# Characterizing musical signals with Wigner-Ville interferences

William DeMeo williamdemeo@yahoo.com

ICMC 2002 Göteborg, Sweden

www.harmonicanalysis.com

www.incunabula.us

# OUTLINE

1. *Musical Signal Decomposition*
   find representation which facilitates a measure of perceived qualities.

   Review well-known decomposition method: "matching pursuit."

2. *Perceived Qualities*
   · melodic dissonance
   · sensory dissonance

3. *Wigner-Ville interference*
   specious signal energy or revealing of perceived qualities?

# PREVIOUS WORK

**consonance/dissonance analysis**

· Helmholtz, *On the Sensation of Tone*, 1877.

· Plomp & Levelt, JASA 1965.

· Sethares, Springer 1997.

**matching pursuit decomposition**

· Mallat & Zhang, IEEE Trans SP, 1993.

· Gribonval, Depalle, Rodet, Bacry, Mallat, ICMC Proceedings, 1996.

*et* many *al*...

# INTRODUCTION

Let $x(t)$ be a musical signal. Measure "dissonance" of $x$ at time instant $t = t_i$ as a function of estimated frequency components of $x$ at $t_i$.

$$\begin{aligned} D(t_i) &= \text{ dissonance of } x \text{ at } t_i \\ &= \text{ fn of frequencies at } t_i \\ &= D(t_i, \alpha_0, \alpha_1, \ldots) \end{aligned}$$

Can account for "instantaneous" dissonance. Cannot account for melodic contour.

## PERCEIVED QUALITIES

- **Melodic Consonance (CDC-1):**
  depends on melodic context; must be fn of $x(t_i)$ as well as $x(t_j)$ for $j \neq i$; accounts for relations among pitches sounded at different times.

- **Sensory Consonance (CDC-5):**
  equates consonance with smoothness/absence of beats; dissonance with roughness/presence of beats; accounts for relations among pitches sounded simultaneously.

## TIME-FREQUENCY REPRESENTATION

**Wigner-Ville Transform (WVT)**:
correlates a signal with a time and frequency shifted version of itself.

$$\mathsf{W}_x(t,\nu) = \int x\left(t+\tfrac{\tau}{2}\right) x^*\left(t-\tfrac{\tau}{2}\right) e^{-i2\pi\tau\nu}\, d\tau$$

It is the Fourier transform of

$$\phi_x(t,\tau) = x^*\left(t-\tfrac{\tau}{2}\right) x\left(t+\tfrac{\tau}{2}\right)$$

**Covariance Property:**
If energy of $x$ is concentrated around $(t_0, \nu_0)$, then energy of $\mathsf{W}_x$ is centered at $(t_0, \nu_0)$ with time-frequency spread equal to that of $x$.

**Cross Wigner-Ville Transform:**

$$\mathsf{W}_{xy}(t,\nu) = \int x\left(t + \tfrac{\tau}{2}\right) y^*\left(t - \tfrac{\tau}{2}\right) e^{-i2\pi\nu\tau}\, d\tau$$

$$\mathsf{W}_{yx}(t,\nu) = \int y\left(t + \tfrac{\tau}{2}\right) x^*\left(t - \tfrac{\tau}{2}\right) e^{-i2\pi\nu\tau}\, d\tau$$

## INTERFERENCE STRUCTURE

WVT does not conform to linear superposition; instead,

$$\begin{aligned} \mathsf{W}_{x+y}(t,\nu) &= \mathsf{W}_x(t,\nu) + \mathsf{W}_y(t,\nu) \\ &\quad + \mathsf{W}_{xy}(t,\nu) + \mathsf{W}_{yx}(t,\nu) \end{aligned}$$

**Interference term:**

$$\begin{aligned} I_{xy}(t,\nu) &= \mathsf{W}_{xy}(t,\nu) + \mathsf{W}_{yx}(t,\nu) \\ &= 2\,\mathrm{Re}\,[\mathsf{W}_{xy}(t,\nu)] \end{aligned}$$

creates non-zero values at interesting locations of the time-frequency plane.

