JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

4aPP8: Open-Source Baby Monitor

181st Meeting of the Acoustical Society of America

Shane Lani*, Jennifer Cooper, **T.J. Flynn**, Adaleena Mookerjee*,
Jordan Schleif*, Olivia Ott

*Welcomed a new baby into their family in the last 6 months

Background

Original Proposal: Baby breathing monitor

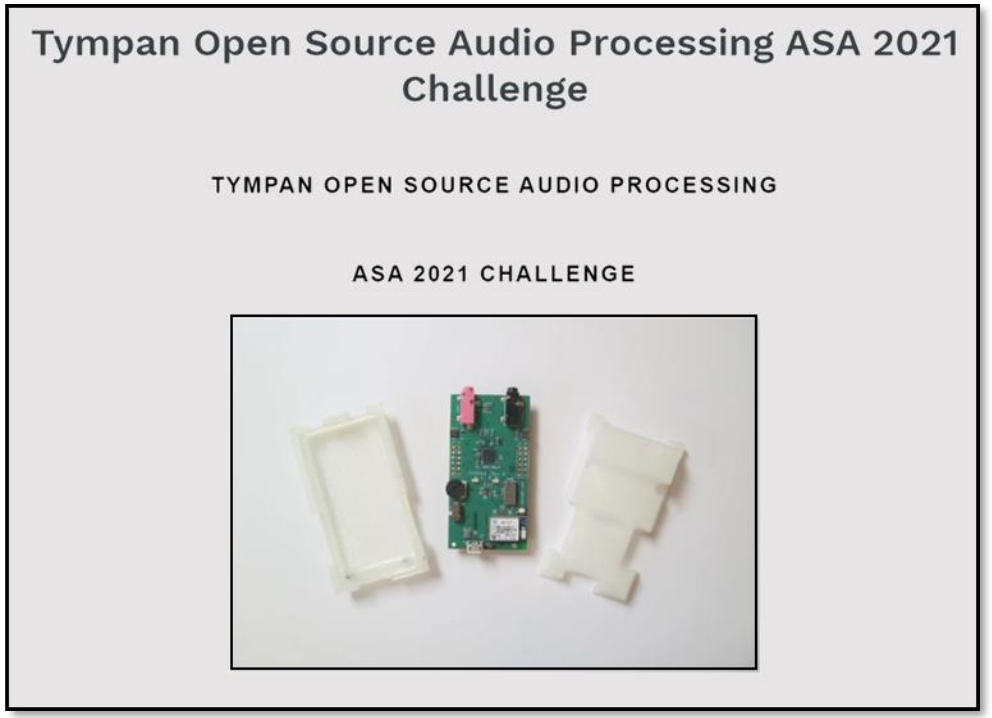
Revised Project Title: Open-Source Baby Monitor

Team Background:

Acousticians, data scientists, physicists, signal processors

Not C/C++ coders or real-time signal processing gurus

Tessa Lani



Three of the coauthors have welcomed new babies (*and their many sounds*) into their families within the last 6 months!

Motivation

Baby monitors are valuable tools providing peace-of-mind to parents and guardians,

but…

Most off-the-shelf monitors can be triggered by **anything that's loud enough**:

- talking
- traffic
- musical toys

A better baby monitor should distinguish a baby from other loud sounds!

What Does a Baby Cry Look Like?

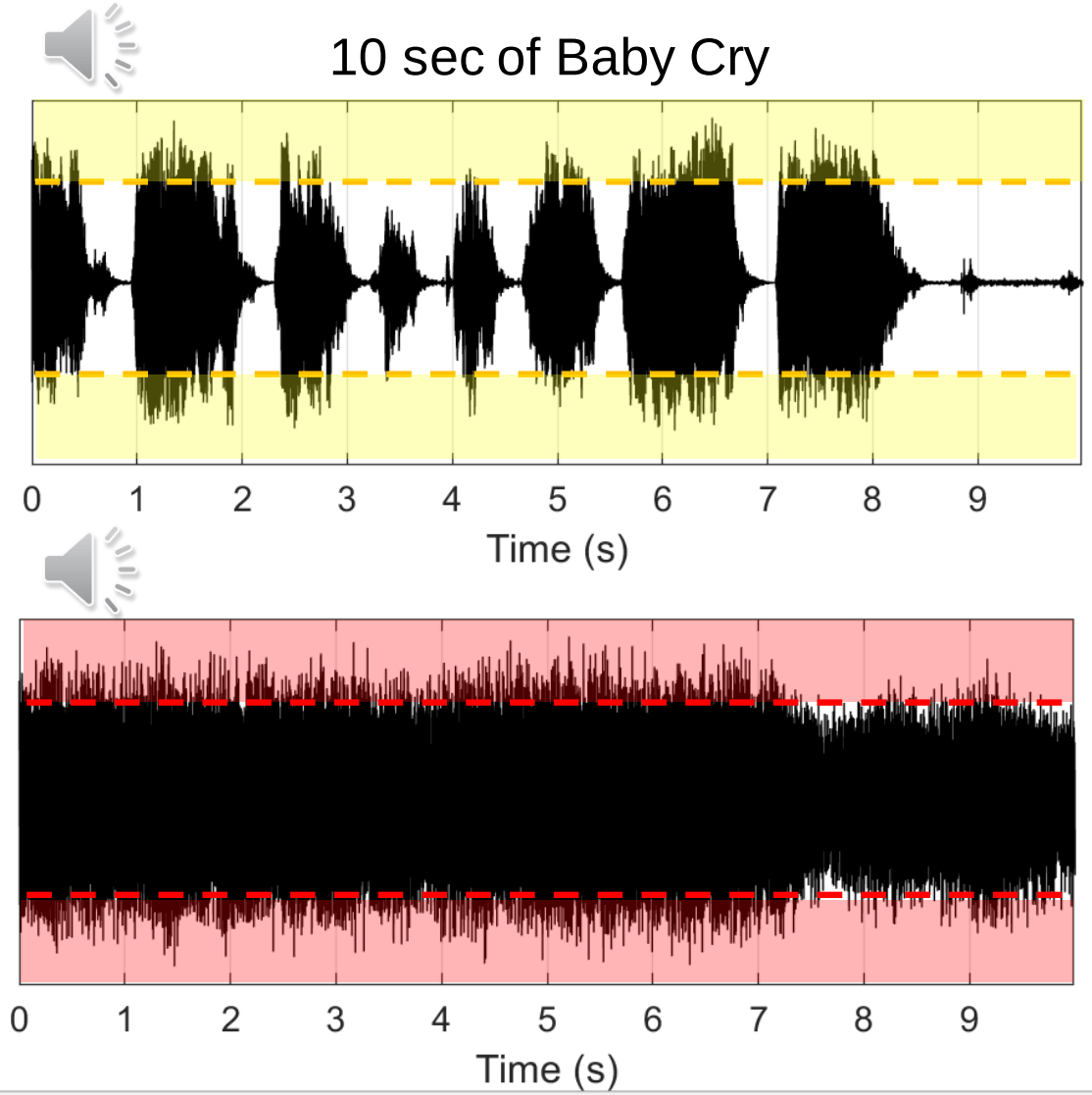
How do we decide when the baby is crying?

