

Revitalizing Classic Illusions: Shepard-Tone Sequences and Shepard–Risset Glissandi, With Various Modifications

ANDREW V. FRANE,* *AES Associate Member*
(avfrane@ucla.edu)

Occidental College, Los Angeles, CA

The *Shepard-tone sequence* and *Shepard–Risset glissando* are classic auditory illusions in which pitch seems to inexhaustibly ascend or descend. Such stimuli have been used in scientific research, as well as for artistic purposes. This paper demonstrates several variations of those illusions, some of which do not appear to have been previously discussed in the literature. Most notably, hybrids of the two illusions are demonstrated, in which discrete Shepard-tone steps are connected by continuous glissandi. It is shown, using a sample of 91 listeners, that such hybrids can disambiguate the perceived direction of motion between two Shepard tones that are a tritone apart, thus overriding what has been called the *tritone paradox*. In other demonstrations, multiple layers of monaural and binaural beats are embedded into a Shepard–Risset glissando to produce *Risset rhythms*. Audio files for these and other examples are provided and discussed. Two original MATLAB functions (and equivalent functions in R) are also provided, which can be used to replicate the examples and explore additional variations.

0 INTRODUCTION

The term *illusion* is often used to refer to a distorted perception of a misleading visual stimulus. However, illusions have also been identified in other sensory modalities, including audition [1]. Like visual illusions, auditory illusions generally work by exploiting vulnerabilities of people's perceptual strategies—strategies that under more typical circumstances would, at least heuristically, provide accurate interpretations.

For instance, the human auditory system relies largely on inter-aural differences of intensity to judge the horizontal position of a sound source. Stereophonic imaging exploits that to create the illusion that a given sound source is located at a single intermediate position between the two speakers. Another commonly encountered auditory illusion that involves distorted localization of a sound source is the *ventriloquist effect*, whereby a sound source is mistakenly perceived as being located at the same position as a corresponding visual stimulus. That occurs when, for example, the dialogue in a movie seems to emanate from the face of the actor whose lips are moving, even though the viewers are well aware that the sound is actually coming from

peripheral loudspeakers or from earphones that they are wearing.

Some auditory illusions that have been defined involve pitch perception. One could argue that pitch cannot be considered as illusory per se, because—unlike the physical location of a sound source—pitch is a subjective experience, so there is no corresponding “true” value that can be compared to a given perceived pitch to determine whether the perception is real or illusory. In that regard, the pitch of a sound is simply whatever it is perceived to be. Nonetheless, because pitch may be considered as, loosely speaking, the mind's interpretation of frequency, it is common to refer to certain pitch-perception anomalies as “illusions” if the typical correspondence between pitch and frequency is disrupted in some way. For instance, the *missing fundamental effect*, whereby the pitch of a complex tone is perceived as unchanged when the fundamental frequency is removed, is sometimes called an illusion [1]. The present paper focuses on two other types of stimuli that are commonly regarded as pitch illusions: the *Shepard-tone sequence* [2] and *Shepard–Risset glissando* [3].

0.1 Shepard-Tone Sequences

While working at Bell Labs, the psychologist Roger Shepard used sound synthesis software (developed by his colleague Max Mathews) to demonstrate a fascinating pitch

*To whom correspondence should be addressed, email: avfrane@gmail.com, Last updated: March 12, 2024

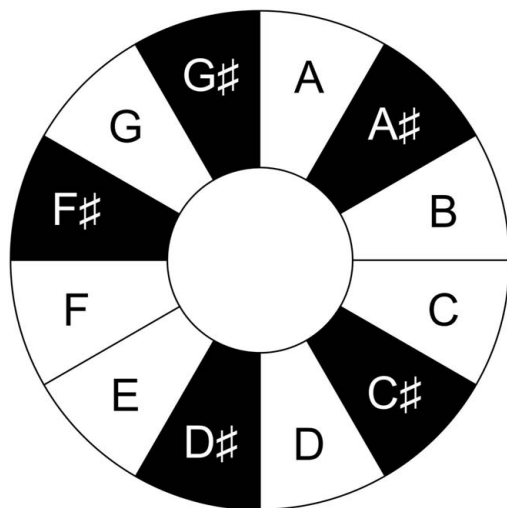


Fig. 1. Circular octave.

illusion that was subsequently reported in a now-classic 1964 paper [2]. In this illusion, the pitches of a series of tones are perceived as inexhaustibly ascending or inexhaustibly descending, whereas in reality the tones are merely cycling over and over, containing exactly the same frequency components each time. Shepard analogized this phenomenon to the impossible Penrose staircase, which appears to perpetually ascend or descend despite being a closed loop (a visual illusion utilized in M. C. Escher's famous *Ascending and Descending* lithograph [4]). Shepard's illusion has also been called an "auditory barber-pole," which is an analogy to the classic barber-pole with helical stripes that appear to perpetually ascend or descend as the pole rotates horizontally [5].

Shepard's illusion works because each tone in the sequence is a *Shepard tone*, i.e., an ensemble of superimposed sine waves whose frequencies are spaced in octave increments and collectively span a broad range. Thus, for example, a Shepard-tone $G\sharp$ is not a $G\sharp_3$, or a $G\sharp_4$, or a $G\sharp$ localized to any other particular octave-register. Rather, it is an octave-ambiguous, "universal" $G\sharp$ that is sounded in all octaves of the spectrum simultaneously. In other words, it has an unambiguous *pitch class* (namely, $G\sharp$) but no inherent *pitch height* (i.e., it is not localized to a particular register of the frequency spectrum). And although a Shepard-tone $G\sharp$ is positioned between G and A as one would expect, that positioning is in a *circular octave*, as shown in Fig. 1. Consequently, the direction of motion from a Shepard-tone $G\sharp$ to, for example, a Shepard-tone A is ambiguous. In the circular octave, that motion could be considered as moving one semitone in the "ascending" (clockwise) direction, but it could also be considered as moving 11 semitones in the "descending" (counterclockwise) direction.

