

CREATING SPATIAL DEPTH USING DISTORTION PRODUCT OTOACOUSTIC EMISSIONS IN MUSIC COMPOSITION

Alex Chechile

Stanford University,
Center for Computer Research in Music and Acoustics (CCRMA),
660 Lomita Court, Stanford, CA, USA
chechile@ccrma.stanford.edu

ABSTRACT

Distortion product otoacoustic emissions (also known as combination tones) are sounds generated within the listener's ears upon physical and physiological interactions between spectral components in a given auditory input. The relationship between sounds generated by loudspeakers and sounds generated in the listener's ears offers fertile ground for the exploration of spatial depth in music. The author's *On the Sensations of Tone* is a series of compositions that provoke ear tones to create an additional spatial dimension in each piece. Following a brief overview of the history and mechanisms behind otoacoustic emissions, this paper discusses techniques for composing with the phenomenon as used in *On the Sensations of Tone VII* (2015).

1. INTRODUCTION

Alex Chechile's *On the Sensations of Tone VII* (2015) is one piece in a series of compositions that explore the physicality of sound and spatial depth through the application of research in psychoacoustics and the biomechanics of hearing. In addition to loudspeaker-based sound diffusion, a second layer of spatial depth is created by provoking the ears to emit frequencies of their own. Spatial depth is experienced through multilayered listening of microscopic and macroscopic levels of sonic material.

Composer and artist Maryanne Amacher (1938-2009) conducted sonic and perceptual experiments with electronic instruments to provoke ear tone responses in her music [1]. Amacher describes the interplay between the tones intentionally generated by physiological mechanisms and the tones sounded in the room as a "perceptual geography" [2]. She writes, "the object of perceptual geography is to learn to compose spatial dimensions in music – the kind of aural dimension we experience in life, but usually not in music" [2]. While Amacher's definition of perceptual geography extends to other forms of spatialization beyond the implementation of ear tones, she continues to describe the effect as, "tactile in presence, [sounds] appear both larger than life and small enough to touch, heard as though miles away, or felt inside the listener" [2]. The tactile quality of sound and the auditory depth of field in Amacher's work inform the methods used for composing with otoacoustic emissions in *On the Sensations of Tone*.

2. CONDITIONS FOR PROVOKING DPOAE

Distortion product otoacoustic emissions (DPOAEs) are sounds that emerge from the ear due to an active process in the inner ear. Upon the simultaneous presentation of two pure tone frequencies f_1 and f_2 ($f_2 > f_1$) called primaries, non-linear distortion products are produced in the cochlea [3]. The most detectable distortion products are at the frequencies of $f_2 - f_1$, which is known as the quadratic difference tone (QDT) and $2f_1 - f_2$, which is known as the cubic difference tone (CDT) [4]. A number of psychoacoustic studies indicate the most salient perception of distortion products are achieved when the two frequencies are within a half-octave from each other at a f_2 / f_1 ratio of 1.22 [3]. The intensity level of f_1 (which is known as L1) is generally greater than f_2 (known as L2) by 10-15 dB SPL [3]. Clinical levels for invoking DPOAE are in the range of 50 - 70 dB SPL [3], however when presented through free-field loudspeakers in music, the levels are typically louder.

3. HISTORY

Combination tones were independently discovered in the mid-eighteenth century by the German organist Georg Andreas Sorge (1703-1778), French dilettante and scientist Jean-Baptiste Romieu (1723-1766), and the Italian violinist and composer Giuseppe Tartini (1692-1770) (whose name is often attributed to the phenomenon) [5]. In the mid-nineteenth century German physicist Hermann von Helmholtz (1821-1894) studied and wrote about combination tones, which he believed to be the result of an overloading process in the middle ear [6]. While the aforementioned musicians and scientists were unknowingly discussing OAE, our understanding of the mechanisms behind the phenomenon was predicted by the British auditory scientist Thomas Gold (1920-2004) in 1948 [7]. Contrary to Georg von Békésy's (1899-1972) widely accepted theory that the cochlea acted passively, Gold believed an active electro-mechanical process was needed in order to counteract the damping that occurs from the viscosity of the liquid in the inner ear [8]. His ideas were not embraced at the time, however in the mid-1970's the British physicist David T. Kemp (b. 1945) also believed that a mechanically active process was taking place in the ear. Kemp discovered otoacoustic emissions after he placed a microphone in the ear and was able to record spontaneous otoacoustic emissions (SOAEs) and DPOAEs [9], [10].