More generally, the composite signal

$$x(t) = \sum_{n=1}^{N} a_n x_n(t)$$

has WVT

$$\mathsf{W}_x(t,\nu) = \sum_{n=1}^{N} |a_n|^2 \, \mathsf{W}_{x_n}(t,\nu)$$

$$+ 2 \sum_{n=1}^{N-1} \sum_{k=n+1}^{N} \mathrm{Re}\left[a_n a_k^* \, \mathsf{W}_{x_n x_k}(t,\nu)\right]$$

## WEYL-HEISENBERG OPERATOR

Consider elementary signal $x \in L(\mathbb{Z}/N)$
· discrete, periodic function with period $N$,
· defined on group $\mathbb{Z}/N \simeq \{0, 1, \ldots, N-1\}$.

**Weyl-Heisenberg operator:**
Define $H : \mathbb{Z}/N \times \mathbb{Z}/N \to L(\mathbb{Z}/N)$ by

$$\begin{aligned} H(\mathbf{a})x(n) &= x_{\mathbf{a}}(n) \\ &= x(n - a_1)\, e^{i2\pi a_2 n/N} \end{aligned}$$

for any $\mathbf{a} = (a_1, a_2) \in \mathbb{Z}/N \times \mathbb{Z}/N$.

## EXAMPLE

Let $x(t) = e^{i2\pi\nu_m t}$

$$\mathbf{a} = (0, -\tfrac{\Delta\nu}{2}) \;\Rightarrow\; x_{\mathbf{a}}(t) = e^{i2\pi(\nu_m - \frac{\Delta\nu}{2})t}$$

$$\mathbf{b} = (0, \tfrac{\Delta\nu}{2}) \;\Rightarrow\; x_{\mathbf{b}}(t) = e^{i2\pi(\nu_m + \frac{\Delta\nu}{2})t}$$

Composite signal is

$$x_{\mathbf{a}}(t) + x_{\mathbf{b}}(t) = 2\cos(2\pi\tfrac{\Delta\nu}{2}t)\, e^{i2\pi\nu_m t}$$

**PHYSICAL INTERPRETATION:**

Composite signal has WVT

$$\mathsf{W}_{x_{\mathbf{ab}}}(t,\nu) = \delta(\nu - (\nu_m - \tfrac{\Delta\nu}{2})) + \delta(\nu - (\nu_m + \tfrac{\Delta\nu}{2}))$$

$$+ \, \delta(\nu - \nu_m)\, 2\cos(2\pi\Delta\nu t)$$

When components $x_{\mathbf{a}}(t)$ and $x_{\mathbf{b}}(t)$ are close in frequency, cosine term is slowly varying.

Composite signal viewed as simple tone with

· frequency $\nu_m$,

· modulated amplitude envelope,

· modulation frequency $\Delta\nu$.

"beats" refer to such amplitude modulations.

# MATCHING PURSUIT (MP)

Mallat & Zhang, *Matching Pursuits with Time-Frequency Dictionaries*, IEEE Trans. SP, 1993.

MP decomposes signal over elements of dictionary $\mathcal{D} = \{g_\gamma\}_{\gamma \in \Gamma}$.

**MP Dictionary:**
Start with Gaussian window

$$f(t) = 2^{1/4} e^{-\pi t^2}$$

Dictionary comprises periodic, scaled, translated, and modulated versions of $f$.

$\mathrm{Per}_B\, f$ denotes *periodization* of $f$ over $B$.

For $n \in \mathbb{Z}$ define discrete $N$-periodic window

$$g(n) = \mathrm{Per}_{N\mathbb{Z}}\, f(n) = \sum_{m \in N\mathbb{Z}} f(n+m)$$

Scaling by $s$,

$$g_s(t) = \frac{1}{\sqrt{s}} g\left(\frac{t}{s}\right)$$

Let $\gamma = (s, a_1, a_2)$ where

$$\begin{aligned} s = &\ \text{scaling factor} \\ a_1 = &\ \text{time translation} \\ a_2 = &\ \text{frequency modulation} \end{aligned}$$

$$g_\gamma(t) = g_s(t - a_1)\langle t, a_2\rangle$$

$$= \frac{1}{\sqrt{s}} g\left(\frac{t - a_1}{s}\right) e^{i2\pi a_2 t}$$

$g_\gamma(t)$ defines a typical atom in the dictionary, with energy concentrated in a neighborhood of $(a_1, a_2)$, whose size is proportional to $s$.

