TIME-FREQUENCY SPECTRA OF MUSIC

James S. Walker
Department of Mathematics
University of Wisconsin–Eau Claire
walkerjs@uwec.edu

Amanda J. Potts
Department of Mathematics
University of Nebraska–Lincoln

Abstract We analyze some musical passages using hanning windowed spectrograms. Preliminary results indicate that the highest peaks in the spectra for these musical passages do not obey a 1/f power law, contrary to results reported by some other workers. Further study is needed to determine the exponents for the spectra for a variety of musical passages from around the world.

Introduction

Initial work on analyzing musical signals was begun in [1], [3], [5], and [6]. The mathematical model for a musical signal $Y(t)$ proposed in [1] is the finite sum

$$Y(t) = \sum_j V_j h\left(\frac{t-\tau_j}{\sigma_j}\right) \cos \lambda_j (t - \tau_j) \tag{1}$$

where $h(\cdot)$ is a taper function of duration σ_j. Formula (1) has constants V_j (amplitude), τ_j (offset in time), and λ_j (frequency).

The Fourier transform $\widehat{Y}(\lambda)$ of $Y(t)$ is then

$$\widehat{Y}(\lambda) = \sum_j \frac{1}{2} V_j \sigma_j \mathrm{e}^{i\phi_j\lambda} \widehat{h}[\sigma_j(\lambda - \lambda_j)] + \sum_j \frac{1}{2} V_j \sigma_j \mathrm{e}^{-i\phi_j\lambda} \widehat{h}[\sigma_j(\lambda_j - \lambda)] \tag{2}$$

where $\widehat{h}[\sigma_j(\lambda - \lambda_j)]$ is a narrow function with peak centered on frequency λ_j. When

$$V_j^2 \sigma_j^2 \;\propto\; \frac{1}{\lambda_j}\,, \quad \text{all } j \tag{3}$$

we say that the signal Y has a $1/f$ spectrum.

In [1], statistical data are gathered to establish the claim that musical signals have $1/f$ spectra. In this paper we shall analyze several musical passages using hanning spectrograms to obtain time-frequency portraits of those signals. *Our preliminary results cast doubt on the validity of* (3) *for these musical passages.* Our musical signals include examples from non-Western music (whereas [1] examined only Western styles of music). The sound recordings we used are all available for downloading at the following website:

$$\texttt{http://www.uwec.edu/walkerjs/TimeFrSpMusic} \tag{4}$$

1. Hanning spectrograms

The basic tool we used to obtain time-frequency spectra for a musical signal is *hanning spectrograms.* For such spectrograms, a Hann windowing function $\omega(\cdot)$ is defined as follows:[1]

$$\omega(t) = \begin{cases} 0.5 + 0.5\cos(\pi t), & \text{for } |t| \leq 1.0 \\ 0, & \text{for } |t| > 1.0 \end{cases} \tag{5}$$

A hanning spectrogram is then obtained by the following three-step process.[2] *Step 1:* Discretely sample the signal $Y(t)$ at the time values $\{t_k\}$, with fixed spacing $\Delta t = t_k - t_{k-1}$, to obtain the (digital) signal $\{Y(t_k)\}$. *Step 2:* Multiply the signal $\{Y(t_k)\}$ by M sliding windows

$$\left\{ \omega\left(\frac{t_k - \tilde{t}_j}{256\Delta t} \right) \right\}, \tag{6}$$

with $\tilde{t}_j - \tilde{t}_{j-1} = 256\Delta t$. *Step 3:* Apply a 512-point FFT to each one of these windowed signals (centered on $\tilde{t}_j$):

$$\left\{ Y(t_k)\omega\left(\frac{t_k - \tilde{t}_j}{256\Delta t} \right) \right\}, \tag{7}$$

[1] For more details on Hann windows, see [4] or [7].
[2] This process is discussed in more detail in [8].

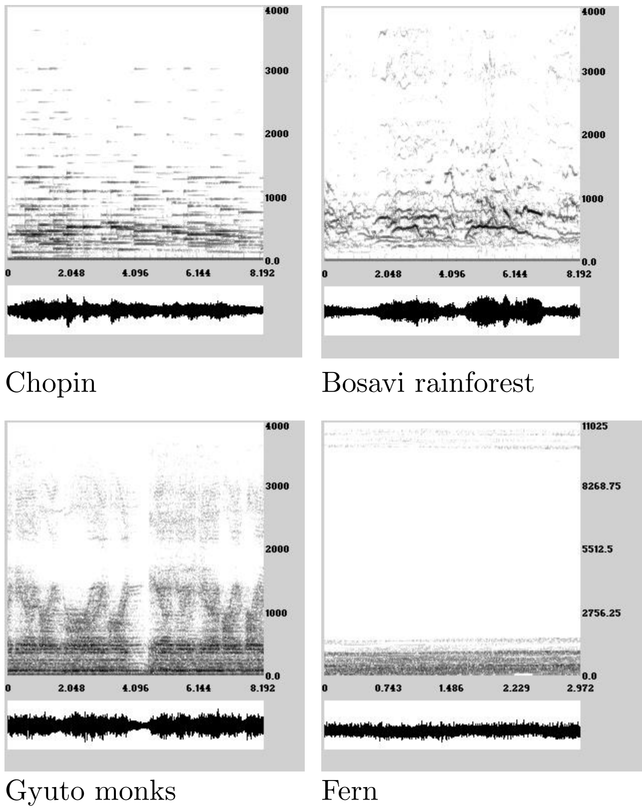


Figure 1. Dynamic power spectra for four musical passages.