- We could check if the sound exceeds some **amplitude threshold**

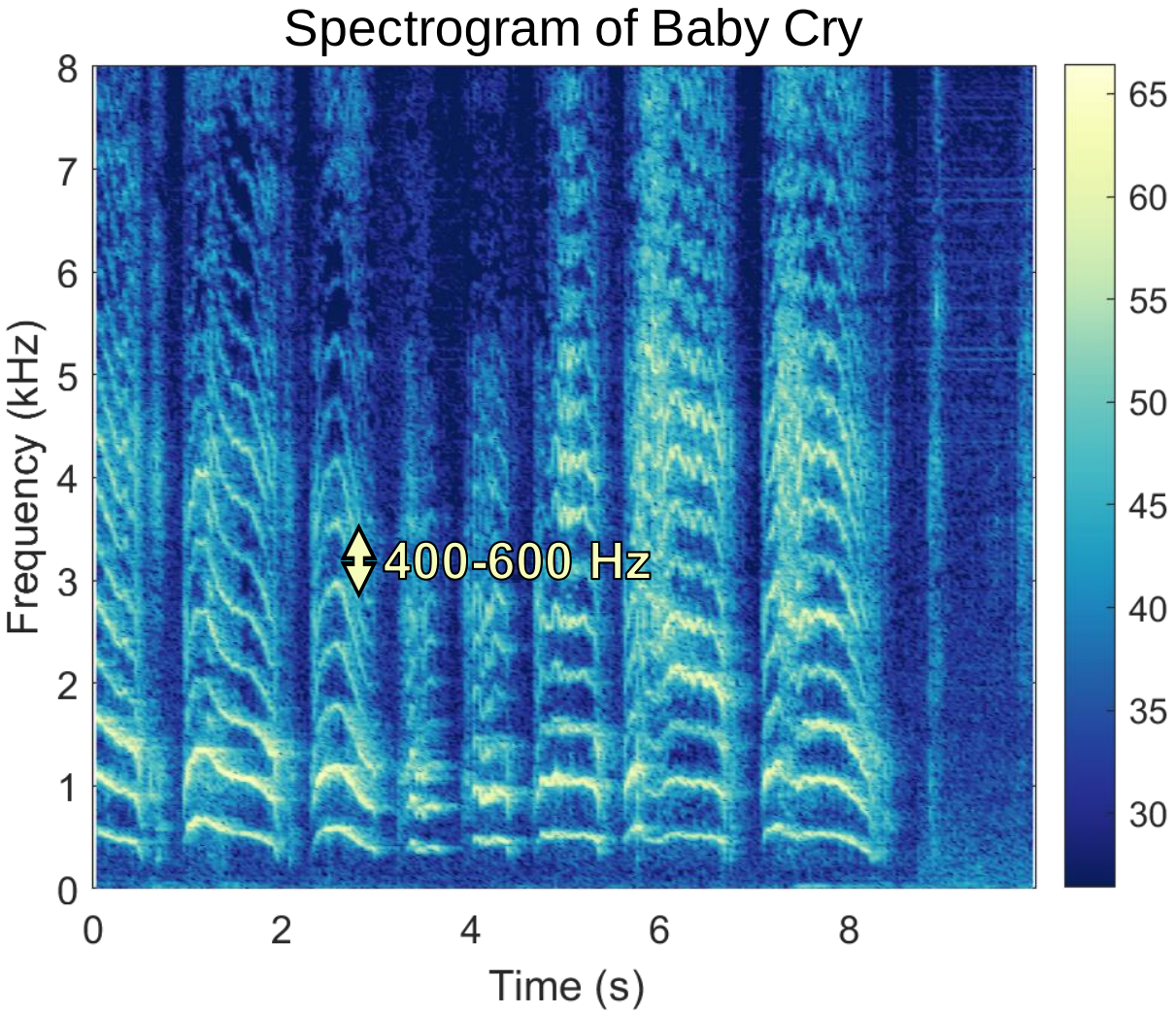
  …*but*

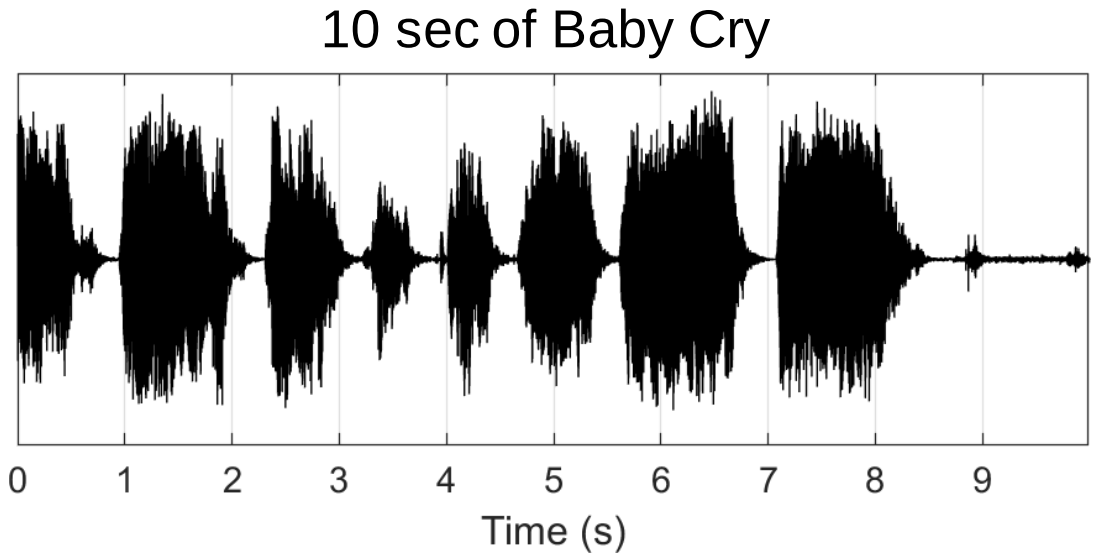
- Any sound that's loud enough will trigger the monitor, leading to many **false alarms** (aka - nuisance alarms)