Let discrete signal $x$ have length $N = 2^{K+1}$.

Select dictionary parameters as follows:

$$s = 2^j, \qquad j \in \{1, 2, \ldots, K\}$$

$$a_1 \in L_1\mathbb{Z}/N, \qquad L_1 = 2^{j-1}$$

where

$$L_1\mathbb{Z}/N = \{0, L_1, 2L_1, \ldots, (M_1 - 1)L_1\}$$

and $$N = L_1 M_1$$

$\therefore$ translations separated by $\frac{s}{2} = 2^{j-1}$ samples.

Suppose modulation parameters are

$$a_2 \in M_1\mathbb{Z}/N = (L_1\mathbb{Z}/N)_*$$

then $(a_1, a_2)$ ranges over

$$L_1\mathbb{Z}/N \times M_1\mathbb{Z}/N = L_1\mathbb{Z}/N \times (L_1\mathbb{Z}/N)_*$$

a *critical sampling subgroup* of $\mathbb{Z}/N \times \mathbb{Z}/N$.

Instead choose $M_2 = 2^{K-j}$, and let $(a_1, a_2)$ range over *integer oversampling subgroup*

$$\Delta_s = L_1\mathbb{Z}/N \times M_2\mathbb{Z}/N$$

**Weyl-Heisenberg (W-H) system:**
For each $s$, $\langle g_s, \Delta_s \rangle$ defines a Weyl-Heisenberg system.

To the set of $K$ W-H systems add
· complex exponentials (Fourier basis)
· $N$ discrete Diracs

**Matching Pursuit (MP)**
iteratively decompose signal over $\mathcal{D}$ as follows:

1. set $R^0 x(t) = x(t)$; assume $R^n x$ computed.

2. choose $g_{\gamma_n}$ which "matches" $R^n x$:

$$|C(R^n x, g_{\gamma_n})| = \sup_{\gamma \in \Gamma} |C(R^n x, g_\gamma)|$$

3. decompose residue $R^n x$ as

$$R^n x(t) = C(R^n x, g_{\gamma_n}) g_{\gamma_n}(t) + R^{n+1} x(t)$$

4. repeat from step 2.

$C$ is a "correlation" function. For usual inner product $C(R^n x, g_{\gamma_n}) = \langle R^n x, g_{\gamma_n} \rangle$ one can show

$$\|R^n x\| \downarrow 0 \text{ exponentially with } n$$

MP results in decomposition

$$x(t) = \sum_{n=0}^{\infty} \langle R^n x, g_{\gamma_n} \rangle g_{\gamma_n}(t)$$

with WVT

$$W_x(t, \nu) = \sum_{n=0}^{\infty} |\langle R^n x, g_{\gamma_n} \rangle|^2 \, \mathsf{W}_{g_{\gamma_n}}(t, \nu) +$$

$$\sum_{m,n} \langle R^m x, g_{\gamma_m} \rangle \langle R^n x, g_{\gamma_n} \rangle^* \, \mathsf{W}_{g_{\gamma_m} g_{\gamma_n}}(t, \nu)$$

Most studies define signal energy as

$$E_x(t,\nu) = \sum_{n=0}^{\infty} |\langle R^n x, g_{\gamma_n}\rangle|^2 \, \mathsf{W}_{g_{\gamma_n}}(t,\nu)$$

· only 1st term of WVT appears;
· assumption: 1st term accounts for energy of "true" signal components.

For $W_x(t,\nu) = E_x(t,\nu) + I_x(t,\nu)$ we define

$$\begin{aligned} E_x(t,\nu) &= \text{ "signal energy"} \\ I_x(t,\nu) &= \text{ "interference energy"} \end{aligned}$$

Cost to compute cross WVT roughly same as cost to compute WVT.

For a $N$-atom MP, must compute $N(N-1)/2$ cross WVT's.

