Speech Compression

- Recommended Reading: J. Harrington and S. Cassidy, *"Techniques in Speech Acoustics"*, Kluwer, 1999
- Contents
  - Uncompressed audio data rates
  - ADPCM
  - SB-ADPCM
  - LPC

Uncompressed audio data rates

- Voice: 8000samples/sec, 8bits/sample, mono

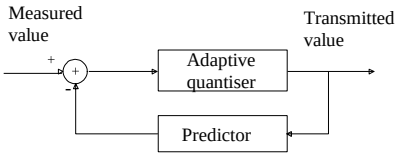
  = 64000bits/sec (64kbps)
- CD: 44100samples/sec, 16bits/sample, stereo

  =1411200bits/sec (~1.5Mbps)

ADPCM (Adaptive Differential PCM)

- Uses the statistical properties of human speech (=> not compatible with fax/modem signals)
- Makes a prediction about the size of the next sample, based on previous info
- Transmitter then sends only the difference between real value and predicted value
- Receiver uses the same prediction algorithm, together with the differences to reconstruct the speech data
- Enables the data rate to be reduced to 32kbps
- Used on international telephone links
- Specified in G.721, G.722, G.723, G.726, G.727

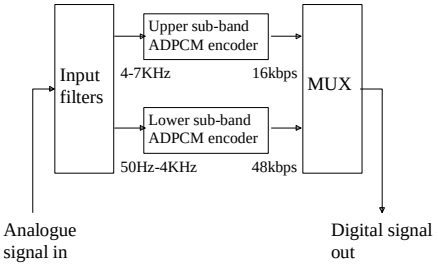
ADPCM



SB-APDCM (Sub-band ADPCM)

- Given 64kbps: ADPCM could produce better than toll voice quality (eg radio)
- Sub-bands are 0-4kHz (given 48kbps), 4-7kHz (given 16kbps)
- Low band contains more audio energy, high band contains intelligibility info.
- Standardised in G.722

SB-ADPCM



Linear Predictive Coding (LPC)

- Introduced in the 1960s
- n^{th} signal sample is represented as a *linear* combination of the previous p samples, plus a residual representing the *prediction* error:

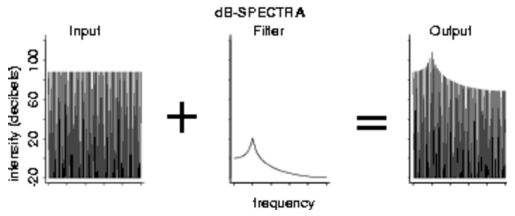
$x(n) = a_1x(n-1) + a_2x(n-2) + \ldots + a_px(n-p) + e(n)$

- If the error ('e') is small enough, we can just transmit the coefficients ('a's)

LPC

- coefficients ('a's) correspond to those of a vocal tract filter and the error signal ('e') corresponds to a source signal
- Source signal will approximate either a voiced signal (which looks like a series of impulses) or a white noise source
- So, LPC involves "exciting" a source signal with a vocal tract filter

Impulses and Filters



LPC – Autocorrelation

- Minimise the error signal by choosing optimal coefficients ('a's)
- Use the autocorrelation criteria (aka root mean squared criterion):

$$\sum_{i=1}^{p} a_i R(i-j) = -R(j)$$

for 1<=j<=p,

where R is the autocorrelation of x(n) defined as

R(i) = E[x(n)x(n-i)]

LPC – Solving the autocorrelation formula

- In matrix form the equation can be written as

  R * a = r

  where the autocorrelation matrix R is a symmetric *Toeplitz* matrix with elements $r_{i,j} = R(i - j)$, vector r is the autocorrelation vector $r_j = R(j)$, and vector a is the parameter vector of a_i
- An algorithm by N. Levinson (proposed in 1947) and modified by J. Durbin (in 1959) recursively calculates the solution to the Toeplitz matrix.
- GSM coder uses an integer version of the Schur recursion (1917)

LPC

- Used in:
  - GSM (Groupe Speciale Mobile) (Residual Pulse Excited-LPC) (13kbps)
  - LD-CELP (Low-Delay Code Excited Linear Prediction) (G.728) (16kbps)
  - CS-ACELP (Conjugate Structure-Algebraic CELP) (G.729) (8kbps)
  - MP-MLQ (Multi Pulse – Maximum Likelihood Quantisation) (G.723.1) (6.3kbps)…