Academic Press.

Figure 1. Examples: Melodies 2 and 4 (diatonic versions).

MUSICAL FORCES, STEP COLLECTIONS, TONAL PITCH SPACE, AND MELODIC EXPECTATION

Steve Larson
School of Music
University of Oregon
Eugene, OR 97403
USA

A theory of expressive meaning in music

A central claim of the theory is that we give meaning to musical sounds by assigning them to categories. The theory represents this claim with the formulation "to *hear* x *as* y" (where x is a musical sound and y is a musical meaning). For example, we might *hear* a pattern of pitches *as* an ascending gesture, or we might *hear* a pattern of durations *as* a syncopated rhythm.

More specifically, the theory argues that we tend to hear music as purposeful action with a dynamic field of musical forces. It is the purposeful quality of this action that led Schenker to speak of "the will of the tones", but, as Fred Maus has observed [1990] such a feeling of purpose does not require that these actions be attributed to a specific agent. Drives that we experience in listening—especially the tendency to hear a percept in terms of the simplest possible interpretation—lead us to feel the "will of the tones" operating in a dynamic field of musical forces that I call "gravity", "magnetism", and "inertia".

Musical forces

Musical gravity is the tendency of an unstable note to *descend*. In other words, in a context where a given note is heard as unstable, listeners will tend to hear that note as one that is pulled down by gravity.

Musical magnetism is the tendency of an unstable note to move (up or down) to the *nearest* stable pitch. Furthermore, magnetism is affected by distance—the closer we get to a goal, the more it attracts us. For example, in a context where D and G are heard as stable and the F between them is heard as unstable, listeners experience musical magnetism as a tendency of the F to ascend to G (because F is closer to G than it is to D). If that F should then move to F#, the magnetic force drawing us to G will intensify.

Musical inertia is the tendency of a pattern of musical motion to continue in the *same* fashion. For example, if a pattern of musical motion begins "C-D-E, D-E-F", then listeners may experience musical inertia as a tendency to continue the pattern "E-F-G", etc. What is meant by "same" depends upon how that musical pattern is represented in our internal hearing. For example, if one listener hears the pattern of pitches "C-D-C-D-E-D-E-F" as four-note groups (i.e., "C-D-C-D, E-D-E-F"), then that first listener will experience inertia as the tendency of the pattern to continue in fours (e.g., "C-D-C-D, E-D-E-F, G-A-G-etc."); but if another listener hears the same pattern as three-note groups (i.e., "C-D-C, D-E-D, E-F-"), then that second listener will experience inertia as the tendency of the pattern to continue in threes (e.g., "C-D-C, D-E-D, E-F-E, etc."). Gestalt psychology calls this tendency "the law of good continuation" [Koehler 1947]. Meyer's work on emotion in music [1956] and Narmour's implication-realization theory [1990, 1992] grant an important role to what I call inertia.

Musical hierarchies

Listeners tend to organize musical passages into what Jamshed Bharucha [1984] calls "event hierarchies". The pattern "C-G-F#-G, D-G-F#-G, E" contains the simpler pattern "C-G, D-G, E", which in turn contains the simpler pattern "C-D-E". Each of these patterns may be called levels of a hierarchical hearing of the pattern that contains them all.

The musical forces operate on all levels of pitch structure in music. Thus, an inertial tendency could develop between notes that are adjacent on the surface of the music—for example, inertia might imply that "C-D-" would continue ascending as "C-D-E". But the same inertial tendency could also develop between these notes even if they appeared on a more abstract level of musical structure—for example, in the previous pattern "C-G-F#-G, D-G-F#-G", inertia might lead us to hear the pattern continue as "C-G-F#-G, D-G-F#-G, E" (or as "C-G-F#-G, D-G-F#-G, E-G-F#-G").

Stability and the displacement of auralized traces

We experience the musical forces as acting upon notes that we hear as unstable. But what does it mean to hear a note as unstable? To explain this, we need the terms "auralize", "trace", and "displace".

To *auralize* means to hear internally sounds not physically present.

The term *trace* means the internal representation of a note that is still melodically active. If I play a short note on the piano for you, it may continue to "ring" in your ears even after the note has stopped physically vibrating. The fact that you can continue to pay attention to the pitch of a note even after its sound has stopped physically sounding means that that sound has left a kind of trace in your memory.

Your attention to that trace can be *displaced* by subsequent traces. If I play one note and then another that is a step away, you may be able to recall both notes. But there is some sense in which the second note will displace the trace of the first (the fact that this displacement occurs more readily with notes that are a step apart leads to a discussion of "steps and leaps" below).

Like the musical forces, displacement also operates on various levels of musical structure. For example, in the first three notes of the pattern "B-A-G, C-B-A, D", the trace of the B may be displaced by A and then by G. However, this pattern contains—at a more basic hierarchical level—the pattern "B-C-D". And—at that more basic hierarchical level—the trace of the B continues until it is displaced by the C.

These terms allow us to distinguish between *stable* and *unstable* notes. To hear a note as unstable means to auralize a more stable pitch to which it tends to move and a path (usually involving step-wise motion) that would take it there, displacing its trace.

Stability also generates structured pitch relationships with hierarchical levels. Bharucha [1984] calls such atemporal structures "tonal hierarchies". Carol Krumhansl's [1990] "major-key profile" is generally understood as a tonal hierarchy of the relative stability of tones within a major-key context.