Yet the motion from a Shepard-tone $G\sharp$ to a Shepard-tone A is in fact typically perceived as ascending one semitone, rather than as descending 11 semitones. That appears to be mainly due to a perceptual bias that favors smaller pitch intervals in the circular octave [2]. Exploiting that bias, a Shepard-tone sequence that repeatedly loops clockwise around the circle, one semitone at a time, can seem to per-

petually ascend—even though the same series of 12 Shepard tones is actually repeating over and over. Conversely, a Shepard-tone sequence that repeatedly loops counterclockwise around the circle can seem to perpetually descend.

In each Shepard tone, the high and low frequencies are gradually rolled off, using an amplitude envelope that is mapped to the log-transformed frequency spectrum. That filtering makes the cycling less conspicuous, i.e., it prevents harmonics from seeming to suddenly appear or disappear at the edges of the spectrum. The filtering also makes the tones considerably more pleasant by attenuating the "screechier" high frequencies. Typically, a bell-shaped envelope is used [2, 6], but other unimodal curves have also been used [7].

Shepard's illusion works with other small increments besides semitones. In fact, when presented with any two Shepard tones, listeners tend to perceive the direction of motion as whichever direction implies a shorter distance around the circle [2]. For example, moving from C to E tends to be perceived as ascending 4 semitones (rather than as descending 8 semitones), whereas moving from C to A tends to be perceived as descending 3 semitones (rather than as ascending 9 semitones).

However, that perceptual bias is less consistent when the difference between the ascending and descending distances is small, e.g., moving 5 semitones in one direction versus seven in the other [2, 8, 9]. And when the interval is a tritone (6 semitones), such as when moving from C to $F\sharp$, the distance is the same in either direction, in which case judgment of direction cannot be based on circular distances alone. Indeed, the direction of a Shepard-tritone may be perceived as either ascending or descending, depending on the specific pitch classes and depending on the listener—a phenomenon called the *tritone paradox* [10]. Various spectral, temporal, contextual, and even cultural factors have been proposed to explain why a Shepard-tritone interval may be perceived as ascending or descending by a given listener in a given case [11–13].

0.2 Shepard–Risset Glissandi

Jean-Claude Risset [3] demonstrated a modified version of Shepard's illusion. In Risset's version, which is often called a *Shepard–Risset glissando*, a tone seems to perpetually ascend or descend as a continuous glide, rather than in the discrete stepped manner that was demonstrated by Shepard. But the same basic mechanisms of Shepard's illusion underly Risset's version. That is, the tone is comprised of layered sine waves that are octaves apart across the spectrum, and the high and low frequencies are rolled off to help mask the cycling. Thus, loosely speaking, a Shepard–Risset glissando is like a Shepard-tone sequence in which the step increments are arbitrarily small.

Shepard–Risset glissandi have been found to have some interesting effects beyond the pitch illusion itself. For example, in some listeners, Shepard–Risset glissandi have been reported to elicit emotions, disequilibrium, and/or *vection* (illusory self-motion) [6, 14, 15]. Capitalizing on those emotive and visceral effects, several films have used Shepard–Risset glissandi to represent—and perhaps even

induce—feelings of tension, disorientation, unreality, or unstoppable acceleration [16–18]. Shepard–Risset glissandi have also been used by composers of experimental music [16] and have shown potential in sonification processes (e.g., in auditory feedback for drivers of electric vehicles [19, 20]).

0.3 Risset Rhythms

The Shepard–Risset glissando can be converted from a pitch illusion to a tempo illusion by applying the same principle of circularity on a slower timescale, i.e., on the timescale of beats-per-minute in a rhythm, rather than on the timescale of cycles-per-second in a tone. For instance, instead of superimposing several sweeping sine-waves whose instantaneous frequencies are always powers of two (i.e., octaves) apart, one could superimpose several accelerating or decelerating pulses whose instantaneous rates are always powers of two apart. That is, one could superimpose a whole-note pulse, half-note pulse, quarter-note pulse, eighth-note pulse, etc., all synchronized at the same constantly accelerating or decelerating tempo. By increasingly attenuating the pulse streams toward the slow and fast ends of the “tempo spectrum,” the progression can then be presented cyclically to create the illusion of perpetual acceleration or perpetual deceleration. This illusion is often called a *Risset rhythm*, though Risset credited unpublished work by Kenneth Knowlton as the first demonstration of it [21].

0.4 Beat Frequencies

When the frequencies of two tones are sufficiently similar, they are not heard as distinct pitches. Instead, the oscillations between near-perfect reinforcement and near-perfect cancellation manifest as a single tone that “beats” (i.e., pulsates/throbs in loudness) at a rate called the *beat frequency* [22]. The beat frequency is equal to the absolute difference between the two tone frequencies. For example, a 100-Hz tone and 95-Hz tone combine to form a tone that pulsates at a beat frequency of $100 - 95 = 5$ beats per second.

That phenomenon is illustrated in Fig. 2, in which Fig. 2(a) shows a 100-Hz sine wave, Fig. 2(b) shows a 95-Hz sine wave, and Fig. 2(c) shows the sum of those two component sine waves. On the left side of the figure (at time 0), the two component sine waves are 180° out of phase, so they perfectly cancel each other, thus minimizing the summed amplitude envelope. Then the phases of the two component sine waves gradually become more aligned over time—causing the summed amplitude envelope to swell—until there is near-perfect reinforcement at approximately 0.1 s, at which point the summed amplitude is maximized. Then the two sine waves gradually slip back out of phase—shrinking the summed amplitude envelope—until the two sine waves are back to near-perfect cancellation at approximately 0.2 s, at which point the summed amplitude envelope is again minimized. Thus, in the summed amplitude envelope, one complete cycle of swelling and shrinking (i.e., one complete beat) lasts 0.2 s, confirming that the beat frequency is equal to $1 \text{ beat} / 0.2 \text{ s} = 5$ beats per second.