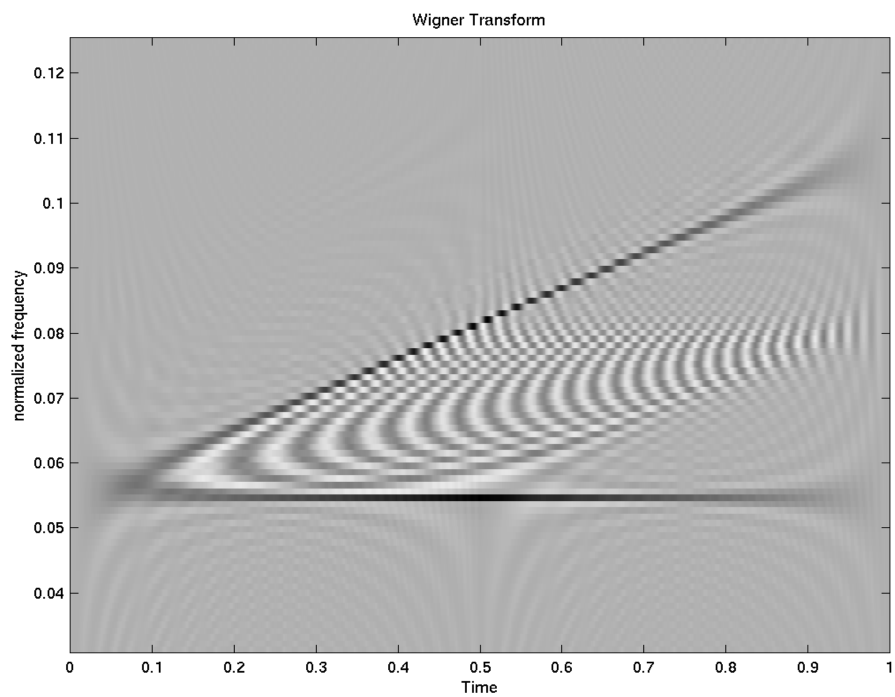
Assuming:
· fast implementation based on FFT
· MP expansion has fewer than 1000's of atoms
$\Rightarrow$ computational burden still "manageable."

EXAMPLE

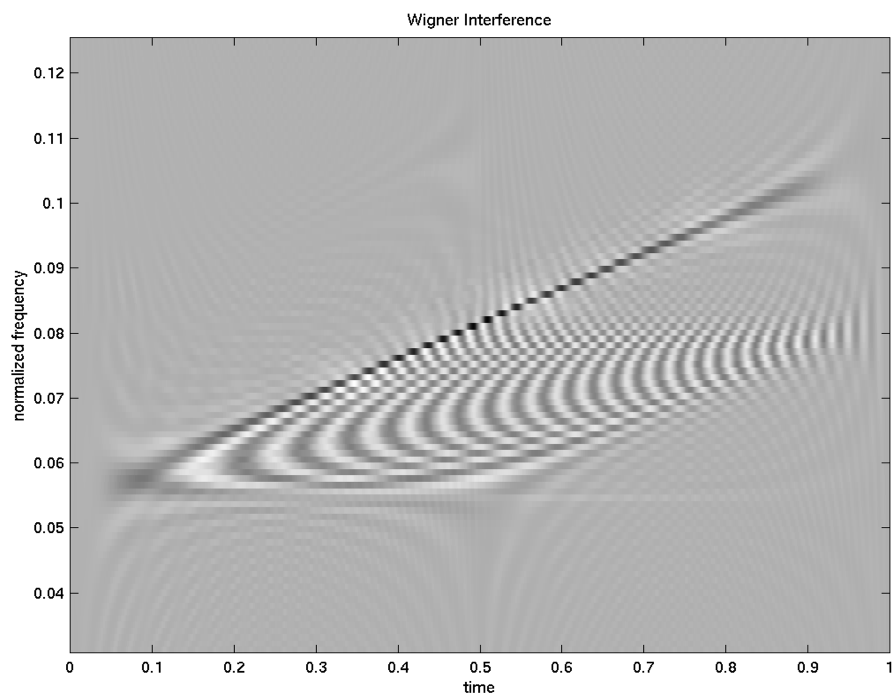
Construct signal by adding constant tone to "linear chirp."

Start at equal frequencies; chirp frequency increases linearly until double that of constant tone.

Demonstrates how constant tone interacts with frequencies in the octave above it.




WVT of constant frequency modulation plus linear chirp.




Interference energy of constant frequency modulation plus linear chirp.

