

Why Filtering Out Higher Harmonics Makes It Easier to Carry a Tune

Griselda Acosta¹, Eric Freudenthal²,
Eric Smith³, and Vladik Kreinovich²

¹Department of Electrical and Computer Engineering

²Department of Computer Science

²Department of Industrial, Manufacturing, and
Systems Engineering

University of Texas at El Paso

500 W. University

El Paso, TX 79968, USA

gvacosta@miners.utep.edu, efreudenthal@utep.edu,
esmith2@utep.edu, vladik@utep.edu

Abstract

A recent patent shows that filtering out higher harmonics helps people sing in-tune. In this paper, we use the general signal processing ideas to explain this empirical phenomenon. We also show that filtering out higher harmonics is the optimal way of increasing the signal-to-noise ratio – and thus, of making it easier for people to recognize when they are signing out of tune.

1 Formulation of the Problem

A helpful invention. According to the patent description, the patent [1] “greatly improves the singing abilities of both novice and experienced singers by amplifying the fundamental frequency of one’s voice to correct tone deafness.”

It works but why? The device has been successfully tested, it clearly works, but why? Why does amplifying the fundamental frequency – or, equivalent;y, filtering out the higher harmonics – helps a person carry a tune?

This is a question to which we plan to find an answer in this paper. This answer will be based on the general engineering signal processing ideas; see, e.g., [2, 3]. Moreover, not only we explain why this works, we also show that such filtering is the optimal way to make it easier for a person to carry a tune.

2 Our Explanation

Higher harmonics: a brief reminder. Each note corresponds to a certain fundamental frequency f_0 . The resulting signal is periodic with the same frequency. Thus, if we perform the Fourier transform – i.e., if we represent the signal as a linear combination of sinusoids of different frequencies, then we only get components corresponding to multiples of the fundamental frequency f_0 . Components corresponding to frequencies $2f_0$, $3f_0$, etc., are known as *higher harmonics*.

The fact that the frequency f_0 is fundamental means that the component corresponding to this frequency has the largest energy $S(f_0)$; the energies of the higher harmonics are smaller:

$$S(2f_0) < S(f_0); \quad S(3f_0) < S(f_0), \dots \quad (1)$$

Why is it not always easy to carry a tune: signal-processing analysis.

In general, in signal processing, the quality of signal detection depends on the signal-to-noise ratio. Thus, if a singing person does not understand that he/she is singing out of tune, this means that for the sound produced by this singing person, the signal-to-noise ratio is too low to detect this.

The overall energy S of the signal can be computed by adding the energies corresponding to the fundamental frequency and to the higher harmonics:

$$S = S(f_0) + S(2f_0) + S(3f_0) + \dots \quad (2)$$

Similarly, the overall energy N of the noise can be computed by adding the energies of the noise on all these frequencies, i.e.:

- the energy $N(f_0)$ of the noise on the fundamental frequency f_0 ,
- the energy $N(2f_0)$ of the noise on the double frequency $2f_0$, etc.:

$$N = N(f_0) + N(2f_0) + N(3f_0) + \dots \quad (3)$$

In contrast to signal whose energy changes drastically from one frequency to another, the energy of the noise is usually changing very little from one frequency to another. In the first approximation, we can therefore simply assume that this energy is the same for all the involved frequencies:

$$N(f_0) = N(2f_0) = N(3f_0) = \dots \quad (4)$$

Thus, the formula (3) has the form

$$N = k \cdot N(f_0),$$

where k is the overall number of harmonics.

The corresponding signal-to-noise ratio of the original signing signal is thus equal to

$$\frac{S}{N} = \frac{S(f_0) + S(2f_0) + S(3f_0) + \dots + S(k \cdot f_0)}{k \cdot N(f_0)}. \quad (5)$$

The fact that a person has difficulty correctly carrying a tune means that this signal-to-noise ratio is too small. We need to increase it.

Let us apply filtering. In signal processing, a usual way to increase the signal-to-noise ratio is to perform some filtering. Filtering means that we either amplify or decrease certain frequencies. This amplification or damping is applied to the combination of signal and noise, so it equally affects both. In both cases of amplification or damping, the energy of both the signal component and of the noise component is multiplied by the same coefficient $c(f) \geq 0$ depending on the frequency f :

- for the fundamental frequency f_0 , the energy of the signal changes from $S(f_0)$ to $c(f_0) \cdot S(f_0)$ and the energy of the noise changes from $N(f_0)$ to $c(f_0) \cdot N(f_0)$;
- for the frequency $2f_0$, the energy of the signal changes from $S(2f_0)$ to $c(2f_0) \cdot S(2f_0)$ and the energy of the noise changes from $N(2f_0) = N(f_0)$ to $c(2f_0) \cdot N(f_0)$;
- for the frequency $3f_0$, the energy of the signal changes from $S(3f_0)$ to $c(3f_0) \cdot S(2f_0)$ and the energy of the noise changes from $N(3f_0) = N(f_0)$ to $c(3f_0) \cdot N(f_0)$; etc.

