# Semi-automated extraction of expressive performance information from acoustic recordings of piano music

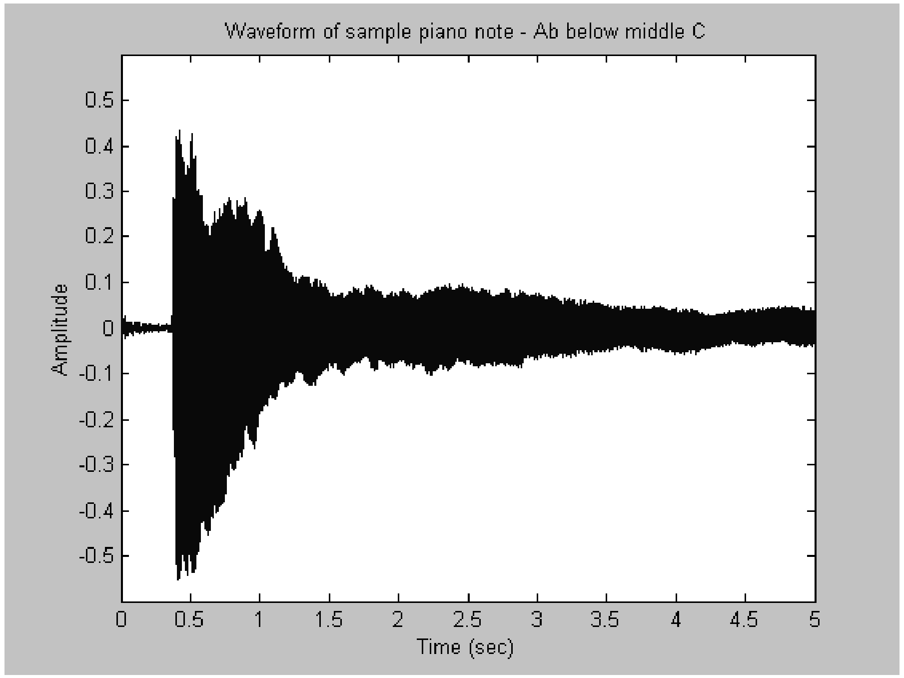
**Andrew Earis**

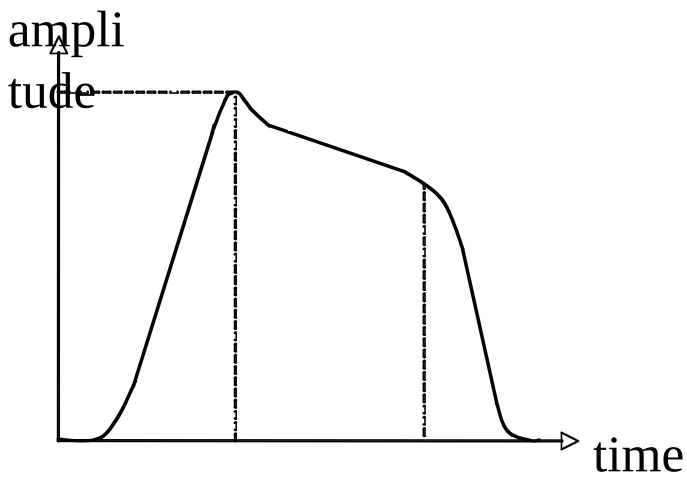
# Outline

- Parameters of expressive piano performance
- Scientific techniques: Fourier transform & wavelet transform
- CHARM: The semi-automated system

# Measurable features of expressive performance

- *Timing (note onset and offset)*
- *Dynamics (variations in intensity of notes and chords)*
- *Pitch*



# Timing parameterisation

- Note onset time. The time at which the note is played.
- Note offset time. The time at which the note is released.
- Attack duration. The time from the initial sound until its amplitude is at a maximum.
- Sustain duration. The time from maximum sound amplitude until the note is released.
- Decay time. The time from when the note is released until the sound ceases.

# Dynamics parameterisation

- Power of note
- Maximum amplitude (amplitude of attack)
- Sustain amplitude (amplitude, or change of amplitude, of the sustain)

