r(t) = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} u(\tau) \cdot \delta(\tau - nT_s) \cdot h(t - \tau) d\tau$$
$$u_r(t) = \sum_{n=-\infty}^{\infty} u(nT_s) \cdot h(t - nT_s)$$
$$u_r(t) = \sum_{n=-\infty}^{\infty} u(nT_s) \cdot F_s \frac{\sin \pi F_s(t - nT_s)}{\pi F_s(t - nT_s)}$$



Note that exactly the same spectrum results for any signal frequency band displaced by  $m \cdot F_s$  (for integer *m*)

This goes by the name of *sub-sampling* 

$$G(f) = \int_{-\infty}^{\infty} U(\Phi + mF_s) W(f - \Phi) d\Phi$$

$$G(f) = \sum_{n=-\infty}^{\infty} U(f - \frac{n}{T_s} + mF_s)$$

$$G(f) = \sum_{n=-\infty}^{\infty} U(f + \frac{m-n}{T_s})$$

$$G(f) = \sum_{n=-\infty}^{\infty} U(f - \frac{n}{T_s})$$







We could also choose a different spectral image to (re)construct the signal:

First work out the time-domain representation of the filter:

$$H(f)=1$$
 for  $F_s < |f| < \frac{3}{2}F_s$   
 $H(f)=0$  everywhere else



$$h(t) = \mathscr{F}^{-1} \{ H(f) \} = \int_{-\infty}^{\infty} H(f) e^{j2\pi ft} df$$
$$h(t) = \int_{\frac{-3}{2}F_s}^{\frac{3}{2}F_s} e^{j2\pi ft} df - \int_{-F_s}^{F_s} e^{j2\pi ft} df$$

$$h(t) = 3F_s \operatorname{sinc}(3\pi t F_s) - 2F_s \operatorname{sinc}(2\pi t F_s)$$





Then convolve the sample stream with the filter function:

$$u_r(t) = \int_{-\infty}^{\infty} g(\tau) \cdot h(t-\tau) d\tau$$

 $u_r(t) = g(t) * h(t)$ 

 $\infty$ 



$$u_{r}(t) = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} u(\tau) \cdot \delta(\tau - nT_{s}) \cdot h(t - \tau) dt$$
$$u_{r}(t) = \sum_{n=-\infty}^{\infty} u(nT_{s}) \cdot h(t - nT_{s})$$

$$u_r(l) - \sum_{n=-\infty} u(nI_s) \cdot n(l-l)$$

$$u_{r}(t) = \sum_{n=-\infty}^{\infty} u(nT_{s}) \{ 3F_{s} \operatorname{sinc} (3\pi(t-nT_{s})F_{s}) - 2F_{s} \operatorname{sinc} (2\pi(t-nT_{s})F_{s}) \}$$



#### 1915: E.T. Whittaker : Interpolation theory

Couldn't get my hands on that one. Everyone refers to him, so I mention him as well.

## 1928: H. Nyquist : Telegraph transmission theory

Classic! Deals with signal distortion in transmission channels like undersea cables, which were a hot subject, at the time.

#### 1933: V.A. Kotelnikov : Carrying capacity of the ether

Detailed demonstration that band-limited signals can be represented by a sum of sinc functions, apparently independently from Nyquist and Whittaker.

## 1949: C. Shannon : Communication in the presence of noise

Classic! Gives transmission capacity of a channel as a function of bandwidth and signal to noise ratio. The sampling theorem is dealt with in section II.

Other names: R.V.L Hartley, J.M. Whittaker, C-J. de la Vallée Poussin, ...



Although mathematically Dirac deltas, brick-wall filters and infinite sums are quite nice to handle, in real electronic circuitry, you can't have them.

The Dirac  $\delta$  is replaced by an (almost) rectangular pulse of one sampling period duration, and filters are described by finite polynomials, with finite-slope band edges. So, the output is held constant during each sampling period, which is functionally called a zero-order hold, and a low-pass filter smooths over the steps.





As a consequence, the reconstructed signal spectrum is convolved with a sinc( $f/F_s$ ) function and some energy from adjacent spectral images leaks into the desired band. Note that at the Nyquist frequency,  $F_N = F_s/2$ , the response is down by 3.9dB. If this is a problem, the reconstruction filter may be designed to compensate. (You can also pre-compensate in the digital domain.)