10 sec of Aircraft Noise →



Frequency Domain Processing

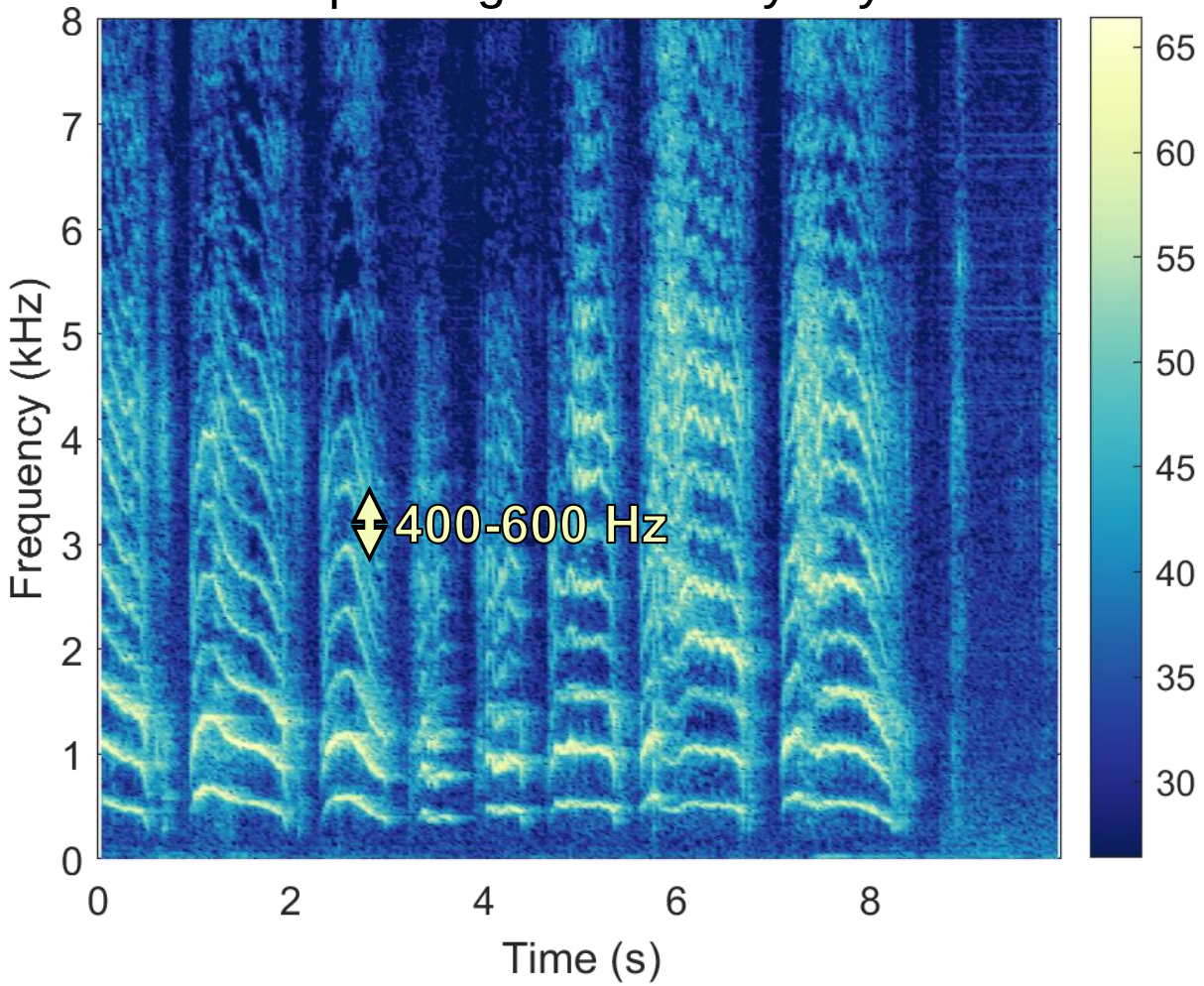




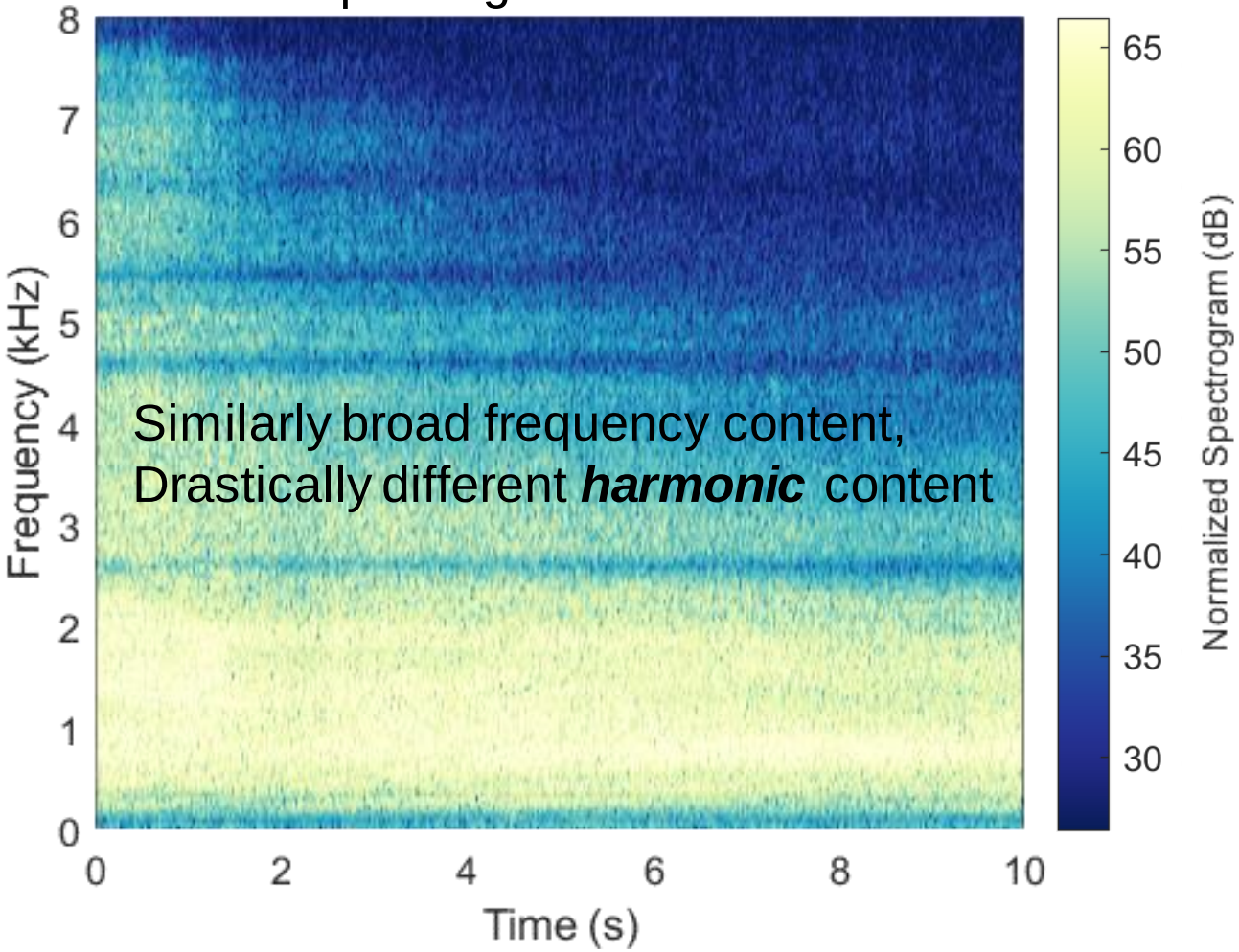
- Spans from **~500 Hz well up to 10 kHz**
- Very **Harmonic** (periodic tonals)
- Fundamental Harmonic: **400-600 Hz**
- Over **15 harmonics** are clearly visible