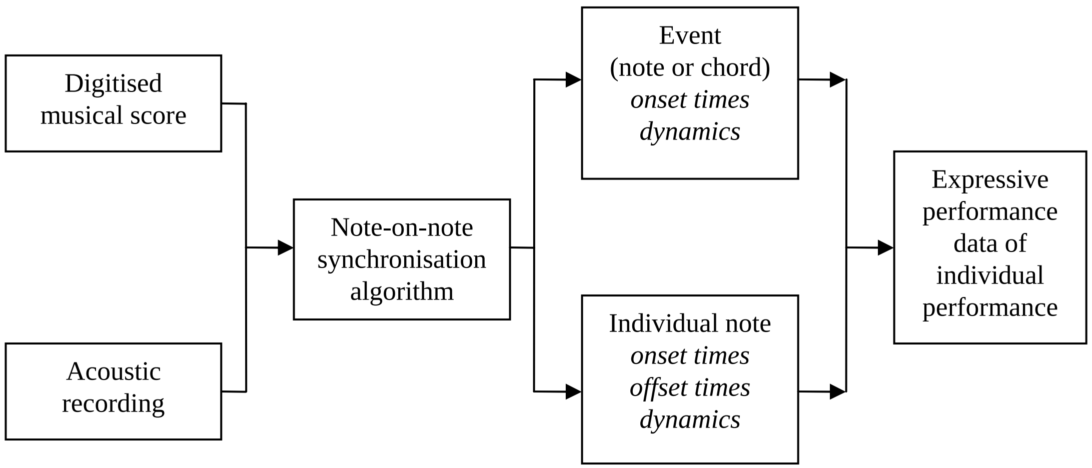
# Pitch parameterisation

- Fundamental frequency of note
- Frequency of individual partials (and how they deviate from the precise harmonic frequencies)
- Vibrato
  - Frequency of vibrato
  - Amplitude of vibrato
- Glissando

# Aims

- to extract the expressive performance information from acoustic recordings of piano music with reference to a digitised version of the musical score of the work being performed.

# Simple flowchart



# Methods

The location in time of individual notes and chords within an acoustic recording must be identified and related to a digitised score of the work being performed.

This is a classical time-frequency analysis problem

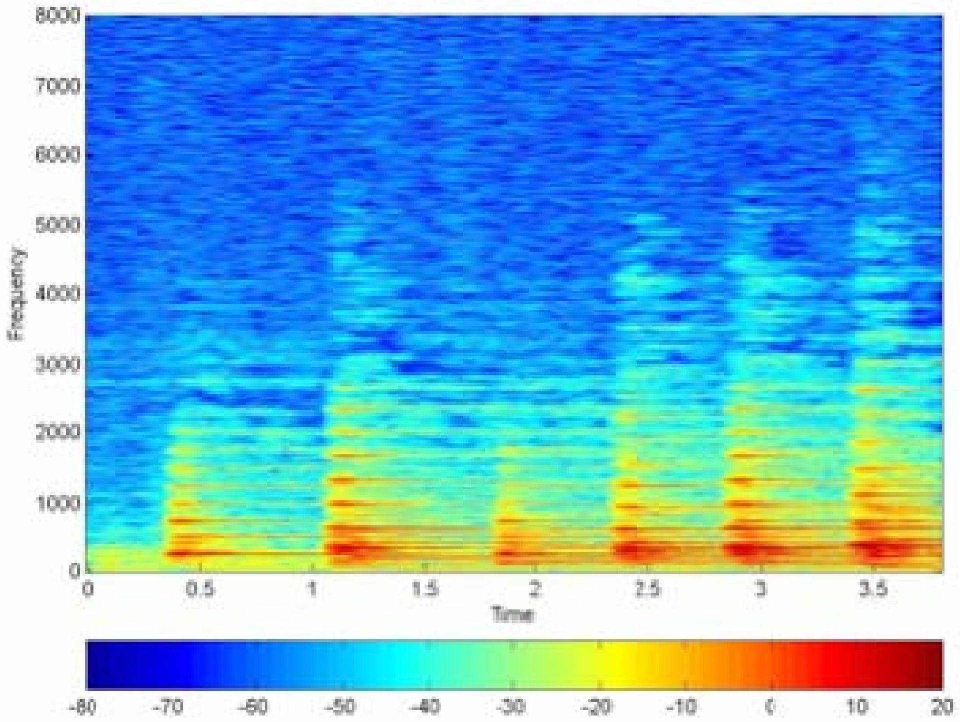
# Fundamental research issues

- A number of constraints on the time-frequency resolution required.
  - Required frequency resolution
    - determined by the interval between adjacent semitones
    - proportional to the frequency of the lower semitone
    - Spacing varies from about 1.6Hz at the bottom of the piano keyboard to over 120Hz at the top
    - the fundamental frequencies ranging from 27.5Hz to just over 4kHz.
  - Required timing resolution
    - The shortest note duration in rapid note passages is around 50ms.
    - The human ear is capable of perceiving differences of as little as 10ms.