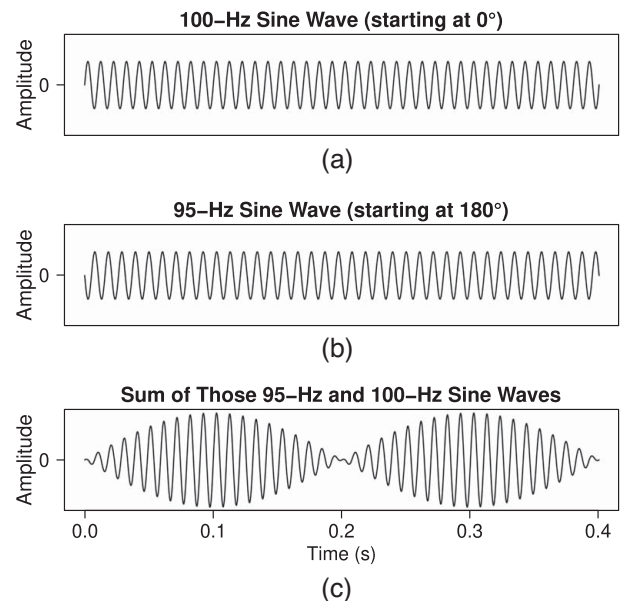


Fig. 2. Two sine waves summing to produce a beat frequency. (a) 100-Hz sine wave (starting at 0°), (b) 95-Hz sine wave (starting at 180°), and (c) sum of those 95-Hz and 100-Hz sine waves.

Beat frequencies are often heard when musicians are tuning their instruments. For example, when a guitarist is tuning a string by playing it in unison with a reference frequency (such as a harmonic on another string) and has not quite gotten the two frequencies to match, beats are produced. The beat frequency gets slower and slower as the string that is being tuned gets closer and closer to the reference frequency, until finally the beating essentially comes to a halt when the target is reached.

If several tones with very similar frequencies are played concurrently, then some beats will overlap and “smear” together into larger undulations. Fig. 3 shows three examples. Fig. 3(a) shows the sum of four sine waves with equally spaced frequencies, all initiated at a phase angle of 0° . Fig. 3(b) shows the sum of four sine waves with unequally spaced frequencies—spreading the smears throughout more of the wave—all initiated at 0° . Fig. 3(c) also shows the sum of four sine waves with unequally spaced frequencies, but in this case with different starting-phases, spreading the smears more asymmetrically throughout the wave.

If two tones with similar frequencies are presented dichotically (e.g., a 100-Hz tone in one ear and 95-Hz tone in the other), they can produce *binaural beats*, which not only pulsate in loudness but also seem to oscillate positionally in the stereo field, i.e., in the listener’s head [22]. Interesting psychotropic effects of binaural beats have been reported by some listeners, including alterations of mood and cognitive state, but individual experiences are highly varied [23]. A fascinating aspect of binaural beats is that, unlike monaural beats, they are not produced by actual interactions between tones in the signal path or the air, since the tones are presented directly to different ears through different channels. Rather, the beats are constructed by the listener’s brain.

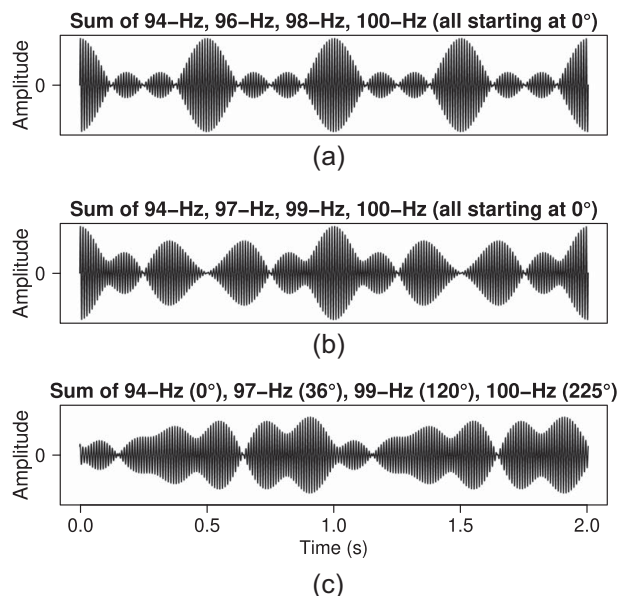


Fig. 3. “Smeared” beats produced by summing several sine waves of given frequencies and given starting-phases. (a) Sum of 94-Hz, 96-Hz, 98-Hz, 100-Hz (all starting at 0°); (b) sum of 94-Hz, 97-Hz, 99-Hz, 100-Hz (all starting at 0°); and (c) sum of 94-Hz (0°), 97-Hz (36°), 99-Hz (120°), 100-Hz (225°).

0.5 Beat Frequencies in Shepard–Risset Glissandi

There has been little to no discussion in the literature regarding beat frequencies in Shepard–Risset glissandi. But such beat frequencies have certainly been heard. For example, if two identical Shepard–Risset glissandi are time-shifted slightly apart and mixed together, then the instantaneous frequencies in the two glissandi will be slightly different, thus producing beats.

An interesting property of beat frequencies that are produced that way is that they inherently create Risset rhythms, as will be demonstrated. Indeed, because the ratios between concurrent sine-wave frequencies in the glissando are constant, the absolute differences between concurrent sine-wave frequencies are greater when the frequencies are higher. Consequently, the beat frequencies will cyclically accelerate if the glissando is cyclically ascending or cyclically decelerate if the glissando is cyclically descending. The ratios between concurrent beat frequencies will be powers of two, since the ratios between the concurrent sine waves are powers of two. Thus, the beat frequencies will form a Risset rhythm, in which concurrent streams of beats function like whole notes, half notes, quarter notes, eighth notes, etc.

0.6 Barber-Pole Phasing and Barber-Pole Flanging

“Inverted” versions of the Shepard–Risset glissando have been described, in which a continuously sweeping array of frequencies is filtered *out* of the signal, rather than included *in* the signal [24]. For example, a broadband signal (such as noise or distorted drums) can be inputted to an array of logarithmically sweeping notch filters that are octaves apart, with the amount of gain reduction for the notch filters

determined by a Gaussian-like function of the filters’ center frequency. The result is a phasing effect that can seem to perpetually sweep in one direction. That *barber-pole phasing* effect can be converted to a barber-pole flanging effect by systematically varying the number of notch filters and mixing in delayed versions of the signal, with both the number of notch filters and the lengths of the delay lines cyclically varying in synchrony with the notch filters’ sweep cycle.