This work is licensed under Creative Commons Attribution – Non Commercial 4.0 International License. The full terms of the License are available at <http://creativecommons.org/licenses/by-nc/4.0/>

4. AUDITORY MECHANISMS BEHIND OAE

In 1983, physiologist Hallowell Davis (1896-1992) described a “cochlear amplifier” that involved an active process with the outer hair cells (OHCs) [11]. First observed by biophysicist William E. Brownell, the OHCs on the basilar membrane of the cochlea are capable of electromotility, which is the most significant mechanism for the generation of otoacoustic emissions [12]. The active motility of the OHCs increases stimulus-specific vibrations of the basilar membrane [3]. Hearing sensitivity and frequency selectivity is heightened when the vibrations are transmitted to the inner hair cells (IHCs), which are sensory receptors [3]. Hall concisely states, “stimulus-induced macromechanical vibrations of basilar membrane and organ of Corti produce micromechanical vibrations of hair cells. These can in turn exert a ‘feedback’ influence on the macromechanics [3].” The frequency response of the electromotility is nonlinear and extends to an audible range [12].

How exactly do the outer hair cells oscillate? The stimulus-induced movement of the OHCs is the result of a voltage change across the plasma membrane of the hair cell, which in turn leads to a change in length of the stereocilia [3]. Hyperpolarization causes the OHCs to shorten and become fatter, and depolarization causes the OHCs to elongate and become thinner [3]. Due to stiff structures above and below the OHCs, the resultant oscillations lead to increased movement of the basilar membrane, which aids the loss of energy from damping and friction effects of the fluids in the cochlea [3].

Kemp explains that otoacoustic emissions are the result of a leakage of energy from the cochlear amplifier [13]. He states, “in reality, as outer hair cell gain is increased, there comes a point where any small irregularities in outer hair cell arrangement activity become magnified and significant stimulus frequency energy travels backward, to cause OAEs” [10]. Since the OHC movement is non-linear, the basilar membrane distorts. Intermodulation components will appear when a mixture such as the pure tone stimulus f_1 and f_2 are presented, producing tones at f_2-f_1 , $2f_1-f_2$, $2f_2-f_1$, etc. The tones created through this intermodulation distortion of the basilar membrane are called distortion products. Apart from their immense contribution to auditory health diagnostics, DPOAE provide fertile ground for new forms of spatial and textural music composition.

5. OAE AS SPATIAL DEPTH IN MUSIC

Since acoustic primary tones in the room produce distortion products in the ear, spatial depth between sound sources is inherent to the process of working with DPOAE. Provoking distortion products creates an active listening experience for the audience. By listening, the audience members become active participants in the production of the composition since their ears generate a layer of the musical material. Spatial focus between the layers of acoustic primary tones, the ear-tones, and the macroscopic combination provides the listener with an additional mode of participating with the composition.

A multichannel speaker array allows further spatialization, prevents speakers from overloading, and can be used to create a spectrum of distortion products. While a single pair of speakers can produce multiple primaries, if the speakers overload the distortion products will be produced in the speaker cone rather than in the listener’s ears. Using multiple speakers allows for an additional layer of spatial



Figure 1: Section two of *On the Sensations of Tone VII* arranges sustained chords and short bursts of combination tones. The image depicts the composition as recorded in Ableton Live.

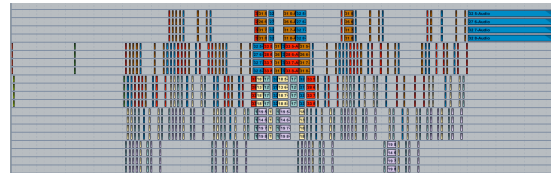


Figure 2: Detail for the second half of section two of *On the Sensations of Tone VII*.

depth by sound source location, and also allows for more primary tones to be used. The increased number of primary tones facilitates a denser spectrum of distortion products.

Physical space is also explored as slight head movement greatly effects how the work is sounded. By changing one’s position in the sonic space, ear tones will appear, disappear, and change timbre. Cupping hands around one’s ears will amplify the intensity of the distortion products.