Frequency Domain Processing

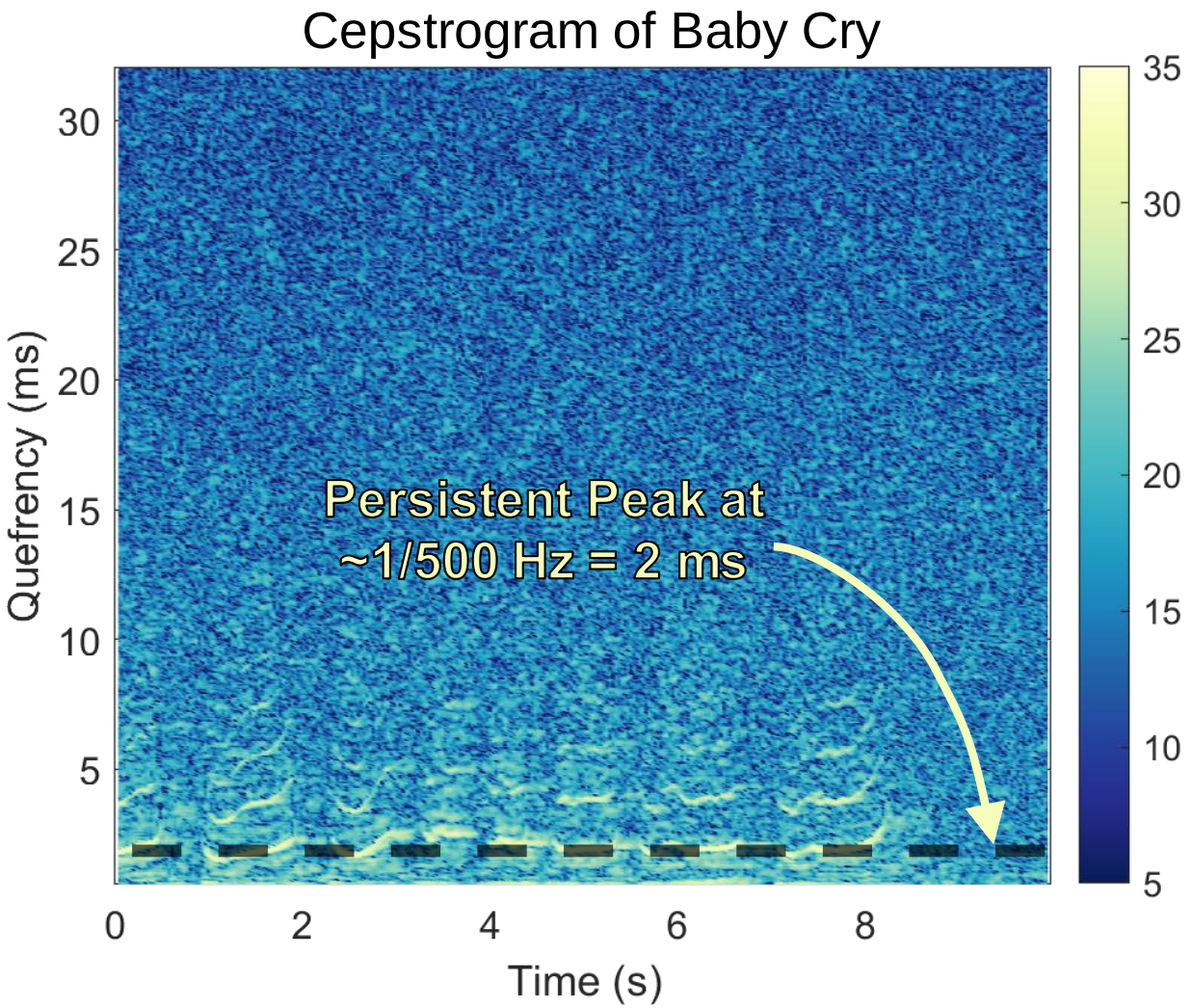
Spectrogram of Baby Cry

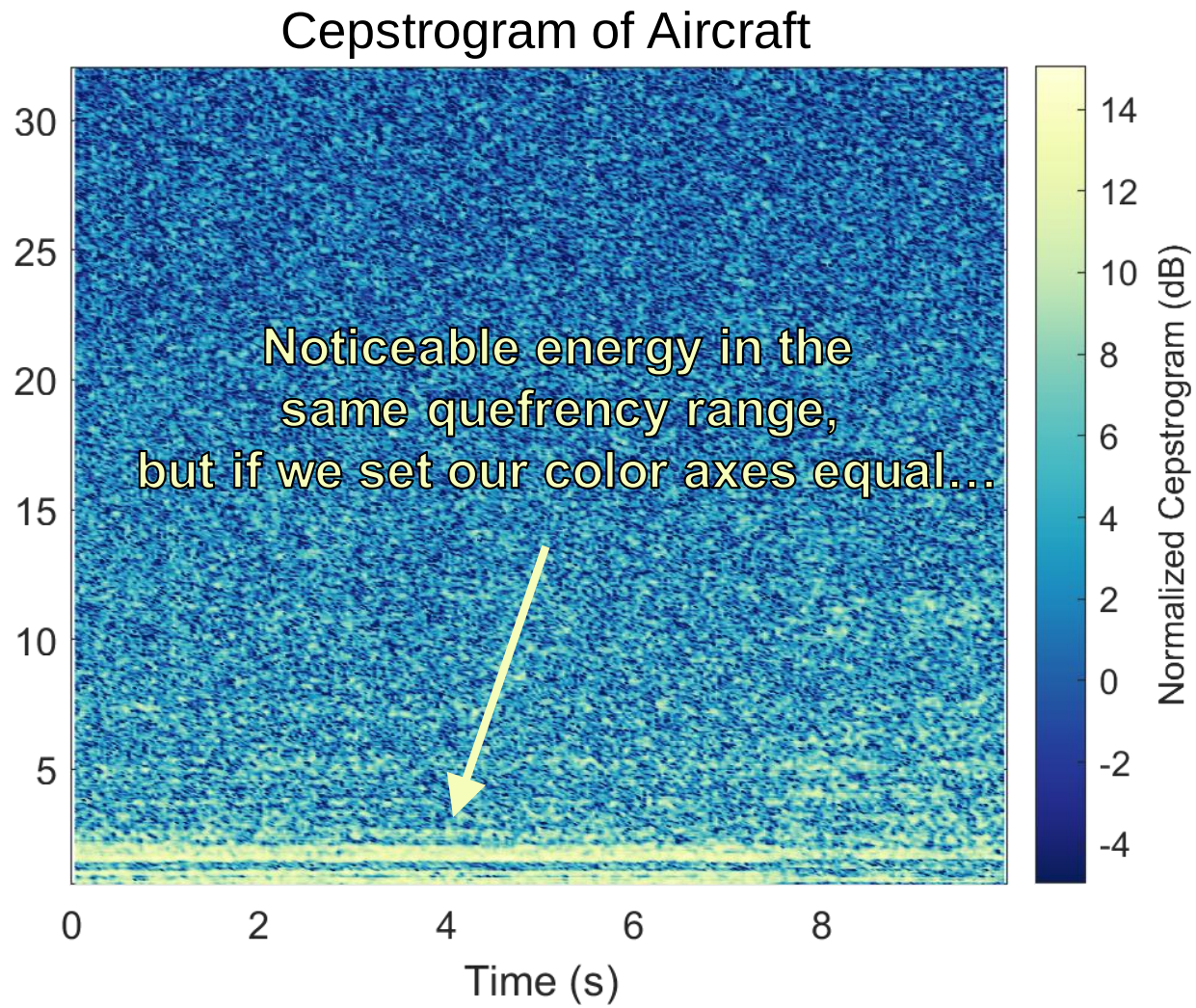


Spectrogram of Aircraft

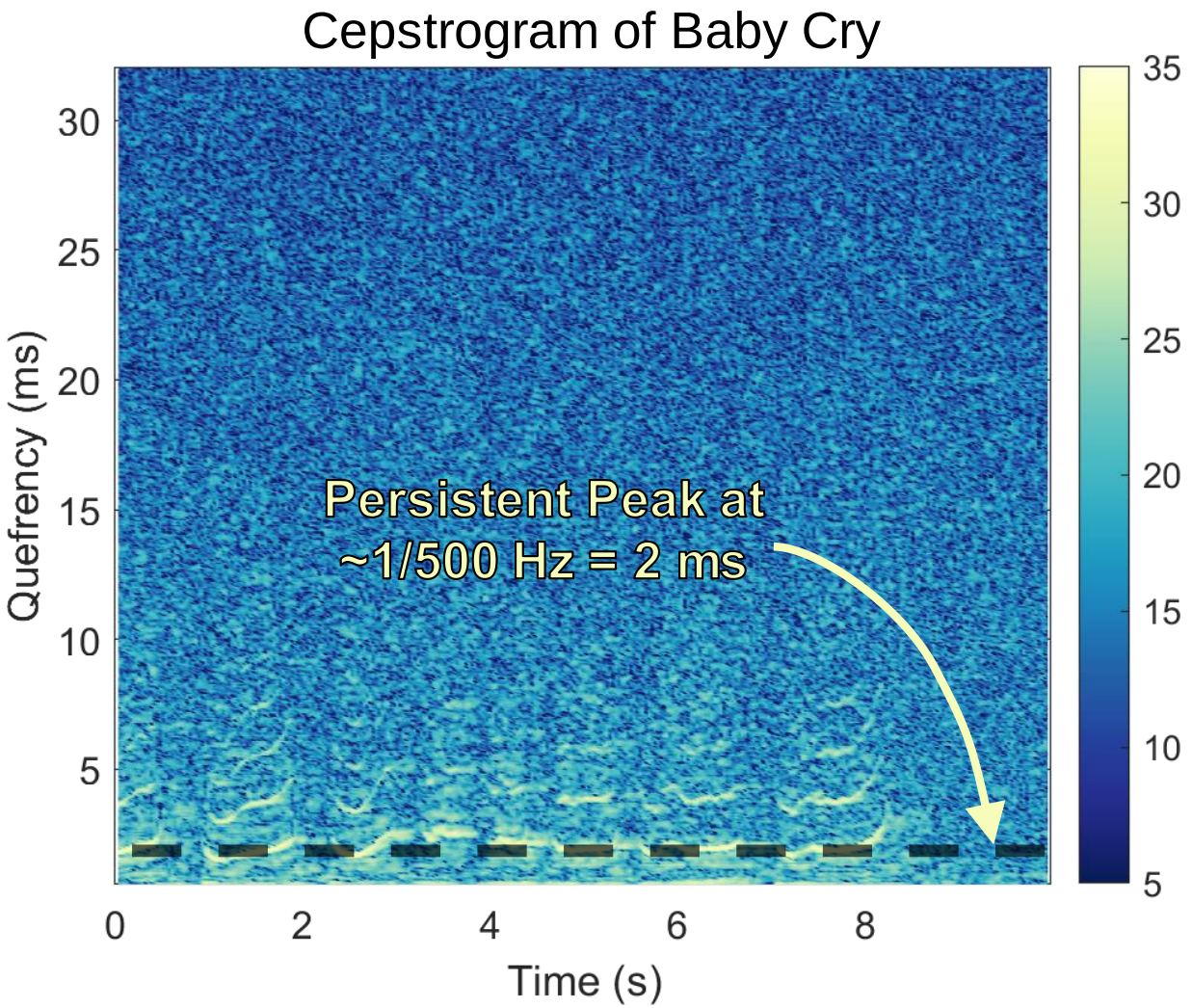


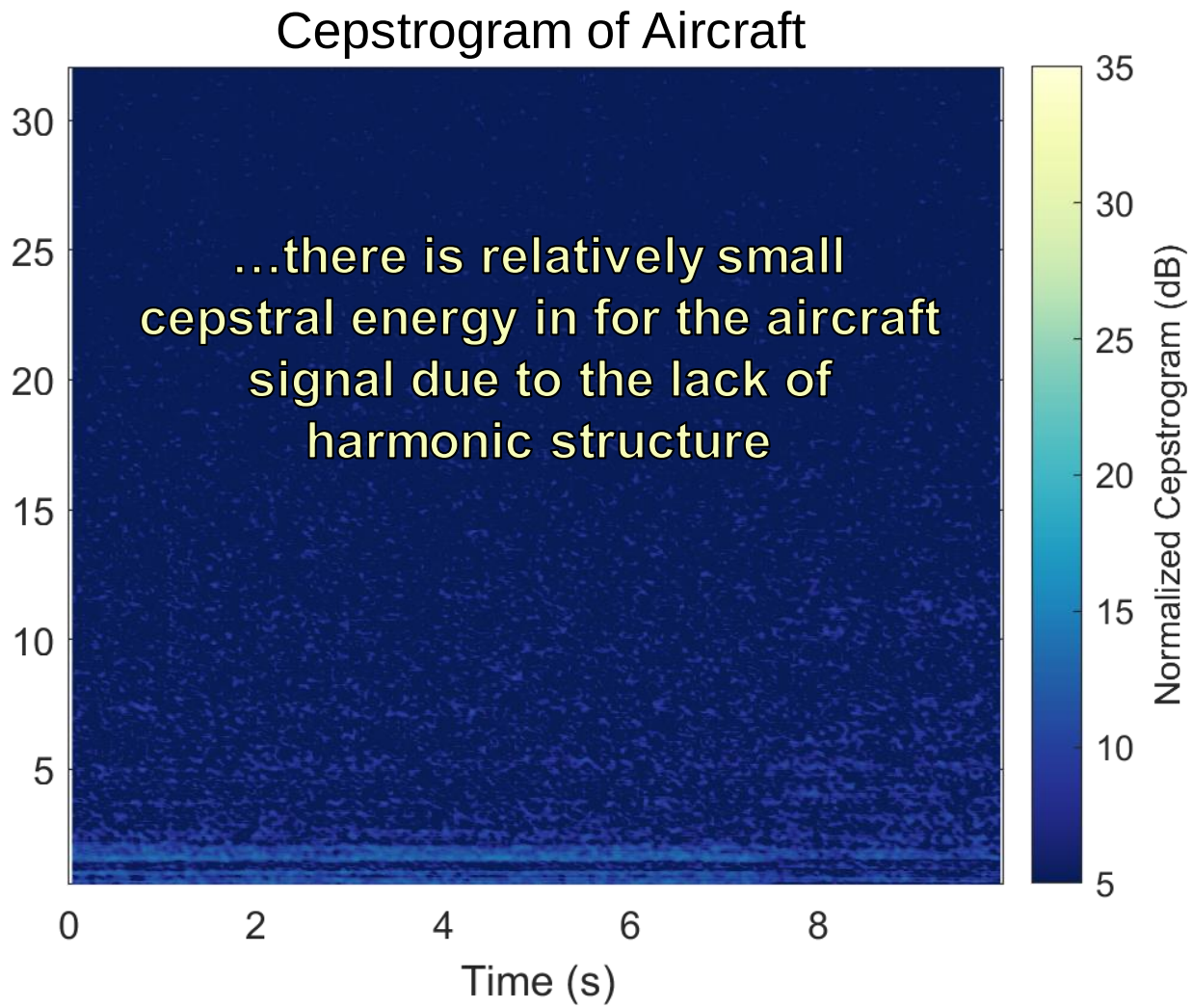
Quefrency Domain (Cepstral) Processing



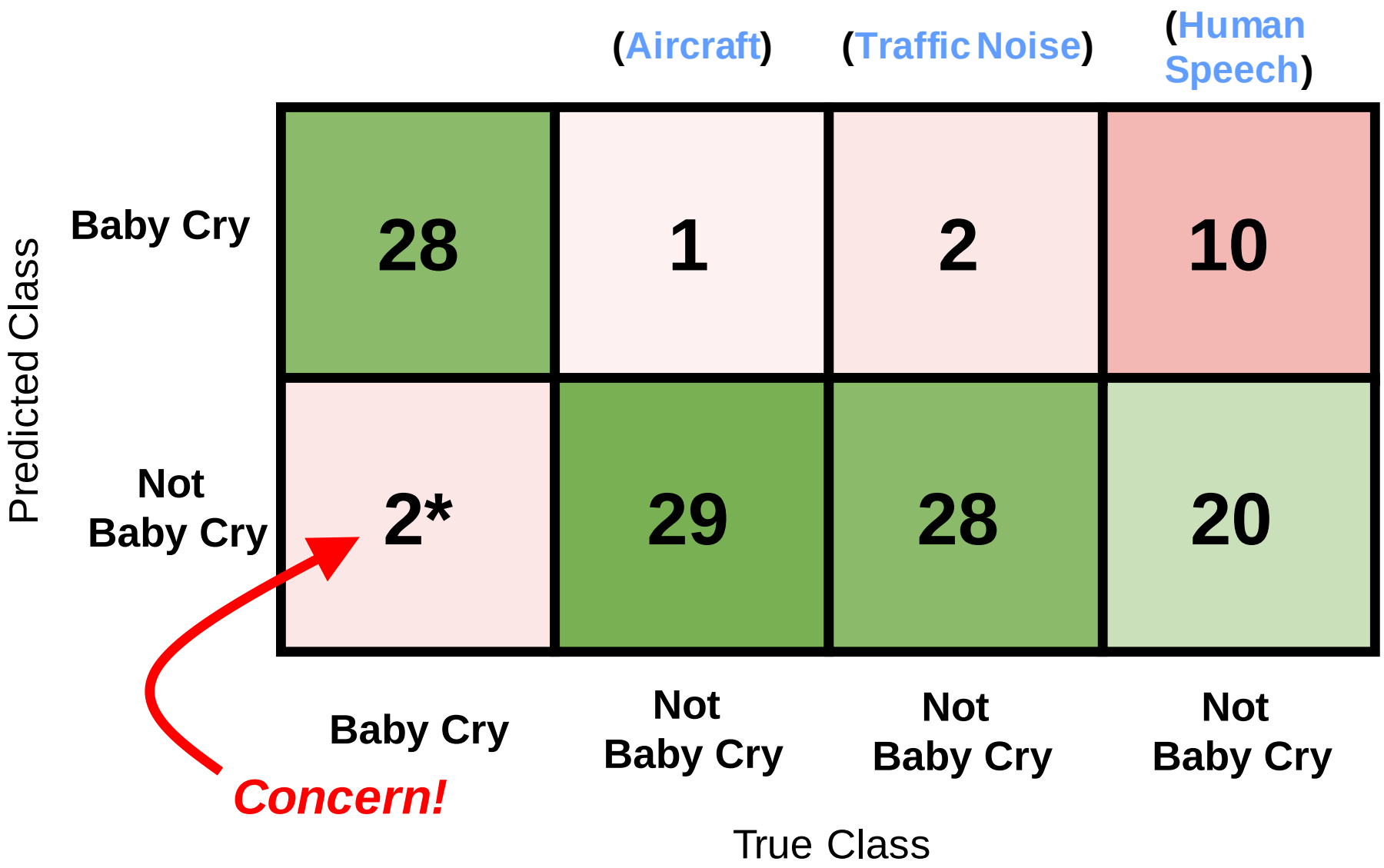


Quefrency Domain (Cepstral) Processing





Preliminary Results



The cepstral processor performed well against typical external sources of household noise, such as aircraft and traffic noise.

Distinguishing human speech from baby cries was more challenging, particularly for children's voices.