0.7 The Present Work

The present study demonstrates and discusses various modifications of the classic Shepard [2] and Risset [3] illusions, some of which do not appear to have been discussed previously in the literature. The modifications include: incorporation of non-octave harmonics, use of beat frequencies to create Risset rhythms, manipulation of dynamics and stereo imaging, use of nonmonotonic and anisochronous sequences, use of *rough* (rapidly amplitude-modulated) tone textures, interleaving staccato and legato articulation, and using continuous glides between steps to disambiguate the perceived direction of a Shepard-tritone interval.

The demonstrations are created using a programmatic method that, in its generalized form, can generate a Shepard-tone sequence, Shepard–Risset glissando, or hybrid of the two (i.e., discrete Shepard-tone steps connected by continuous portamento glides). Two original MATLAB functions, and equivalent R functions [25], are provided that are based on that method. These functions can be used to efficiently generate a wide variety of precisely calibrated stimuli for scientific, educational, or artistic purposes.

1 METHOD

1.1 Programmatic Generation of a Shepard–Risset Glissando

- **Step 1: Define a logarithmic sine-wave sweep** that spans exactly ten octaves and generously covers the full range of human hearing. Using a minimum frequency of 19.6 Hz and a maximum frequency of $19.6 \times 2^{10} = 20,070.4$ Hz satisfies those criteria.
- **Step 2: Define a truncated Gaussian vector of amplitude-multipliers** representing the envelope that will be used to gradually roll off the high and low frequencies. The goal is to map the curve to the log-transformed frequency spectrum, but that is equivalent to mapping the curve directly to the sweep in the time domain, since the sweep is logarithmic and spans the full spectrum. Thus, the “filtering” can be achieved by simply multiplying the envelope elementwise by the sweep. The location of the curve’s peak, which determines which region of the spectrum is most emphasized, can be shifted to achieve the desired spectral balance. For instance, Fig. 4 shows a truncated Gaussian envelope, spanning 7 standard deviations, that has been left-shifted so its peak is 1.5 octaves below the log-scale center of the spectrum. The highest octave and lowest

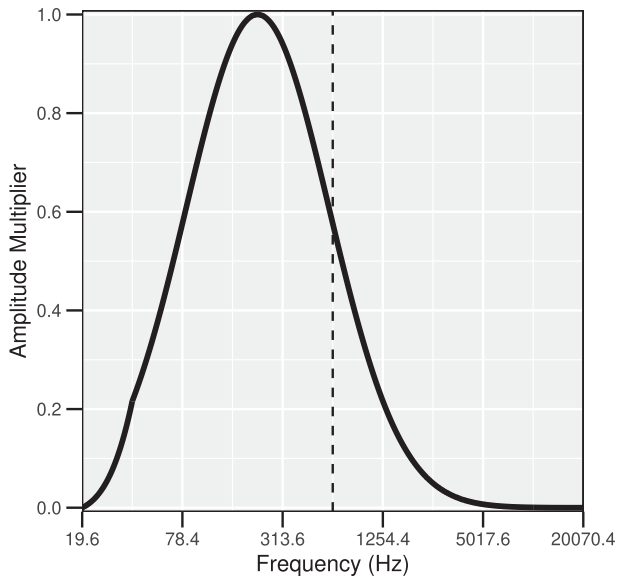


Fig. 4. Example of a truncated Gaussian amplitude envelope mapped to instantaneous log-frequency, with a span of 7 standard deviations and a location-shift of -1.5 octaves. The highest octave and lowest octave of the curve are multiplied by linear ramps to coerce the curve to zero. Vertical gridlines are at octaves. The dashed line marks the log-scale center of the spectrum (627.2 Hz).

octave of the truncated Gaussian are multiplied by linear ramps that coerce the tails to zero to prevent transient “pops” at the truncation points (though the ramp’s effect is too subtle to see on the right side of the figure).

- **Step 3: Reshape the sweep into a ten-column matrix** (one column for each octave) that has a height equal to the number of samples per octave, with time going from top to bottom. That is, the first column contains the first octave of the sweep, and then the sweep “wraps around” to continue down the second column for the next octave, and so on.
- **Step 4: Sum that matrix across octaves to collapse it into a “snippet”** (a seamlessly loop-able vector).
- **Step 5: Normalize the snippet amplitudes** (after DC-offsetting, if desired, to improve gain).
- **Step 6: Loop the snippet** using vector replication.

1.2 Incorporating Additional Harmonics

Consider the concurrently sweeping frequencies in the snippet, which are separated by octaves, as comprising an ensemble of sweeping “fundamental” frequencies. Before normalizing and looping the snippet, one might want to add ensembles of additional harmonics, at a designated interval “above” those fundamentals. Because the octaves-per-second in the snippet is constant, such harmonics can be incorporated by mixing (summing) the snippet with circularly shifted versions of itself.

For example, to add harmonics at 7 semitones (i.e., $7/12$ of an octave) “above” the fundamentals, make a copy of the snippet that has the first $7/12$ of the vector moved to the end (if the glissando is ascending) or has the last $7/12$ of the vector moved to the beginning (if the glissando is descending),

and then sum that shifted vector with the original snippet. Note that because the shift is circular, adding harmonics at 7 semitones “above” the fundamentals is equivalent to adding harmonics at 5 semitones “below” the fundamentals.

1.3 Incorporating Beat Frequencies

Beat frequencies will be created if harmonics are added that are very close to the fundamentals or very close to other added harmonics. For instance, if the snippet is circularly shifted by 0.1 or 0.2 semitone, that creates beating when the shifted snippet is summed with the original snippet. Or if two similarly shifted snippets (e.g., one shifted “up” by 7 semitones and another shifted “up” by 7.2 semitones) are summed with the original snippet, that will create beating as well.