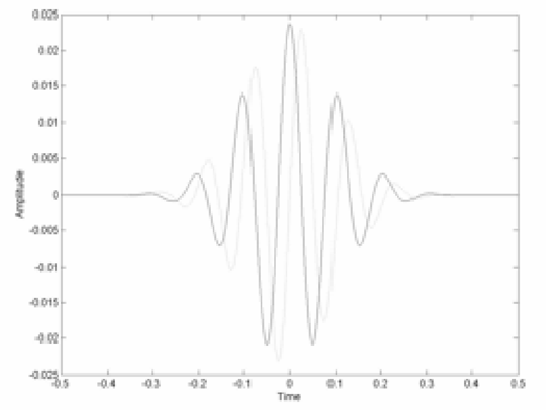
# Possible approaches

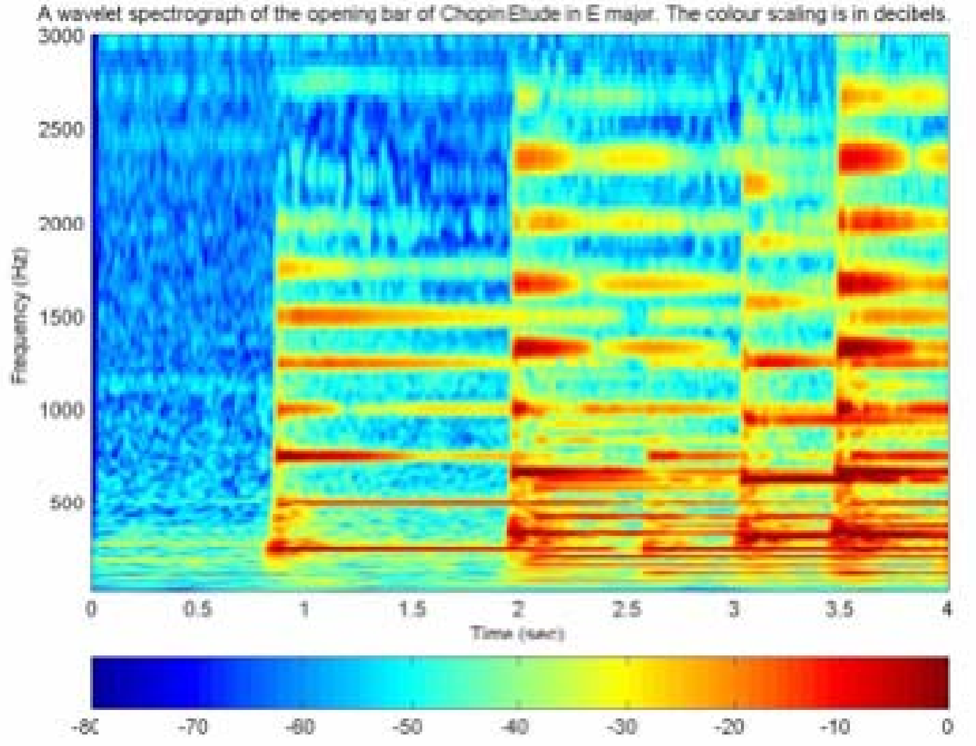
- Filterbank techniques
- Time-frequency analysis (using Fourier Transforms, wavelets and spectrographs)
- Modelling

# Short Time Fourier analysis



# Wavelet analysis



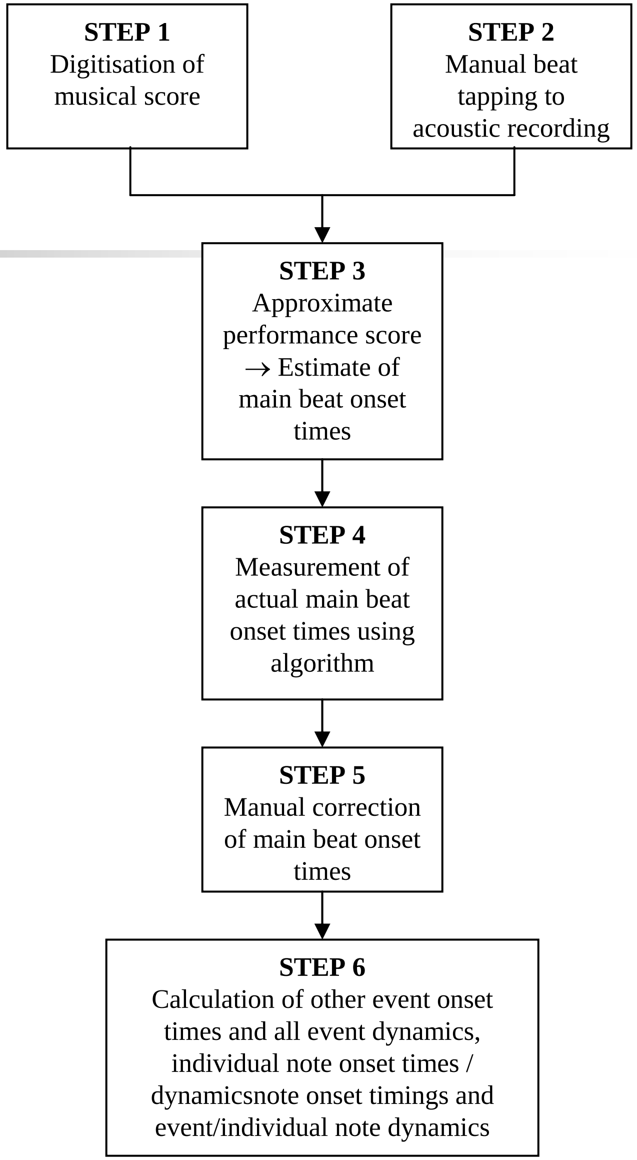
# Problems with ‘simple approach’

- Synchronization between score and acoustic recording relies on stepping through each event in the score in turn
- If a particular note or chord in the score is not played, the whole system can break down
- Other problems
  - too much noise in the recording
  - a high degree of rubato
  - chords are played with too much asynchrony
- The longer the passage of music to be analysed, the more likely it is that such a system will break down.
- Time-consuming manual checking of the data would also required to ensure accuracy.