Other misclassifications:

- *Whistling*
- *Guitar (upper register)*
- *Angry cat meows*

Conclusions

- Proof-of-concept for a ‘smart’ baby monitor has been shown using the **Tympan Rev-E** open-source development platform.

- **Cepstral processing** is reasonably effective for distinguishing baby-like sounds from non-baby-like sounds. However, human speech can still be misclassified as ‘crying baby’ in many instances, particularly for children.

- There are **many** avenues for improving this technology:
  - Time domain processing (Looking for rhythms, like hiccoughs)
  - Looking at multiple metrics simultaneously
  - Machine Learning (Transferring image classification techniques to audio applications)

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

JOHNS HOPKINS

APPLIED PHYSICS LABORATORY

Conceptualization of Baby Monitor

Tympan REV-E

Audio capture (real time)

Decision Switch

Processing suggests **YES** baby is crying

Processing suggests **NO** baby is crying

Bluetooth®

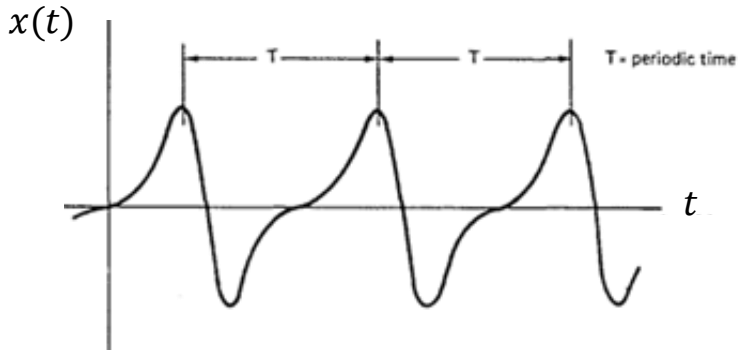
Audio transfer (real time) to speaker

Vibration/Flash Update to phone

Keep Listening