## DISSONANCE MEASURES

Define

$$\phi_{g_{\gamma m} g_{\gamma n}}(t,\tau) = g_{\gamma_m}\left(t+\tfrac{\tau}{2}\right) g^*_{\gamma_n}\left(t-\tfrac{\tau}{2}\right)$$

$\mathsf{W}_{g_{\gamma m} g_{\gamma n}}$ is Fourier transform of $\phi_{g_{\gamma m} g_{\gamma n}}(t,\tau)$

$\therefore$ inverse FT of $\mathsf{W}_{g_{\gamma m} g_{\gamma n}}$ is $\phi_{g_{\gamma m} g_{\gamma n}}$

$$\phi_{g_{\gamma m} g_{\gamma n}}(t,\tau) = \int \mathsf{W}_{g_{\gamma m} g_{\gamma n}}(t,\nu) e^{i2\pi\nu\tau}\, d\nu$$

Integrating $\mathsf{W}_{g_{\gamma m} g_{\gamma n}}(t, \nu)$ over all $\nu$ is equivalent to evaluating $\phi_{g_{\gamma m} g_{\gamma n}}$ at $\tau = 0$

$$\begin{aligned}\int \mathsf{W}_{g_{\gamma m} g_{\gamma n}}(t, \nu)\, d\nu &= \phi_{g_{\gamma m} g_{\gamma n}}(t, 0) \\ &= g_{\gamma_m}(t) g^*_{\gamma_n}(t)\end{aligned}$$

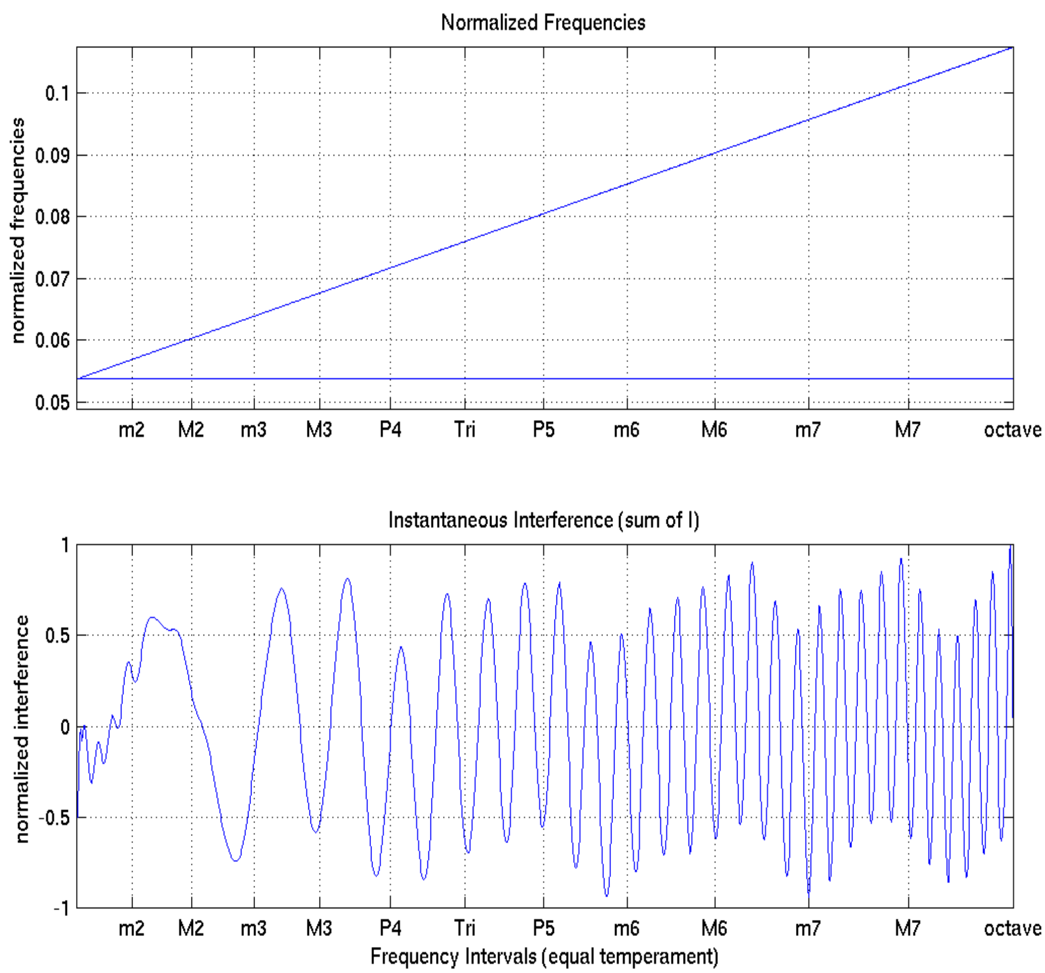
**1st Dissonance Measure:**
"instantaneous interference"