# CHARM: Semi-automated system

- A semi-automated expression extractor, requiring user input for checking and verifying results as the algorithm proceeds i.e. note-on-note.
- Input to system: MIDI score, acoustic waveform
- Technology incorporated: filterbank/wavelet-based note onset detector
- Note/event parameter measurements: onset time, dynamics

# Algorithm plan



# Advantages

- Quick to use, results can be verified as they go along – know they are correct!
- Very easy to use
- Wide range of scores can be used
- Quick and easy to develop and automate further
- Results are verified by the user as measurements are made, ensuring that all results are correct and can be used reliably for musicological analysis
- Further enhancements can be made (e.g. increase level of automation, sophistication of parameters)

# STEP 1

- Digitisation of musical score
  - Score is scanned using a standard flatbed scanner
  - Converted to symbolic data with Sharpeye
  - Errors are corrected manually
  - Symbolic data is then converted to the Humdrum data format
  - This is outputted as ASCII tabular data

# STEP 2

- Manual beat tapping to acoustic recording
  - Standard beat tapping procedures
  - backbeat.exe
  - average of 20 complete taps initially!

# STEP 3

- Averaged manual beat tapped data combined with the digitized musical score
  - -> Approximate performance score

| Column | Description | |
|---|---|---|
| 1 | abstime | Average absolute time in ms of beat tapped timings |
| 2 | duration | Expected duration of note in ms based on score duration |
| 3 | note | MIDI number of note pitch |
| 4 | metric level | 1 = downbeat<br>0 = beat<br>-1 = offbeat |
| 5 | bar number | |
| 6 | absolute beat | measured from beginning of work |

# Approximate performance score: Mazurka in A minor op 7 no 2

| | | | | | |
|---|---|---|---|---|---|
| 1912 | 646 | 76 | 1 | 0 | 0 |
| 2558 | 463 | 77 | 0 | 1 | 1 |
| 3021 | 154 | 76 | -1 | 1 | 1.75 |
| 3175 | 603 | 57 | 0 | 1 | 2 |
| 3175 | 603 | 62 | 0 | 1 | 2 |
| 3175 | 603 | 65 | 0 | 1 | 2 |
| 3175 | 603 | 74 | 0 | 1 | 2 |
| 3778 | 652 | 57 | 1 | 1 | 3 |
| 3778 | 652 | 62 | 1 | 1 | 3 |
| 3778 | 652 | 65 | 1 | 1 | 3 |
| 3778 | 652 | 77 | 1 | 1 | 3 |
| 4430 | 1111 | 77 | 0 | 2 | 4 |
| 4914 | 627 | 60 | 0 | 2 | 5 |
| 4914 | 627 | 65 | 0 | 2 | 5 |