From analog to digital



For 'large enough' and 'busy enough' signals, the quantization error is a random variable with a flat distribution.

 $\rightarrow$  Quantization noise is white and spread out evenly over  $0 < f < F_s/2$ .





Unfortunately, quantization noise isn't always white:

- Simple ratios between  $F_{in}$  and  $F_s$  cause some of the quantization noise power to concentrate in discrete spectral lines
- ADC non-linearities cause harmonics of the input signal

Spurs appear in the spectrum:

SFDR is the distance between the input signal and the greatest spur.

A little bit of dither can help to reduce spurs.

(Dither is the intentional injection of a little bit of noise.)



Non-linearity creates harmonics

THD is the rms sum of the first 6 harmonics compared to the input signal, in dB









In real ADCs, the quantization function isn't perfectly uniform

For an input signal with a uniform distribution, the distribution of output values is no longer uniform.

The DNL measures the normalized error of the nominal size of each quantization step.

1.00







Non-monotonicity



INL measures the deviation of the ADC characteristic from a straight line through the end points (or sometimes from a least squares fit)



21



Apply a nearly full-scale sinusoidal signal. Measure  $P_{\varepsilon}$ , as the rms sum over all frequencies, ignoring DC and the first five harmonics, and solve for *n*:

$$SNR = \frac{P_s}{P_{\varepsilon}} = 1.5 \cdot 2^{2n}$$

If you choose to also add in all harmonics into the calculation of  $P_{\varepsilon}$ , you would get the SINAD. (SIgnal over Noise-And-Distortion) (Which looks a little bit worse, of course)



Non-linearity also causes inter-modulation distortion. (Creating sum and difference frequencies from two applied tones  $f_{imd} = \pm n f_1 \pm m f_2$ .)

IMD is the rms sum of the intermodulation products compared to the rms sum of the input signals (usually in dB).

The order of an IMD product is |n|+|m|





Importance of clock jitter depends on rate of change of <u>analogue input</u> signal



A clock timing error  $\Delta t$  yields an amplitude error:

$$\Delta U = \frac{d u(t)}{dt} \cdot \Delta t$$

This is a severe condition!



24



Tolerable jitter:

$$\Delta t = \frac{1}{\frac{d u(t)}{dt} \cdot 2^n}$$

Ex: Suppose we digitize a 100MHz sinusoid to 10 bits:

$$u(t) = \sin 2\pi 10^8 t \qquad \Rightarrow \qquad \frac{d u(t)}{dt} = 2\pi 10^8 \cos 2\pi 10^8 t$$

So at the steepest slope:

$$\Delta t = \frac{1}{2\pi 10^8 \cdot 2^{10}} \approx 1.6 \, ps \; \Rightarrow$$

A good quartz or ceramic resonator oscillator is needed





With 1 ps of jitter, a 100 MHz signal can only be digitized to 10.5 bits Relaxed by root(decimation ratio) for  $\Sigma$ - $\Delta$  converters



Even using the best clock, the resolution reaches a limit. For example, for an actual 12-bit ADC, the ENOB vs.  $F_{in}$  plot might look like this:







Upconverted thermal, schottky and 1/f noise

$$\Delta t_{rms} = \frac{T_0}{2\pi} \sqrt{\int_0^\infty S_\varphi(f) \cdot 4\sin^2(\pi f \tau) df}$$

$$L(f)$$
  
 $L(f)$   
 $L(f)$ 

 $S_{\varphi}(f)$  is the spectral density of the phase noise

(Function of Fourier frequency *f* and sum of both sidebands)

 $\sin^2(\pi f \tau)$  is a weighting function

(For low frequencies of *S*, the phase can't drift very far, when  $\tau = n/f$ , the contribution cancels, and there are maxima in between.)

 $\tau$  is the time between two events (Usually  $\tau = T_0$ )

The term 
$$\frac{T_0}{2\pi}$$
 converts phase into time



#### Ground noise between clock source and ADC aggravates jitter



(Complex logic circuits cause all sorts of interference)

But if you must, then resynchronize with the original clean clock



- $\cdot$  ADCs (and DACs too!) need good quality clock sources
- · Digital electronics is not optimized for low crosstalk
- PLLs in FPGAs usually have \*very\* poor jitter specs

## Treat your clock oscillator like a sensitive analogue circuit

- Filter and bypass clock generation & distribution power supply extra carefully
- Keep PCB layout tight and compact, minimize loop areas
- $\cdot$  Refer clock source to the same GND as the ADC
- $\cdot$  Do not route an ADC clock through an FPGA
- · Don't use left-over gates in clock buffer package for other purposes





You were expecting this:

but you got this:

Reason: The ADC delivers 2's complement numbers and the sign bit is not in the right place.