Recall that when a Shepard–Risset glissando is combined with a slightly time-shifted version of itself, the resulting beats inherently create a Risset rhythm. That Risset rhythm can be made polyrhythmic by adding multiple sets of beating harmonics. Or, as will be demonstrated, the beats in the Risset rhythm can be smeared into gradual, rolling swells by adding a tight cluster of several harmonics that are close together (e.g., 0.01, 0.02, 0.03, 0.04, and 0.05 semitones “above” the fundamentals). Conveniently, the high and low beat frequencies in the Risset rhythm will be automatically rolled off because the high and low sine-wave frequencies that produce them were already rolled off when filtering the sweep. Of course, the cost of that convenience is that the roll-offs for the beat frequencies cannot be adjusted independently of the roll-offs for the sine-wave frequencies.

If a constant beat frequency is desired, rather than Risset-rhythm beat frequencies, that can be achieved by generating a second snippet, using a sweep for which the entire vector of instantaneous frequencies is shifted slightly up or down by a scalar number of hertz. Mixing the two snippets together will then produce a constant beat frequency equal to that scalar number of hertz.

1.4 Generalization of the Method

The same six-step method can be generalized to accommodate sequences that are stepped and/or nonmonotonic. In the generalized method, in lieu of computing a sweep per se, a “pseudo-sweep” vector is constructed. That is, a pure-tone sequence (stepped, continuous, or a mixture of both) is repeated ten times, an octave higher or an octave lower each time, such that the entire replicated sequence spans the full ten-octave spectrum (just as the sweep did in the originally described method). Any frequencies in the replicated sequence that are outside of the spectrum should be “wrapped around” to the opposite end of the spectrum, e.g., by using a modulus-after-division function. As with the sweep in the original method, the pseudo-sweep is then reshaped into a samples-by-octaves matrix (so that the different octaves run concurrently), and then the matrix is summed across octaves to create a loop-able snippet, and then the snippet is normalized and looped.

The octaves-per-second is not constant in a stepped and/or nonmonotonic sequence. Consequently, in the gen-

eralized method, the truncated Gaussian envelope must be directly applied as a function of instantaneous frequency, rather than using time as a proxy for frequency. But that is straightforward if a vector of instantaneous frequencies was already computed in the process of generating the pure-tone sequence.

Another consequence of the octaves-per-second not being constant is that any desired additional harmonics must be generated as new pseudo-sweeps, rather than by circularly shifting the snippet. But here again, if a vector of instantaneous frequencies was already computed for the original pseudo-sweep, that is straightforward: If f is the vector of instantaneous frequencies for the fundamentals (i.e., for the original pseudo-sweep), then the vector of instantaneous frequencies for h semitones above the fundamentals is $f \times 2^{h/12}$. As with the frequencies for the original pseudo-sweep, any frequencies that end up outside the spectrum should be wrapped around to the opposite end of the spectrum.

2 SOFTWARE

Two original MATLAB functions are provided: `rissetgliss` and `shepardglide`, each of which may be downloaded from <https://osf.io/rxyd5>. The coding in each function is thoroughly commented to explain the process step-by-step, and the opening comments (which are displayed when the `help` function is used) serve as documentation. Both functions were written in MATLAB 9.14 (2023a) but are compatible with any version that is 9.0 (2015a) or later. No add-ons or toolboxes are required, as verified by MATLAB's dependency analyzer. Equivalent functions in R are also provided, which were written in R Version 4.3.0. To play sound, the R functions require the `audio` package [26], but they have no dependencies for generating the audio data.

In addition to playing the generated signal, each function outputs two objects, each of which is a stereophonic (two-column) matrix representing amplitudes (in the time domain) that are offset and normalized to range exactly from -1 to 1 . The first outputted matrix is the entire signal (the repeating snippet), which has 20-ms fades at the very start and very end to avoid transient pops. The second outputted matrix is just the loop-able snippet, which may be convenient for efficient storage of the audio data or for indefinite real-time looping “on the fly.”

It should be acknowledged that previously designed software, called “Endless Series,” could generate variations of the Shepard and Risset illusions. But as of this writing, that software appears to be unavailable for contemporary operating systems [27].

2.1 rissetgliss

The `rissetgliss` function generates a Shepard-Risset glissando. It uses the method described in this paper, adapted for stereo so that different sets of harmonics may be placed at different positions in the stereo field. The func-

tion takes inputs (all of which have defaults) to specify the following parameters:

- Rate and direction of the glissando.
- Harmonic ensembles, and amplitudes of those ensembles, specified for each stereo channel independently.
- Span and location-shift for the truncated Gaussian amplitude envelope that rolls off the high and low ends of the log-transformed frequency spectrum.
- Starting pitch for the glissando.
- Number of snippet iterations in the signal.
- Sample rate.
- Logical flag indicating whether to play the signal.

2.2 shepardglide

Based on the generalized method described in this paper, the `shepardglide` function generates a series of discrete Shepard-tone steps with seamless portamento glides between those steps. That allows for the creation of stimuli that essentially hybridize Shepard's original stepped version of the illusion with Risset's continuous version. All steps and glides in the sequence are independently adjustable in various ways. The function allows all the same parameter customizations as `rissetgliss` (except glissando rate, glissando direction, and glissando starting pitch, none of which are applicable) and also allows customization of the following:

- Pitch sequence for the steps.
- Duration of each step and each glide.
- Amplitude for each step in each channel.
- Curvatures of the amplitude-ramps across glides that connect steps of different amplitude.

Purely discrete steps can be generated by setting glide durations to 0. But even in that case, the tones that the `shepardglide` function produces are not identical to the tones in Shepard's original study [2]. Most notably, in each tone that Shepard used, all component sine waves were initialized to a phase angle of 0° . By contrast, the `shepardglide` function is designed to accommodate seamless transitions between contiguous steps and thus generates the step sequence as a single continuous wave in which phase relationships between component sine waves vary continuously. Another distinction is that Shepard used a raised cosine pulse, rather than a truncated Gaussian, for the frequency spectrum's amplitude envelope. A raised cosine pulse and a truncated Gaussian have very similar bell shapes, but a truncated Gaussian allows more flexibility because its span may be expanded indefinitely without exhausting the tails of the curve.

3 DEMONSTRATIONS

Twelve example audio files are provided, each of which was produced using the `rissetgliss` or `shepardglide` function. In each example, the snippet

is looped four times, which is the default for both functions. The audio files, as well as the MATLAB and R syntax that can be used to replicate them, are available at <https://osf.io/rxyd5>. Brief discussions of the examples are below.