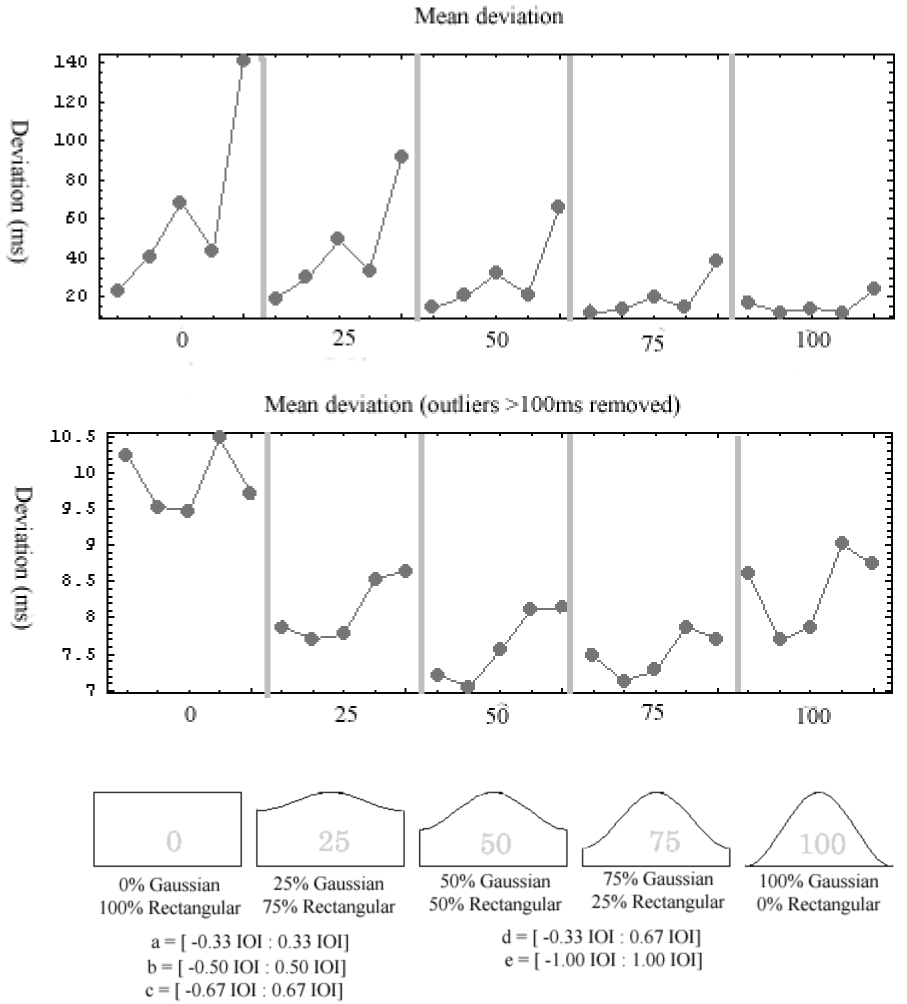
# STEP 4

- Correction of ‘main beat’ onset times (from beat tapped times) using wavelet-based algorithm, programmed in MATLAB

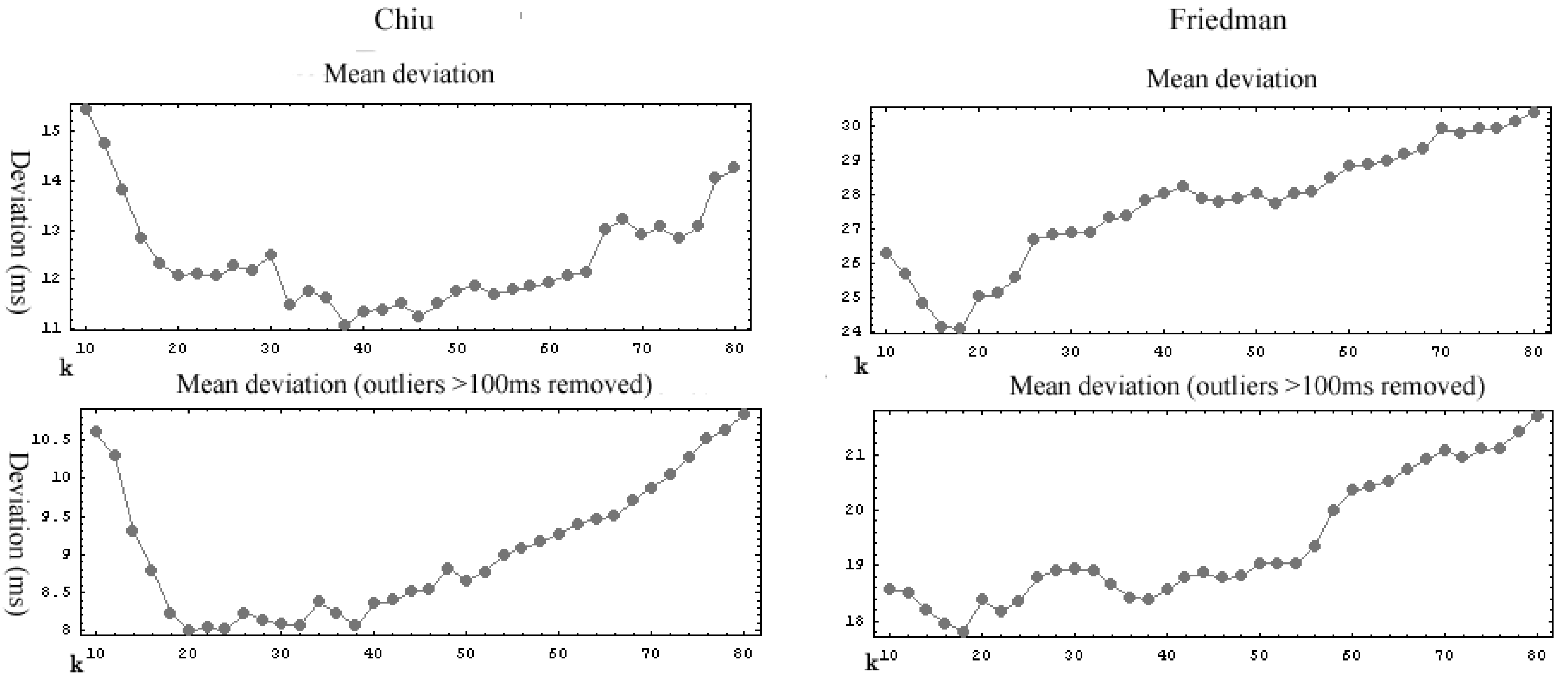
# Parameters

- Search window parameters
- Wavelet parameter - *k*
- Wavelet parameter - *LPF*
- Wavelet tuning
- Feedback analysis