The sign bit isn't in the right place:



Apply sign extension:  $a = (a \wedge 0 \times 800) - 2048;$ 



or logical left shift: a<<=4;









BCD and display driver outputs. Handy for stand-alone instruments, panel meters, hand-held multi-meters.

Gray code: Only one bit changes between adjacent values. No glitches. Resolver disks. Angular and linear transducers.





· Inductive coupling

Coupling paths:

· Capacitive coupling

I

(Keep high-Z nodes and nodes with high dE/dt far apart, or put grounded shields between them)



... or how to get your signal digitized cleanly:







Signal integrity





Buffer amplifier adapts signal to ADC and prevents sampler kickback RC circuit at ADC input isolates amplifier from ADC input capacitance

Instrumentation amplifier: Good common mode rejection, high gain, but only at low frequency

Baluns and transformers: Good common mode rejection at high frequencies





Single-ended to differential conversion (Often used for high-performance or low voltage ADCs)

Baluns and transformers can also be used for this.

Or use monolithic differential buffer amplifiers  $4006 \overline{4}$  (ADA4941, AD8351, LT6411, THS4503, etc.)





#### Summary:

- <sup>•</sup> Adapting signal range to ADC input range (Scaling & level shifting)
- <sup>•</sup> Conversion between single-ended and differential signals
- <sup>•</sup> Protecting the ADC input from overload
- <sup>•</sup> Terminate long cables into their characteristic impedance...
- " ...or, to the contrary, provide a high impedance to avoid loading the source
- <sup>•</sup> Rejecting interference
- <sup>•</sup> Filtering out-of-band frequencies (Anti-alias filter, noise reduction)
- " Holding input constant while conversion takes place



Architecture	Speed	Resolution	Linearity	Applications
Flash	Very fast (GS/s)	Poor (8 bits)	Poor	Oscilloscopes
				Transient recorders
Successive approximation	Fast (MS/s)	Fair (14 bits)	Fair	DSP, video, digital receivers, instrumentation
Σ-Δ	Slow (kS/s)	Excellent (24 bits)	Excellent	Process control, audio, weight, pressure, temperature measurement
Dual-slope Integration	Very slow (S/s)	Very good (18 bits)	Very good	Bench-top and hand-held measuring instruments, battery powered devices

#### Other architectures:

- Mixed forms (E.g. flash with SA, or flash with  $\Sigma$ - $\Delta$ )
- Tracking ADC
- Voltage-to-frequency converters









SAR ADC binary decision tree picture

Sequence of operation:

- <sup>•</sup> Compare input with half scale, keep if greater.
- Add one quarter and compare, keep if input greater.
- " Add 1/8 etc...

Usually clocked or strobed Serial data interface Often fixed conversion rate Sometimes poor DNL







Flash ADCs are the fastest:

- A resistor divider chain creates all possible decision levels from a single reference.
- 2<sup>n</sup>-1 comparators compare each level with the input signal.
- Digital logic converts "thermometer" code into binary.
- Sensitive to 'sparkle' codes
- Metastability
- Input capacitance

Usually 8 bits, rarely more than 10





CERN

Segmented or pipelined ADC:

- Sample rate comparable to flash ADCs, but with several clock periods of latency.
- Resolution comparable to successive approximation architecture.





**Differential Nonlinearity** 



LTC2242-12

224212 703

100

0.1aF

T1 = MA/COM ETC1-1T

RESISTORS, CAPACITORS

ANALOG

Vom

AN' 2DF

12-bit, 250 MS/s, 5-stage pipeline ADC Differential LVDS <u>or</u> de-multiplexed outputs 2pF sampling capacitor Differential analogue input, BW 1.2 GHz





Increase sample rate:

47

→ Quantization noise is spread over a larger BW. Numerically low-pass filter the sample stream:

→ Out-of-band quantization noise power is removed. Conclusion:

→ SNR gets better by 3dB/octave of oversampling rate.
Dither may be necessary for very quiet ADCs.