$$\mathcal{I}_x(t) = \int I_x(t, \nu)\, d\nu$$

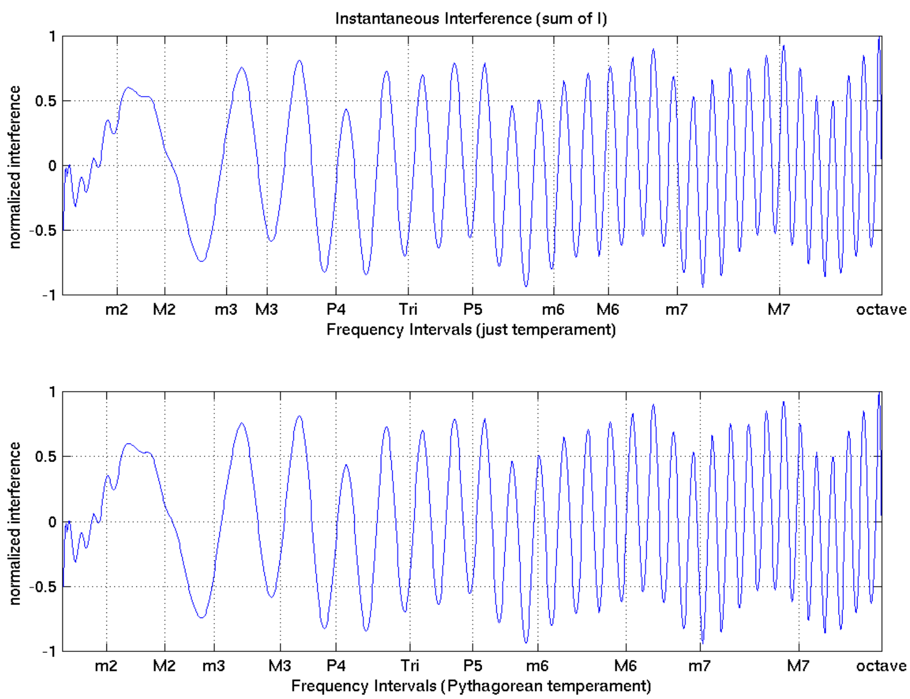
$$= \sum_{m,n} \langle R^m x, g_{\gamma_m} \rangle \langle R^n x, g_{\gamma_n} \rangle^* \int \mathsf{W}_{g_{\gamma m} g_{\gamma n}}(t, \nu)\, d\nu$$

$$= \sum_{m,n} \langle R^m x, g_{\gamma_m} \rangle \langle R^n x, g_{\gamma_n} \rangle^* g_{\gamma_m}(t) g^*_{\gamma_n}(t)$$

Normalized instantaneous frequencies (top); Instantaneous interference $\mathcal{I}_x(t)$ (bottom)



$\mathcal{I}_x(t)$ with time axes delimited by just (top) and Pythagorean (bottom) tunings.



$\mathcal{I}_x(t)$ essentially sums $I_x(t,\nu)$ over $\nu$.
$\Rightarrow$ measures interferences only at time $t$.

**2nd Dissonance Measure:**
Fourier transform the interference energy.

$$\mathcal{I}_x(t,\tau) = \int I_x(t,\nu) e^{i2\pi\nu\tau}\, d\nu$$

$$= \sum_{m,n} \langle R^m x, g_{\gamma_m}\rangle \langle R^n x, g_{\gamma_n}\rangle^* \phi_{g_{\gamma_m} g_{\gamma_n}}(t,\tau)$$

$\mathcal{I}_x(t,\tau)$ leads to measure of interference among signal components at different points in time.