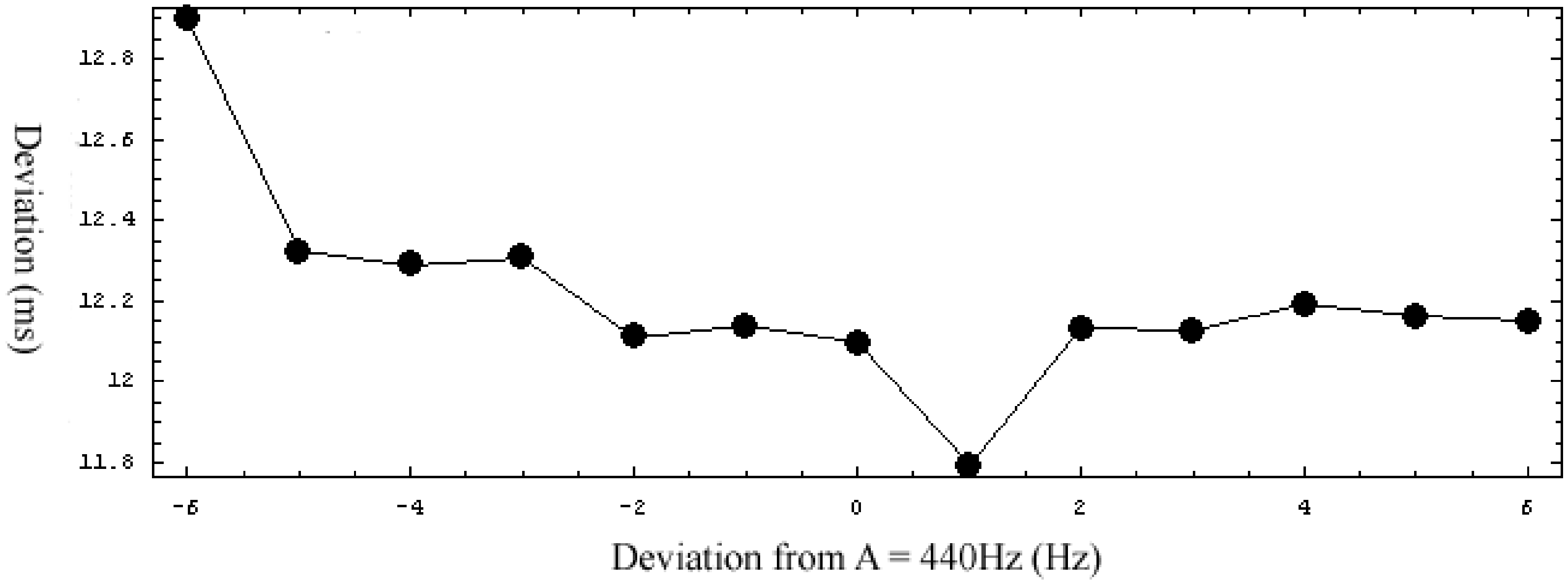
# Search window parameters



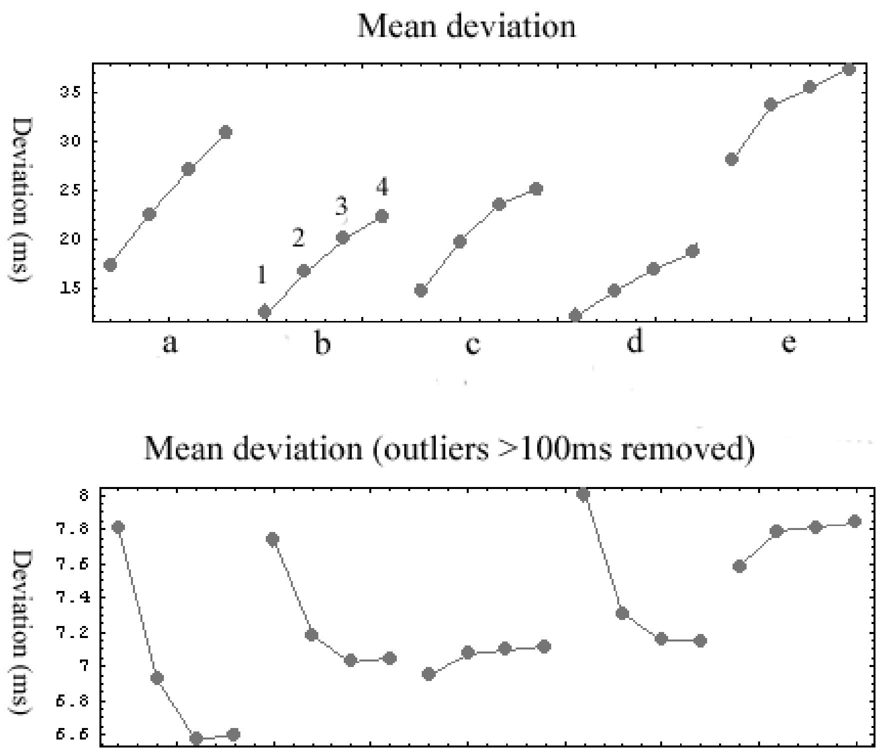
# Wavelet parameters - *k*



# Wavelet tuning



# Feedback analysis




a = [ -0.33 IOI : 0.33 IOI]
b = [ -0.50 IOI : 0.50 IOI]
c = [ -0.67 IOI : 0.67 IOI]

d = [ -0.33 IOI : 0.67 IOI]
e = [ -1.00 IOI : 1.00 IOI]