3.1 Basic Shepard–Risset Glissando

`RissetUp.mp4` demonstrates a simple, “ascending” Shepard–Risset glissando.

3.2 Shepard-Tones Connected by Glides

`ShepardStepGlideUp.mp4` demonstrates what is essentially a hybrid of a stepped Shepard-tone sequence and a continuous Shepard–Risset glissando. The sequence “ascends” the circular octave in stepped semitone increments, with seamless portamento glides between steps.

3.3 Risset-Rhythm Beat Frequencies Embedded in a Shepard–Risset Glissando

`RissetMonauralBeatsDown.mp4` demonstrates a “descending” Shepard–Risset glissando that contains a Risset rhythm comprised of decelerating beat frequencies. The beat frequencies are produced by adding harmonics at 0.2 semitone “above” the fundamentals.

`RissetBinauralSmearsUp.mp4` demonstrates an “ascending” Shepard–Risset glissando that contains a Risset rhythm comprising accelerating beat frequencies. The beats are smeared into gradual, rolling swells by using tight clusters of several harmonics that are close to the fundamentals. Specifically, in the left channel, there are harmonics at 0.06, 0.07, 0.08, 0.09, and 0.1 semitone “above” the fundamentals, and in the right channel, there are harmonics at 0.01, 0.02, 0.03, 0.04, and 0.05 semitone “above” the fundamentals, and both channels contain the fundamentals. Thus, there are both monaural and binaural components to the beating.

3.4 Disambiguation of Shepard-Tritone Direction Using Glides

Recall that the direction of Shepard-tritone intervals can be perceived differently by different listeners in different cases [11–13]. However, one might expect that the direction could be reliably disambiguated by inserting a continuous glide connecting the two notes. And indeed, that can be demonstrated using stimuli generated by the `shepardglide` function.

`shepardTritoneGlideUp.mp4` produces a Shepard-tone C, held for 2 s, followed by Shepard-tone F \sharp , held for 2 s. An upward glide, with a duration of 3 s, connects the two notes to give the impression that the second note is higher than the first. `shepardTritoneGlideDown.mp4` produces exactly the same interval—with exactly the same frequency components in each step—as the previous example. But the glide is downward, to give the impression that the second note is lower than the first.

To confirm that the glides reliably disambiguate the perceived direction of the interval, the two versions of the C-

Table 1. Number of students who responded that the tone sounded higher at the “beginning,” “end,” or “neither,” after hearing a C-to-F \sharp Shepard-tone interval with an upward or downward glide between notes.

		Downward Glide		
		Beginning	End	Neither
Upward Glide	Beginning	5	0	0
	End	75	5	3
	Neither	3	0	0

-to-F \sharp Shepard-tritone were presented once in randomized order to 91 undergraduate students, in an online survey that was completed for extra credit in a cognitive psychology course. The students were not informed of the purpose of the survey in advance and had not previously learned about auditory perception in the course. For each version of the interval, the students were asked, “Did the tone sound higher at the beginning, at the end, or neither?” For the version with the upward glide, 83 students (91.2%) said the tone sounded higher at the end. For the version with the downward glide, 83 students (91.2%) said the tone sounded higher at the beginning. See Table 1 for details.

It is perhaps not surprising that the glides so effectively implied the direction of the interval. After all, it has been shown that when there actually is an objective difference in pitch height between notes (i.e., when tones are not octave-ambiguous), perceived direction is more accurate if portamento glides are inserted [28]. That effect has been analogized to the Gestalt principle of *good continuation*, which dictates that when the connections between segments in a visual stimulus are occluded or ambiguous, perception heuristically favors the simplest, most continuous interpretation [29].

3.5 Stereo Shepard-Tone Illusion in a Nonmonotonic, Anisochronous Sequence

The Shepard-tone illusion has conventionally been demonstrated using an isochronous sequence that implies motion in a consistent direction around the circular octave. But as will be demonstrated here, an overall trend in the ascending or descending direction can be implied even if the implied step-to-step motion is not strictly monotonic. The sequence in `ShepardStereoNonmonotonicUp.mp4` “ascends” a whole-tone scale, but takes one whole-step “down” after every two whole-steps “up.” Additionally, the sequence is made anisochronous by systematically varying the step durations. And different notes are panned to different positions in the stereo field while maintaining approximately equal subjective loudness by using *constant-power panning* [30]. That is, for each step, the squared step amplitudes in the two channels sum to 1. For example, each channel’s step amplitude is set to $\sqrt{0.5}$ when the note is centered, and whenever the step amplitude is set to 0 in one channel, it is set to 1 in the other channel (and vice versa). The glide durations (2 ms each) are short enough not to

produce perceptible portamento but long enough to prevent transient pops that otherwise might occur when abruptly changing pan position from one step to the next.

3.6 Quasi-Continuous, Stereo Shepard-Tone Sequence With Beat Frequencies and a Rough Texture

`ShepardRoughBinauralSmearsUp.mp4` demonstrates a stepped, “ascending” Shepard-tone sequence, but because there are glides between the steps and the step increments are tiny (0.01 semitone), the rise in pitch sounds essentially continuous. The rough, granular texture is achieved by using very short step and glide durations (10 ms each) and making the glides pass through 0 amplitude between steps (as the step amplitudes glide between -1 and 1 on the left and between 1 and -1 , respectively, on the right). A Risset rhythm comprised of smeared, accelerating beats (with monaural and binaural components) is created by using a different tight cluster of harmonics in each channel.

3.7 Stereo Shepard–Risset Glissando in a Harmonically Rich Chord

`RissetBinauralChordUp.mp4` demonstrates an “ascending” Shepard–Risset glissando with harmonics at a major second, perfect fourth, perfect fifth, and minor seventh “above” the fundamentals. Binaural beats are created by making each right-channel harmonic 0.1 semitone “above” or “below” a respective left-channel harmonic. But because the harmonics overall are diverse, the combined beating from all the different pairs is quite polyrhythmic, and thus the beating essentially becomes a wash that manifests more as active stereo spread than as discernable Risset rhythms. To create a more interesting timbre, different amplitudes are used for different pairs of beating harmonics. The span of the truncated Gaussian (i.e., the amount of attenuation in the tails) is relatively reduced, creating a “lusher,” harmonically richer sound that brings out more of the high-frequency “shimmer.”