6. ON THE SENSATIONS OF TONE VII (2015)

While the general discussion in the previous section applies to *On the Sensations of Tone VII*, the following section discusses more detailed techniques used in the construction of the piece. The composition is built from a custom non-repeating scale system derived from particularly resonant ear tones selected by the author from sweeping the frequency spectrum with a constant CDT value. Each note of the scale produces two acoustic frequencies (f_1 and f_2) and the resulting quadratic and cubic difference tones. The balance between the dominant difference tone varies across the scale, however the QDT is often the most detectable. Based on the scale, polyphonic melody and harmony is crafted to generate a complex spectrum of distortion products. Table 1 shows the combination tone scale system used in the composition.

The sections of *On the Sensations of Tone VII* alternate between non-ear tone material generated with a modular synthesizer and ear tone sections generated by custom software built in Max/MSP and the Chuck programming language. The ear tone material is categorized into two groups: sustained chords, and short bursts. Each of these two categories overlap at times, resulting in additional primary tone combinations that produce a denser ear tone spectrum.

Distortion products first appear in the second section of the piece, which begins with sustained ear tones layered as distortion product chords (see figure 1). Each chord consists of eight acoustic tones (four notes from the scale) playing through eight loudspeakers, which produces a dense distortion product spectrum consisting of each combination’s quadratic difference tone, cubic difference tone, and the

note	f1	f2	QDT	CDT
1	1289	1378	89	1,200
2	1301	1402	101	1,200
3	1339	1478	139	1,200
4	1349	1498	149	1,200
5	1376	1552	176	1,200
6	1446	1692	246	1,200
7	1464	1728	264	1,200
8	1514	1828	314	1,200
9	1560	1920	360	1,200
10	1600	2000	400	1,200
11	1650	2100	450	1,200
12	1709	2218	509	1,200
13	1756	2312	556	1,200
14	1828	2456	628	1,200
15	1850	2500	650	1,200
16	1890	2580	690	1,200
17	1923	2646	723	1,200
18	1960	2720	760	1,200
19	2029	2858	829	1,200
20	2081	2962	881	1,200
21	2110	3020	910	1,200
22	2142	3084	942	1,200
23	2304	3408	1,104	1,200
24	2375	3550	1,175	1,200
25	2480	3760	1,280	1,200
26	2518	3836	1,318	1,200
27	2634	4068	1,434	1,200
28	2689	4178	1,489	1,200
29	2728	4256	1,528	1,200
30	2814	4428	1,614	1,200
31	2870	4540	1,670	1,200
32	2947	4694	1,747	1,200
33	3017	4834	1,817	1,200

Table 1: Scale system of acoustic stimulus pairs and associated quadratic and cubic difference tones as measured in Hz. The cubic difference tone was held at a constant frequency to limit the combined spectrum, find and emphasize the quadratic difference tone, and create a common pitch linking the notes of the scale.

resulting interactions between each note. The distortion product spectrum increases in density when polychords are introduced (such as the blue/green pairing in figure 1). The orange chord in figure 1 is designed to resemble a tonic chord and is frequently repeated. While the sustained material is sounded, the listener is encouraged (through either program notes or a pre-concert talk) to move their head or change their position in the room as the ear tone layer changes. This interactive component gives the listener further control of the sounds that their ears are generating.

The second half of section two (see figure 2) combines chords from the sustained section with new chords, and introduces the sustained tones as truncated short bursts. The bursts are organized into two repeating melodies that evolve through a rotating order of each burst. The two melodies play simultaneously, producing short bursts of distortion products. Duration is explored as the bursts elongate and overlap, producing different combinations of ear tone material.

Ear Tone Chord 1

note	f1	f2	QDT	CDT
8	1514	1828	314	1,200
9	1560	1920	360	1,200
10	1600	2000	400	1,200
11	1650	2100	450	1,200

Ear Tone Chord 2

note	f1	f2	QDT	CDT
12	1709	2218	509	1,200
13	1756	2312	556	1,200
14	1828	2456	628	1,200
15	1850	2500	650	1,200

Table 2: Deconstruction of each of the two chords for section four of *On the Sensations of Tone VII*. The note numbering is consistent with the scale system outlined in Table 1, and the frequencies are displayed in Hz. The QDT and CDT are provided for each note, however the internal combinations between each note in the chord are not shown.



Figure 3: Section four of *On the Sensations of Tone VII* showing the composite recording of the two chord progression. The red lines represent amplitude of the final mix.