# STEP 5

- Manual correction of main beat onset times
  - Sonic Visualiser
- Output – table of corrected main beat onset times

# STEP 6

- Estimation of subbeat onset times from main beat onset times
- Correction of estimated subbeat onset times using algorithm
- Estimation of main beat and subbeat dynamics
- Estimation of individual note onset times

# End product

- MATLAB code converted to standard .exe files
- Input data required
  - Excel control spreadsheet - for batch processing of data
    - Input/output file names
    - Analysis parameters
  - ‘Main beat’ beat tapped input data spreadsheet
  - .wav sound recording

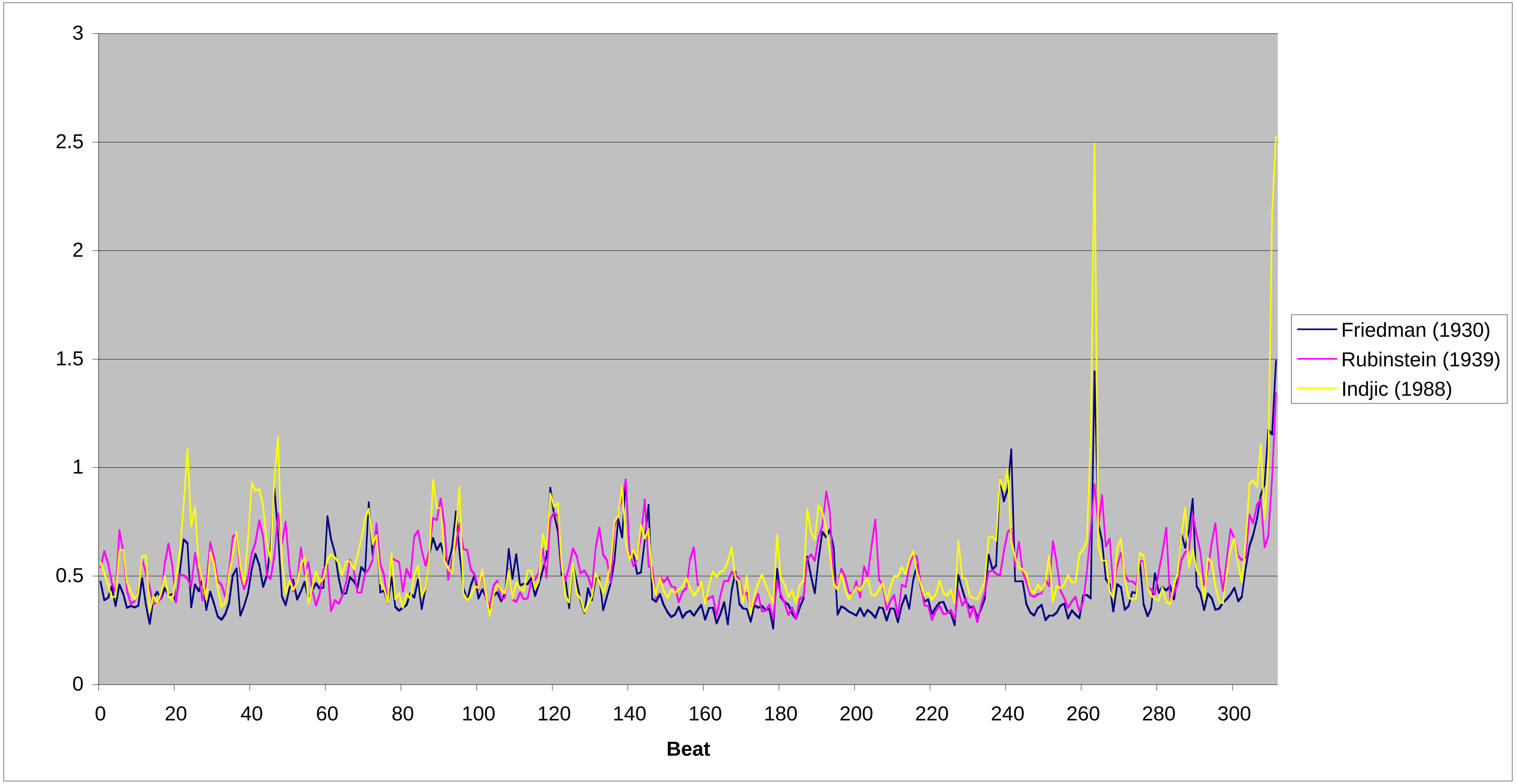
# End product

- Three different algorithms
  - mainbeatcorrection.exe
    - correction of ‘main beat’ tapped onset times
    - measurement of main beat dynamics (in dB)
  - subbeatmeasurement.exe
    - measurement of sub beat onset times
    - measurement of sub beat dynamics
  - individualnotemeasurement.exe
    - measurement of individual note onset timings and dynamics
- Will be available (or a variation of this) on CHARM website as downloadable resource