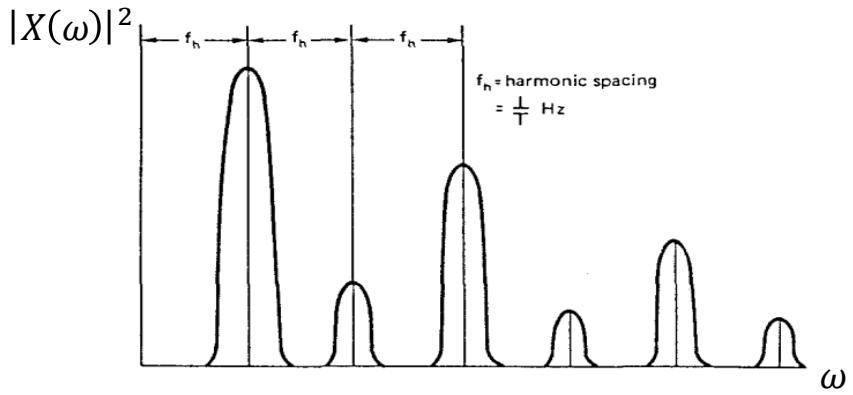
Quefrency Domain (Cepstral) Processing

Time Domain: $x(t)$



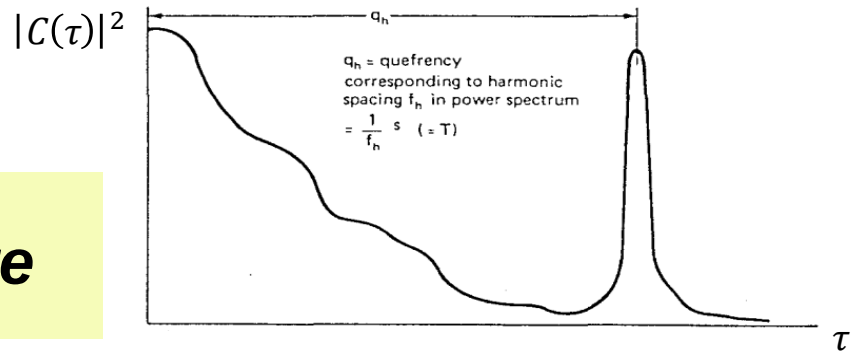
Frequency Domain: $X(\omega) = \mathcal{F}\big(x(t)\big) \equiv \int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt$

Power Spectrum: $|X(\omega)|^2$



Quefrency Domain: $C(\tau) = \mathcal{F}^{-1}\left(\log\left(\left|\mathcal{F}\big(x(t)\big)\right|^2\right)\right)$

Power Cepstrum: $|C(\tau)|^2$



Cepstra let us focus on specific *harmonic structure*

Figures referenced from R.B. Randall, 'BKSV Cepstrum Analysis & Gearbox Fault Diagnosis', 1977.