Section two was composed in Max/MSP by developing a keyboard that generated the ear tone scale, and it was recorded and arranged in Ableton Live. The distortion product chords were built through exploring ear tone harmony in an octophonic sound system. Once the chords were selected, their duration was modified into sustained chords and short burst melodies.

The fourth section of *On the Sensations of Tone VII* starts with the tonic chord from section two but soon transitions to new material. The section was composed in the ChucK programming language with a script that generated a repeating two-chord progression of primary tones in an eight-channel sound system (see Table 2). The script was added multiple times in asynchrony to generate rhythm that phases upon each repetition. The overlapping chords once again produce a complex distortion product spectrum with each of the four pairs of primaries creating their respective quadratic difference tones, cubic difference tones, and internal interactions. The chords change when the amplitude of each channel of audio is adjusted to create a shifting dynamic

between primary tones (see figure 3).

7. CONCLUSION

A consistent compositional technique held throughout *On the Sensations of Tone VII* is the generation of a dense distortion product spectrum created from layering multiple primary tones in a multichannel sound system. The dynamic between the primary acoustic frequencies and the resulting ear tones creates a sonic environment where listeners can shift their attention between a microscopic layer (ear tones, or primaries) and the macroscopic layer (the composite sonic environment). In addition to producing sounds by simply listening, participants are asked to further engage with the sound space by exploring their physical space to change their ear tones. The use of multiple speakers allows for a denser spectrum of sounds, and increases spatial depth through loudspeaker localization.

8. ACKNOWLEDGMENT

On the Sensations of Tone VII was composed under the mentorship of Brian Ferneyhough, and the underlying scientific research was supported by Takako Fujioka. In addition to the aforementioned individuals, the author would like to thank Chris Chafe, John Chowning, Cathleen Grado, Jaroslaw Kapuscinski, Sasha Leitman, Fernando Lopez-Lezcano, Romain Michon, Sunil Puria, Malcolm Slaney, Ge Wang, Carr Wilkerson, and his colleagues at the Stanford Center for Computer Research in Music and Acoustics (CCRMA) for their invaluable support. Additionally, the author is indebted to Maryanne Amacher for her friendship, guidance, and perpetual inspiration.

9. REFERENCES

- [1] M. Amacher, "Psychoacoustic Phenomena in Musical Composition: Some Features of a 'Perceptual Geography'," in *Arcana III*, New York, NY: Hips Road, 2008.
- [2] M. Amacher, "Synaptic Island: A Psybertonal Topology," in *Architecture as a translation of music*, E. Martin, Ed. New York: Princeton Architectural Press, 1994.
- [3] J. W. Hall III, *Handbook of Otoacoustic Emissions*. San Diego, California: Singular Publishing Group, 2000.
- [4] D. W. Purcell, B. Ross, T. W. Picton, and C. Pantev, "Cortical responses to the 2f1-f2 combination tone measured indirectly using magnetoencephalography," *J. Acoust. Soc. Am.*, vol. 122, no. 2, pp. 992–1003, Aug. 2007.
- [5] R. Plomp, *Experiments on tone perception*. Soesterberg: National Defense Research Organization TNO. Institute for Perception RVO-TNO, 1966.
- [6] H. Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, 2nd English ed. New York: Dover Publications, 1954.
- [7] S. Dhar and J. W. Hall III, *Otoacoustic Emissions: Principles, Procedures, and Protocols*. San Diego, California: Plural Publishing Inc, 2012.
- [8] T. Gold, "Hearing. II. The Physical Basis of the Action of the Cochlea," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 135, no. 881, pp. 492–498, Dec. 1948.
- [9] D. T. Kemp, "Stimulated acoustic emissions from within the human auditory system," *J. Acoust. Soc. Am.*, vol. 64, no. 5, pp. 1386–1391, Nov. 1978.
- [10] D. T. Kemp, *The OAE Story*. Hatfield, UK: Otodynamics, 2003.
- [11] H. Davis, "An active process in cochlear mechanics," *Hear. Res.*, vol. 9, no. 1, pp. 79–90, Jan. 1983.
- [12] W. E. Brownell, "Outer hair cell electromotility and otoacoustic emissions," *Ear Hear.*, vol. 11, no. 2, pp. 82–92, Apr. 1990.
- [13] D. T. Kemp, "Otoacoustic Emissions: Concepts and Origins," in *Active Processes and Otoacoustic Emissions in Hearing*, G. A. Manley, R. R. Fay, and A. N. Popper, Eds. New York: Springer, 2008.