An interesting perceptual phenomenon can be heard in Shepard–Risset glissandi, such as the one above, that include diverse harmonics. Namely, at least to the present author, frequency components that are not the primary target of attention can take on an illusory drone-like quality, such that they seem to have a sustained pitch even though they are sweeping at the same rate as the other components. One explanation for this subtle illusion is that the pitch of the lower-frequency components is perceived not only in an absolute way but also relative to the other frequencies. Another explanation is that the human auditory system has difficulty tracking constant frequency changes in multiple streams simultaneously and thus simplifies perception of streams that are not the primary target of attention.

3.8 Shepard Tones Connected by Glides, With Hard-Panning

`ShepardDichoticStepGlideDown.mp4` demonstrates a Shepard-tone sequence in which the

notes “descend” chromatically, with portamento glide between notes. The tones in the left channel are at tritone intervals from the tones in the right channel. The peak location of the truncated Gaussian envelope is especially downshifted in this example, thus strengthening the illusion (by making the reentrance of the highest frequencies in each repetition less conspicuous) and creating a deeper, “darker” sound.

`ShepardActivePanStepGlideDown.mp4` is similar to the preceding example. But instead of panning each harmonic ensemble to a different channel, the steps alternate between channels, and the glides actively traverse the panoramic spectrum between steps.

3.9 Stereo Shepard-Tone Sequence With Varied Articulation and Dynamics

`ShepardStereoStaccatoLegatoUp.mp4` demonstrates a more musical example of a Shepard-tone sequence. It not only enriches the timbre with harmonics but also varies the articulation and dynamics. It consists of the following monotonically “ascending” pattern: two staccato notes at the same pitch (the second quieter than the first), followed by a sustained note at the same pitch, which then smoothly glides up to repeat that pattern a half-step “higher,” and so on. The harmonics are essentially a perfect fifth “above” the fundamentals. But to create stereo spread, the left-channel harmonics are given slightly different frequencies and amplitudes from the right-channel harmonics.

4 CONCLUSION

This study demonstrates several modifications of the Shepard-tone sequence [2] and the Shepard–Risset glissando [3]. Some of those modifications do not appear to have been previously discussed in the literature.

For example, discrete Shepard-tone steps were combined with continuous Shepard–Risset glides in various ways. And in particular, inserting glides between Shepard-tones that were a tritone apart appeared to reliably disambiguate the perceived direction of motion between notes, thus overriding the tritone paradox. Future studies may investigate that phenomenon in more detail by systematically varying the parameters (such as frequency and duration) of the steps and glides.

Another modification that was demonstrated is the use of beat frequencies to embed Risset rhythms into Shepard–Risset glissandi. There are a variety of “smeared” beat patterns that may be produced by combining different frequencies in different ways, so there remains much to explore in that regard.

More generally, it is hoped that the techniques and functions that are presented here will be useful to scientists, educators, sound engineers, and sound artists who wish to use these types of stimuli in their work. It is likely that there are many more modifications, of both scientific and creative interest, that can be produced using octave-ambiguous tones.