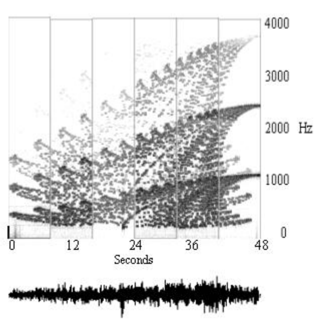


Figure 2. Complete dynamic power spectra for the Fern musical passage.

obtaining the *hanning spectrogram,*

$$\{\widehat{Y}_j(\lambda_m)\}. \tag{8}$$

With band-limited signals, if the time-values $\{t_k\}$ are close enough to each other, Steps 1 to 3 are invertible. With discretely sampled data $\{Y(t_k)\}$ to begin with (as in a `.wav` audio file), Steps 2 and 3 are reversible, i.e. the map

$$\{Y(t_k)\} \mapsto \{\widehat{Y}_j(\lambda_m)\} \tag{9}$$

is invertible.[3] When displaying these spectrograms, one typically takes the moduli-squared of all values, $\{|\widehat{Y}_j(\lambda_m)|^2\}$, which are called *dynamic power spectra.* In plotting these dynamic power spectra, one displays values on a logarithmic intensity scale with larger values plotted as darker pixels. See Fig. 1 for dynamic power spectra for four musical passages [the ones available for downloading at website (4)]. In Fig. 1, the Chopin signal is a recording of a Chopin piano passage; the Bosavi rainforest signal is a recording of some members of a South American indigenous tribe singing; the Gyuto monks signal is a recording of Buddhist monks singing; and the Fern signal is a computer generated fractal-music signal. See Fig. 2 for a longer dynamic spectra for this last passage.

In Fig. 3 we graph the FFT magnitudes $|\widehat{Y}(\lambda_m)|$ for each of these four test signals. By inspection it is easily seen that, for each signal as a whole, the FFT magnitude is *not* proportional to $1/\lambda^{0.5}$, hence the power spectrum $|\widehat{Y}(\lambda_m)|^2$ is *not* $1/f$.

In Fig. 4(a) we show the dynamic spectra for the Chopin passage. In Figures 4(b) to (d) we show the magnitudes $\{|\widehat{Y}_j(\lambda_m)|\}$ for the times $\tilde{t}_j = 1.0$, 1.5, and 2.0, respectively. The least-square curves for the log-log values of the 5 highest peaks from the FFT magnitudes $\{|\widehat{Y}_j(\lambda_m)|\}$ are also plotted in Figures 4(b) to (d). The exponents at $\tilde{t}_j = 1.0$, 1.5, and 2.0 are 2.6, 1.4, and 3.6, respectively. These exponents are not close to 0.5 as they would be if there were a $1/f$ power law in effect. In Table 1 we show the estimated exponents for $\{|\widehat{Y}_j(\lambda_m)|\}$ at times $\tilde{t}_j = 0.5$, 1.0, 1.5, 2.0, and 2.5. These exponents are widely at variance from the $1/f$-exponent of 0.5. This casts doubt on the proposed $1/f$ power law for spectra of musical signals.

[3]See [8] for more details on the inversion of (9). The map in (9) is a *Gabor transform.* See [2] for a thorough treatment of Gabor transforms and their inverses.

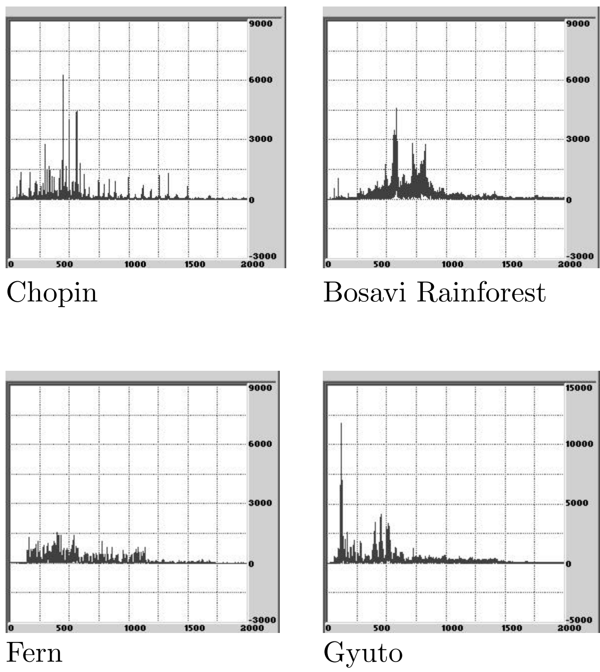
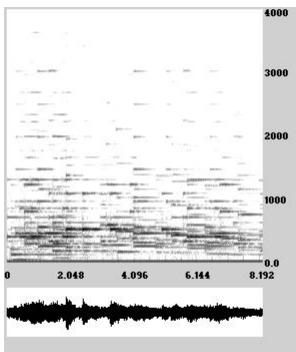


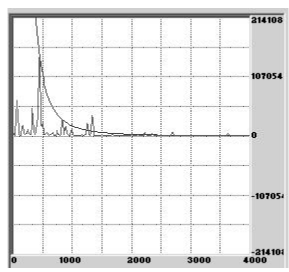
Figure 3. Global transform amplitudes.