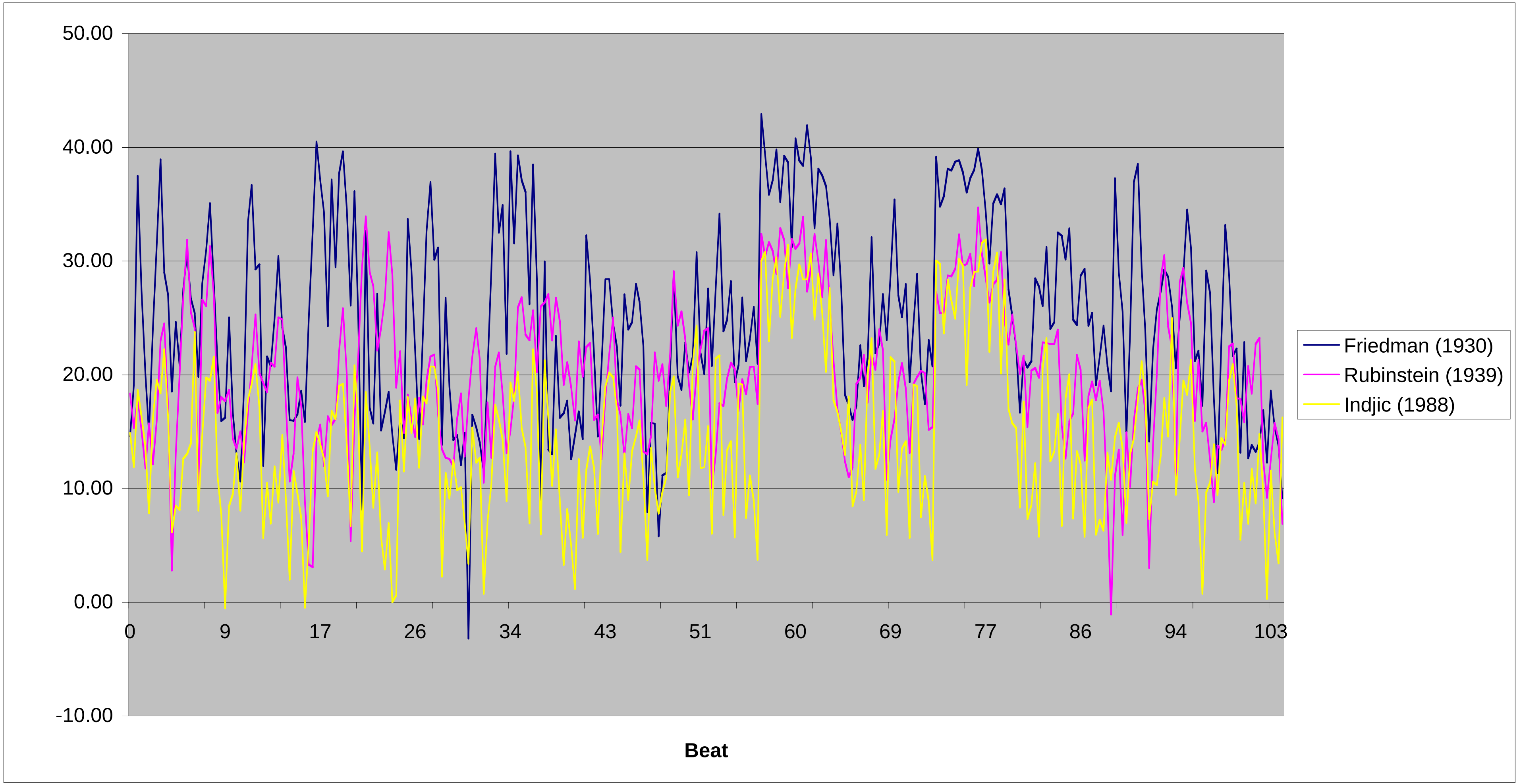
# End product

- For example…
  - 3 recordings of Chopin’s Mazurka in A minor op 7 no 2
    - Friedman (1930)
    - Rubinstein (1939)
    - Indjic (1988)
  - Graphs of
    - inter-onset interval for each event in score
    - dynamics of each event (scaled in dBs)

# Inter-onset interval



# Event dynamics



# Conclusions and future work

- Further scientific development (wavelet parameters, shape, dynamics measurement)
- Extension to note offset times
- Testing and refinement of user interface