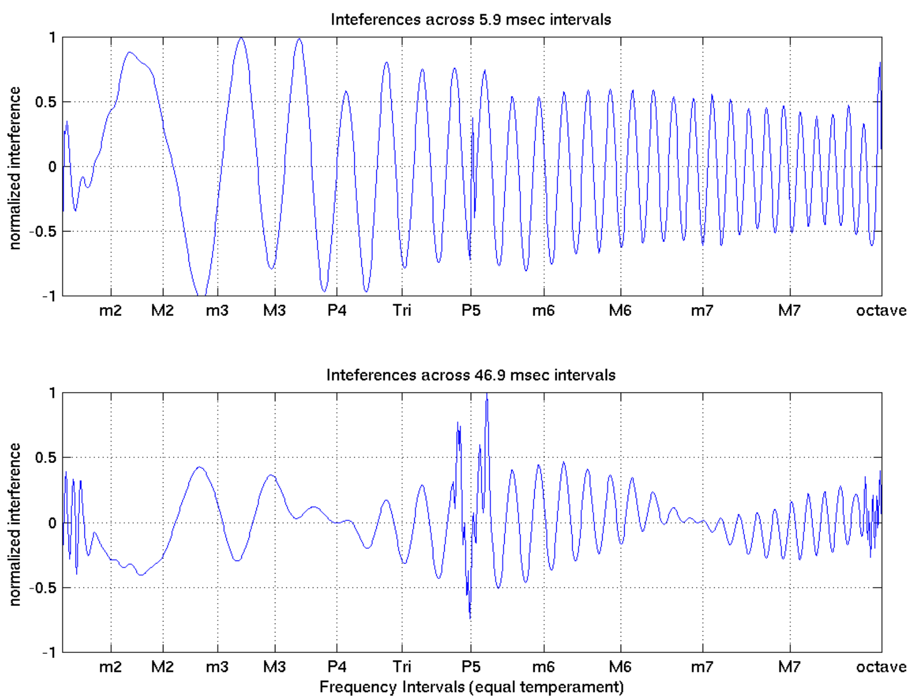
Measure interferences separated by up to $\tau_0$ units of time:

$$\mathcal{I}_x^{\tau_0}(t) = \int_0^{\tau_0} \mathcal{I}_x(t,\tau)\, d\tau$$

$$= \sum_{m,n} \langle R^m x, g_{\gamma_m}\rangle \langle R^n x, g_{\gamma_n}\rangle^* \times \int_0^{\tau_0} g_{\gamma_m}\left(t + \tfrac{\tau}{2}\right) g_{\gamma_n}^*\left(t - \tfrac{\tau}{2}\right)\, d\tau$$

$\tau_0$ determines maximum time interval across which to measure interferences.

$\mathcal{I}_x^{\tau_0}$ for two values of $\tau_0$.




$\mathcal{I}_x^{\tau_0}$ for $\tau_0 = 6$ milliseconds (top), and $\tau_0 = 47$ milliseconds (bottom).

**General Dissonance Measure:**

Put distribution $\mu$ on domain of time intervals; $\mu$ describes relative importance of interference among time intervals.

Define

$$\mathcal{I}_x^{\mu}(t) = \int_{-\infty}^{\infty} \mathcal{I}_x(t,\tau)\, d\mu(\tau)$$

1st and 2nd measures are special cases.

$$\mathcal{I}_x^{\tau_0} \;\leftrightarrow\; d\mu(\tau) = \chi_{[0,\tau_0)}(\tau)\, d\tau$$

$$\mathcal{I}_x \;\leftrightarrow\; \mu(\tau) = \delta(\tau)$$

## SUMMARY & CONCLUSIONS

- defined "interference energy" function $I_x(t,\nu)$ to be sum of cross terms in atom expansion of $W_x$.

- defined measures $\mathcal{I}_x$ and $\mathcal{I}_x^{\tau_0}$ quantify interference among signal components at points and over intervals, resp.

- defined more general measure $\mathcal{I}_x^{\mu}$ employing a distribution on the domain of time differences between signal components.

- observed interesting behavior for simple composition of pure tones.

## FUTURE WORK

1. optimize atomic decomposition algorithm (e.g. employ better MP)

2. experiment with real music examples

3. experiment with various distributions $\mu$