After the filtering, the overall energy of the signal is equal to

$$S' = c(f_0) \cdot S(f_0) + c(2f_0) \cdot S(2f_0) + \dots + c(k \cdot f_0) \cdot S(k \cdot f_0), \quad (6)$$

the overall energy of the noise is equal to

$$\begin{aligned} N' &= c(f_0) \cdot N(f_0) + c(2f_0) \cdot B(f_0) + \dots + c(k \cdot f_0) \cdot N(f_0) = \\ &= (c(f_0) + c(2f_0) + \dots + c(k \cdot f_0)) \cdot N(f_0), \end{aligned} \quad (7)$$

and thus, the new signal-to-noise ratio is equal to

$$\frac{S'}{N'} = \frac{c(f_0) \cdot S(f_0) + c(2f_0) \cdot S(2f_0) + \dots + c(k \cdot f_0) \cdot S(k \cdot f_0)}{(c(f_0) + c(2f_0) + \dots + c(k \cdot f_0)) \cdot N(f_0)}. \quad (8)$$

Which filter is optimal: formulation of the problem. We want to find the coefficients $c(f_0)$, $c(2f_0)$, \dots , $c(k \cdot f_0)$ for which the signal-to-noise ratio (8) attains the largest possible value – this will lead to the best possible chance of a person recognizing inaccuracies in his/her own signing.

The optimal filter is exactly filtering out higher harmonics. Let us prove that the optimal filter is exactly the filter used in [1] – the filter that filters out all higher harmonics, i.e., the filter for which

$$c(f_0) > 0 \text{ and } c(2f_0) = \dots = c(k \cdot f_0) = 0. \quad (9)$$

Indeed, for this filter, the signal-to-noise ratio is equal to

$$\frac{S'}{N'} = \frac{c(f_0) \cdot S(f_0)}{c(f_0) \cdot N(f_0)} = \frac{S(f_0)}{N(f_0)}. \quad (10)$$

What happens if at least one of the higher harmonics is not completely filtered out, i.e., we have $c(i \cdot f_0) > 0$ for some i ? In this case, for all such harmonics i , we have, due to the inequalities (1), $S(i \cdot f_0) < S(f_0)$ hence

$$c(i \cdot f_0) \cdot S(i \cdot f_0) < c(i \cdot f_0) \cdot S(f_0).$$

By adding up these inequalities, we conclude that

$$\begin{aligned} S' &= c(f_0) \cdot S(f_0) + c(2f_0) \cdot S(2f_0) + \dots + c(k \cdot f_0) \cdot S(k \cdot f_0) < \\ &c(f_0) \cdot S(f_0) + c(2f_0) \cdot S(f_0) + \dots + c(k \cdot f_0) \cdot S(f_0) = \\ &(c(f_0) + c(2f_0) + \dots + c(k \cdot f_0)) \cdot S(f_0). \end{aligned} \quad (11)$$

Dividing both sides of this inequality by the expression (7) for the noise N' , we conclude that

$$\frac{S'}{N'} < \frac{(c(f_0) + c(2f_0) + \dots + c(k \cdot f_0)) \cdot S(f_0)}{(c(f_0) + c(2f_0) + \dots + c(k \cdot f_0)) \cdot N(f_0)}. \quad (12)$$

Dividing both numerator and denominator of the right-hand side by the same sum $c(f_0) + c(2f_0) + \dots + c(k \cdot f_0)$, we thus conclude that

$$\frac{S'}{N'} < \frac{S(f_0)}{N(f_0)}. \quad (13)$$

Thus, if at least one of the higher harmonics is not fully filtered out, the resulting signal-to-noise ratio is smaller than the value (10) that we have when all these harmonics are filtered out.

In particular, by taking the values

$$c(f_0) = c(2f_0) = \dots = c(k \cdot f_0) = 1$$

corresponding to taking the original signal as is, we conclude that the signal-to-noise ration of the optimally filtered signal is indeed larger than the signal-to-noise ratio of the original signal.

Conclusion. Thus, we have shown the following:

- We showed that filtering out higher harmonics increases signal-to-noise ratio. Thus, we explain why after this filtering, it is easier for a person to detect when he or she is signing out of tune.

- We also showed that filtering out higher harmonics is indeed the optimal approach – in the sense that it leads to the largest possible increase in the signal-to-noise ratio (and thus, to the best chance of detecting out-of-tune deviations).

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

References

- [1] E. Freudenthal, B. Usevitch, and E. Hanson, *Apparatus and Method to Facilitate Singing Intended Notes*, US Patent No. 9773426, 2017.
- [2] S. R. Laxpati and V. Goncharoff, *Practical Signal Processing and Its Applications*, World Scientific, Singapore, 2018.
- [3] R. G. Lyons, *Understanding Digital Signal Processing*, Prentice Hall, Upper Saddle River, New Jersey, 2011.
- [4] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing*, Pearson, New York, 2006.