Average duty cycle of comparator output reflects input value.

Digital low-pass decimation filter trades sample rate for resolution.

Good DNL, good resolution, slow.





## Higher order loops and noise shaping, E.g., a 2nd order modulator:



CERN

Side tones, idling patterns, birdies

- <sup>\*</sup> For some input values, the  $\Sigma/\Delta$  modulator can produce repeating patterns with repetition rates well below the sampling frequency
- These may leak though the decimation filter, causing a side tone or 'birdie'



 Possible remedies include using higher order modulators and dither, to randomize things



# Programmable, instrumentation ADC 12-16 bits, depending on filter settings





## Programmable, instrumentation ADC 12-16 bits, depending on filter settings





5

10

15

Frequency (MHz)

20

25

30

0

-20

-40

-60

-80

-100

-120

0

Magnitude (dB)



## Integrate input during T1

Then:

Integrate reference until zero is reached, while counting clock pulses Final count is proportional to input

Can be built for low power operation, good for battery power equipment (Multimeters etc.) Integration time often chosen to reject power line interference Excellent DNL





Sometimes the signal source is inherently a pulsed current source (Photomultiplier) Variants: Time to Digital Converter (TDC)



Directly drives an LCD display ±200 mV full scale Conversion time 300 ms Consumes 1 mA @ 9 V





Analogue section of ICL7106



Input voltage is integrated onto Ci A fixed-size packet of charge is removed each time the switch connects Cf to the input.



Applications: Process control Easy to use as integrating converter (Just add a counter) In combination with F-to-V converter: Cheap isolation amplifier

Low-cost Slow! Signal easy to transmit over large distances Very good linearity









Digital potentiometers (DS1669, MAX5438, AD5259, etc. Digital output sensors, temperature, acceleration (AD7414, AD16006) Capacitance-to-digital converters (AD7745)



Architecture	
Kelvin-Varley divider	Accurate, monotonous. Mainly as building block in integrated DACs
Thermometer DAC	Monotonous. Limited number of bits
Binary weighted ladder	Very common, but subject to glitches
R-2R ladder	Widely used. Not very power efficient
$\Sigma$ - $\Delta$	Linear, accurate, but complex.

#### Mixed architectures:

- " Segmented DAC
- <sup>•</sup> Interpolating DAC

#### Variants:

- Multiplying DACs
- " Current or voltage output
- <sup>•</sup> Differential or single-ended





- <sup>•</sup> Still rivals the best modern DACs
- Available as rack mounted units (Expensive!)
- " Used as sub-circuit in IC DACs
- $^{\cdot}$  Variable  $Z_{out}$ : Buffer amplifier needed



Example: Fluke 720ALinearity, resolution:  $10^{-7}$ Input resistance:  $100 \text{ k}\Omega, 0.005\%$ 







- " Inherently monotonic
- " No glitches
- <sup>"</sup> Limited number of bits (2<sup>N</sup> current sources!)
- " Used as a sub-circuit in segmented DACs



Example: THS5641, 8 bit, 100 MS/s, 35 ns settling time



- " Very common architecture
- " Uses only two resistor values
- " Voltage or current output
- " Not very power efficient
- <sup>•</sup> Often as multiplying DAC

$$V_{out} = V_{ref} \frac{N}{2^n}$$

Example: AD5445, 12 bit current output, 20.4 MS/s, 80 ns settling time







#### PWM DAC





- <sup>"</sup> Pulse Width Modulation
- " Inherently linear
- " Limited resolution
- <sup>"</sup> Often used in μ-controller chips
- <sup>•</sup> Applications: Motor control



Combination of other architectures, to optimize speed, linearity, SFDR, glitch energy, etc.

Common for high performance DACs used in instrumentation and communication equipment.

Example: AD9753, (12 bit, 300 MS/s) combines two thermometer (5 and 4 bits) sections and a binary weighted stage (3 bits).





Combining a Kelvin divider with an R2R ladder





- There are ADCs and DACs for almost any imaginable application
- $\cdot$  Performance is ever getting better
- Prices keep going down (15 years ago, a 12 bit 10 MS/s ADC cost 1 k\$. Now it's around 10 \$)

Digital is here to stay!