5 REFERENCES

- [1] R. M. Aarts, “The Application of Illusions and Psychoacoustics to Small Loudspeaker Configurations,” in *Proceedings of the AES UK 22nd Conference: Illusions in Sound* (2007 Apr.), paper 7. <https://www.aes.org/e-lib/browse.cfm?elib=17281>.
- [2] R. Shepard, “Circularity in Judgments of Relative Pitch,” *J. Acoust. Soc. Am.*, vol. 36, pp. 2346–2353 (1964 Dec.). <https://doi.org/10.1121/1.1919362>.
- [3] J. C. Risset, “Pitch Control and Pitch Paradoxes Demonstrated With Computer-Synthesized Sounds” *J. Acoust. Soc. Am.*, vol. 46, no. 1A, p. 88 (1969 Sep.). <https://doi.org/10.1121/1.1973626>.
- [4] S. W. Draper, “The Penrose Triangle and a Family of Related Figures,” *Perception*, vol. 7, no. 3, pp. 283–296 (1978 Jun.). <https://doi.org/10.1068/p070283>.
- [5] W. J. Dowling, “Scale and Contour: Two Components of a Theory of Memory for Melodies,” *Psychol. Rev.*, vol. 85, no. 4, pp. 341–354 (1978 Jul.). <https://doi.org/10.1037/0033-295X.85.4.341>.
- [6] E. Vernooij, A. Orcalli, F. Fabbro, and C. Crescentini, “Listening to the Shepard–Risset Glissando: The Relationship Between Emotional Response, Disruption of Equilibrium, and Personality,” *Front. Psychol.*, vol. 7, paper 300 (2016 Mar.). <https://doi.org/10.3389/fpsyg.2016.00300>.
- [7] P. Patrício, “From the Shepard Tone to the Perpetual Melody Auditory Illusion,” in *Proceedings of the 9th Sound and Music Computing Conference*, paper 38 (Copenhagen, Denmark) (2012 Jul.). <https://hdl.handle.net/10400.14/34929>.
- [8] J. Giangrande, B. Tuller, and J. A. S. Kelso, “Perceptual Dynamics of Circular Pitch,” *Music Percept.*, vol. 20, no. 3, pp. 241–262 (2003 Mar.). <https://doi.org/10.1525/mp.2003.20.3.241>.
- [9] F. Ragozzine, “Correspondence in Perception of the Tritone Paradox and Perfect-Fifth/Perfect-Fourth Intervals,” *Music Percept.*, vol. 30, no. 4, pp. 391–406 (2013 Apr.). <https://doi.org/10.1525/mp.2013.30.4.391>.
- [10] D. Deutsch, W. L. Kuyper, and Y. Fisher, “The Tritone Paradox: Its Presence and Form of Distribution in a General Population,” *Music Percept.*, vol. 5, no. 1, pp. 79–92 (1987 Oct.). <https://doi.org/10.2307/40285386>.
- [11] D. Deutsch, “The Tritone Paradox: Effects of Spectral Variables,” *Percept. Psychophys.*, vol. 41, no. 6, pp. 563–575 (1987 Nov.). <https://doi.org/10.3758/BF03210490>.
- [12] D. Deutsch, “The Tritone Paradox: a Link Between Music and Speech,” *Curr. Dir. Psychol. Sci.*, vol. 6, no. 6, pp. 174–180 (1997 Dec.). <https://jstor.org/stable/20182482>.
- [13] B. H. Repp, “Spectral Envelope and Context Effects in the Tritone Paradox,” *Perception*, vol. 26, no. 5, pp. 645–665 (1997 May). <https://doi.org/10.1068/p260645>.
- [14] R. A. Mursic, B. E. Riecke, D. Aphorpe, and S. Palmisano, “The Shepard–Risset Glissando: Music That Moves You,” *Exp. Brain Res.*, vol. 235, pp. 3111–3127 (2017 Jul.). <https://doi.org/10.1007/s00221-017-5033-1>.
- [15] R. A. Mursic and S. Palmisano, “Something in the Sway: Effects of the Shepard–Risset Glissando on Postural Activity and Vection,” *Multisens. Res.*, vol. 35, no. 7–8, pp. 555–587 (2022 Sep.). <https://doi.org/10.1163/22134808-bja10081>.
- [16] B. Johnston, M. Norris, A. Kapur, and J. Murphy, “Deceptive Sounds: Auditory Illusions & Artefacts in Sound Art,” in *Do Androids Dream of Computer Music? Proceedings of the Australasian Computer Music Conference*, pp. 11–18 (Adelaide, Australia) (2017 Sep./Oct.). <https://computermusic.org.au/proceedings/ACMC2017-proceedings.pdf>.
- [17] E. Rapan, “Shepard Tones and Production of Meaning in Recent Films: Lucrecia Martel’s *Zama* and Christopher Nolan’s *Dunkirk*,” *New Sound-track*, vol. 8, no. 2, pp. 135–144 (2018 Sep.). <https://doi.org/10.3366/sound.2018.0126>.
- [18] B. Kattelman, “The Sound of Evil: How the Sound Design of *Hereditary* Manifests the Unseen and Triggers Fear,” *Horror Stud.*, vol. 13, no. 1, pp. 133–148 (2022 Apr.). https://doi.org/10.1386/host_00050_1.
- [19] S. Denjean, R. Kronland-Martinet, V. Rous-saire, and S. Ystad, “Zero-Emission Sonification Strategy Based on Shepard–Risset Glissando,” in *Proceedings of the International Symposium on Computer Music Multidisciplinary Research: Perception, Representations, Image, Sound, Music*, pp. 709–724 (Marseille, France) (2019 Oct.). https://doi.org/10.1007/978-3-030-70210-6_46.
- [20] T. Dupré, S. Denjean, M. Aramaki, and R. Kronland-Martinet, “Spatial Integration of Dynamic Auditory Feedback in Electric Vehicle Interior”, *J. Audio Eng. Soc.*, vol. 71, no. 6, pp. 349–362 (2023 Jun.). <https://doi.org/10.17743/jaes.2022.0087>.
- [21] J. C. Risset, “Pitch and Rhythm Paradoxes: Comments on ‘Auditory Paradox Based on Fractal Waveform’,” *J. Acoust. Soc. Am.*, vol. 80, no. 3, pp. 961–962 (1986 Sep.). <https://doi.org/10.1121/1.393919>.
- [22] G. Oster, “Auditory Beats in the Brain,” *Sci. Am.*, vol. 229, no. 4, pp. 94–103 (1973 Oct.). <https://jstor.org/stable/24923223>.
- [23] M. J. Barratt, A. Maddox, N. Smith, et al., “Who Uses Digital Drugs? An International Survey of ‘Binaural Beat’ Consumers,” *Drug Alcohol Rev.*, vol. 41, no. 5, pp. 1126–1130 (2022 Mar.). <https://doi.org/10.1111/dar.13464>.
- [24] F. Esqueda, V. Välimäki, and J. Parker, “Barber-pole Phasing and Flanging Illusions,” *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx)*, pp. 87–94 (Trondheim, Norway) (2015 Dec.). <https://dx.doi.org/10.13140/RG.2.1.2489.7368>.
- [25] R Development Core Team, *R: A Language and Environment for Statistical Computing*, version 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria, 2023). <https://www.R-project.org>.
- [26] S. Urbanek, *audio: Audio Interface for R* (2021 Nov.). <https://cran.r-project.org/web/packages/audio>.
- [27] O. Larkin, “Endless Series v3” (2012). <https://olilarkin.co.uk/index.php?p=eseries>.
- [28] A. S. Bregman and G. L. Dannenbring, “The Effect of Continuity on Auditory Stream Segregation,” *Per-*

cept. Psychophys., vol. 13, no. 2, pp. 308–312 (1973 Jun.).
<https://doi.org/10.3758/BF03214144>.

[29] D. Deutsch, “The Perception of Auditory Patterns,” in W. Prinz and B. Bridgeman (Eds.), *Handbook of Perception and Action*, vol. 1,

pp. 253–296 (Academic Press, London, UK, 1996).
[https://doi.org/10.1016/S1874-5822\(96\)80011-9](https://doi.org/10.1016/S1874-5822(96)80011-9).

[30] R. Orban, “Notes on Panpots,” *J. Audio Eng. Soc.*, vol. 19, no. 11, pp. 935–939 (1971 Dec.).

THE AUTHOR



Andrew V. Frane

Andrew Frane is an Assistant Professor of Psychology at Occidental College, Los Angeles, CA, where he teaches psychology, research methods, and statistics. He is also a lecturer at the University of California, Los Angeles, where he teaches psychology and MATLAB programming. And

he is a lecturer at the Colburn Conservatory of Music in Los Angeles, where he teaches an original course on the psychology of music and sound. He received his Ph.D. in psychology from the University of California, Los Angeles in 2019.