Table 1. Estimated Exponents γ for $|\widehat{Y}_j| \propto 1/f^{\gamma}$.

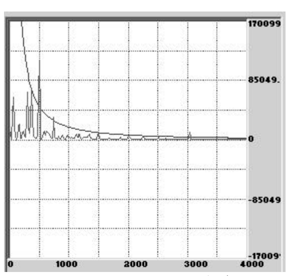
Time (sec)	Chopin	Fern	Bosavi	Gyuto
0.5	1.8	0.0	2.8	1.6
1.0	2.6	1.8	3.5	0.5
1.5	1.4	2.1	0.4	0.6
2.0	3.6	0.2	13.9	2.6
2.5	2.4	0.5	3.4	0.6



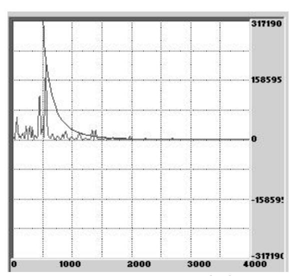
Dynamic Spectra



$t \approx 1.0$ sec, $1/f^{2.6}$



$t \approx 1.5$ sec, $1/f^{1.4}$



$t \approx 2.0$ sec, $1/f^{3.6}$

Figure 4. Analysis of Chopin passage.

2. Synthesized sounds

In the previous section we attempted to calculate exponents for a power law $1/f^{\gamma}$ and found that values for γ are widely dispersed for four test signals. In this section we take a different tack. Suppose that a commonly recognized musical passage, such as the Chopin piano passage considered above, is subjected to a transformation which surely distorts it away from a $1/f$ power law (assuming it has one to begin with). Given that the spectrogram has the form $\{|\widehat{Y}_j(\lambda_m)|\mathrm{e}^{i\theta(\lambda_m)}\}$ we subject it to two transformations:

$$\left\{|\widehat{Y}_j(\lambda_m)|^2\mathrm{e}^{i\theta(\lambda_m)}\right\} \tag{10}$$

and

$$\left\{\sqrt{|\widehat{Y}_j(\lambda_m)|}\mathrm{e}^{i\theta(\lambda_m)}\right\} \tag{11}$$

We then apply the spectrogram inversion procedure to each of the data in (10) and (11). This produces two sampled signals, denoted $\{Y_2(t_k)\}$ and $\{Y_{\sqrt{2}}(t_k)\}$, respectively. The signal $\{Y_2(t_k)\}$ was saved as `PowerTwo.wav`, and the signal $\{Y_{\sqrt{2}}(t_k)\}$ was saved as `Sqrt.wav`, at the website listed in (4). The reader is invited to listen to these two `.wav` files. If they sound like music (maybe not great music, but still music), then this casts further doubt on the $1/f$ hypothesis: Assuming that the Chopin passage `Chopin.wav` is $1/f$, then neither `PowerTwo.wav` nor `Sqrt.wav` are $1/f$ signals.

Conclusion

We have reported some preliminary results on determining exponents γ, for a spectral power law $1/f^{\gamma}$, for various musical passages. Further study is in order, such as increasing the number of peak values used in the spectrogram analysis, and obtaining further data on subjective impressions of synthesized sounds. Preliminary results cast some doubt on the $1/f$ power law proposed by other workers.

References

[1] Brillinger, D.R., and R.A. Irizarry. (1998). "An investigation of the higher-order spectra of music." *Signal Processing,* **65,** 161–79.

[2] Gröchenig, K. (2001). *Foundations of Time-Frequency Analysis.* Boston, MA: Birkhäuser.

[3] Lowen, S.B., and M.C. Teich. (1993). "Fractal renewal processes generate 1/f noise." *Physical Rev.,* **47,** 992-1001.

[4] Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. (1992). *Numerical Recipes,* 2nd Edition. Cambridge, England: Cambridge University Press.

[5] Voss, R.F., and J. Clarke. (1975). "'1/f Noise' in music and speech." *Nature,* **258,** 317–18.

[6] Voss, R.F., and J. Clarke. (1978). "'1/f Noise' in music: music from 1/f noise." *J. Acous. Soc. Am.,* **63,** 258–63.

[7] Walker, J.S. (1996). *Fast Fourier Transforms,* 2nd Edition. Boca Raton, FL: CRC Press.

[8] Walker, J.S., and J.F. Alm. (2002). "Time-Frequency Analysis of Musical Instruments." *SIAM Review*, **44**, 457–76.