Deutsch and Feroe's [1981] "alphabets" and Lerdahl's [1988] "tonal pitch space"

Diana Deutsch and John Feroe [1981] have advanced a model of music cognition that involves what they call "alphabets" of pitches. Their model describes event hierarchies in terms of these alphabets (which include the chromatic scale, the major scale, certain chords, etc.) and operations on those alphabets. The alphabets themselves can also form tonal hierarchies. For example, the chromatic scale contains the hierarchically superior diatonic scale, which contains the hierarchically superior tonic major triad, and so on. The operations on these alphabets indicated motion from one element of the alphabet to another, usually by a relationship of sameness (repetition) or successorship (to the next adjacent element of the alphabet).

Lerdahl's "basic pitch space" uses the hierarchical nature of Deutsch and Feroe's alphabets to make a statement about stability.

Steps and leaps

My definition of "unstable" suggests that intervals may be divided into two classes—steps and leaps. Augmented unisons, minor seconds, and major seconds are called steps. (A diminished third is a step whose name suggests that it is heard as a leap.) Minor thirds and larger intervals are called leaps. (An augmented second is a leap whose name suggests that it is heard as a step.)

In a melodic step, the second note tends to displace the trace of the first, leaving one trace in musical memory. In a melodic leap, the second note tends to support the trace of the first, leaving two traces in musical memory. Many experiments in music perception and published discussions in music cognition support this distinction between steps and leaps, or they seem to rely implicitly on such a separation [for a summary, see Bregman 1990]. And several recent theoretical discussions build explicitly upon this distinction [Westergaard 1975; Larson 1987; Dembski 1988].

My "step collections"

A step collection is a group of notes that can be arranged in ascending pitch order to satisfy the following two conditions: [1] every adjacent pair of notes is a step (that is, a half step or a whole step) apart; and [2] no non-adjacent pair of notes is a step apart. The second condition can be modified slightly to produce a third condition, true of all "proper" step collections: (3) no two pitches—nor any of their octave equivalents—that are not adjacent in the list (except the first and last) are a step apart. The first condition ensures that the collection can be heard as a complete filling in of a musical space (this follows from our recognition that melodic leaps tend to leave the "trace" of a note "hanging" in our musical memories). The second condition ensures that no note will be heard as redundant in the filling of that space (this also reflects our desire to avoid confusion and the fact that either a whole step or half step can be heard as a step). The third condition grants a role to octave equivalence, ensuring that the first and last pitches are less than an octave apart and that adding octave equivalents to a proper step collection will result in a step collection.

Melodic expectation and musical forces

These definitions of gravity, magnetism, and inertia lead to two basic assertions.

The first assertion is that melodic expectations in tonal music depend on the iterated operation of these forces on various hierarchical levels of musical structure. By "iterated operation of these forces", I mean a multi-stage process like the following: (1) take a simple (but in some sense incomplete) melodic pattern, (2) follow the implications of one of the musical forces until a certain degree of stability is achieved, (3) take the resultant pattern, and (4) follow the implications of another of the musical forces until an even greater degree of stability is achieved.

The second basic assertion is that goal-direction is a very important aspect of tonal music, and thus that the patterns of musical motion in which the final note is most strongly predicted by the musical forces are the most fundamental melodic patterns.

But how do we test such assertions? One cannot just pick some pieces and then count the frequency with which patterns occur in them. To count the number of patterns, one must first find them, a process that requires analytically separating the piece into patterns. Finding patterns on all hierarchical levels requires further analysis. In the end, such counting might prove more about the intellectual theory behind the analysis than about the aural and emotive experience of musical forces. Furthermore, we must not confuse frequency with importance—music does not always do what we expect it to do, so it does not necessarily follow that just because a pattern is "more fundamental" it will happen more often. Music often creates some of its most salient effects by diverging from our expectations. In music, making clear distinctions between frequency, structural importance, and salience is a fascinating but complicated problem!

We may find more persuasive support for the theory in a comparison between the behavior of subjects in a psychological study and the behavior of a computer model based on the theory. This comparison exploits the importance of expectation, suggesting that stronger patterns play a more fundamental role in listeners' expectations.

A psychological study

In his dissertation study, William Lake [1987] asked music students at the University of Michigan to sing simple continuations. First, to establish a context, he played a chord and a scale for them. He then played a two-note beginning. Finally, he asked the students to sing that two-note beginning, "adding another tone or tones of your own choosing". After excluding the few responses in which subjects did not correctly reproduce the two-note beginning, he tabulated the frequency with which each third note (the first note of each continuation) was sung (excluding those notes sung in less than 12 per cent of the continuations).

On first inspection, Lake's results seem to support the theory advanced here—that is, the continuations seem fairly easy to rationalize in terms of musical forces. However, comparing Lake's results with the behavior of computer models based on the operation of musical forces provides a better test of the theory.

What Next [1993] a first computer model of musical forces

Elsewhere [Larson 1993], I have described my computer program, called What Next, which generates melodic continuations from two-note beginnings like those that Lake gave to his subjects. What Next uses each of the musical forces to specify a motion in one direction through one of Deutsch and Feroe's alphabets (the "reference level") that continues until arriving at a note contained at a specified deeper level (the "goal level").

A comparison of What Next [1993] with Lake [1987]

Comparisons between the predictions of What Next and the continuations sung by Lake's subjects suggest that gravity, magnetism, and inertia play an important role in melodic expectation.

What Next applies the forces in only the simplest way. In applying gravity and magnetism it considers only the second note of the pattern. It does not perceive or create embellishments of hierarchical musical structure. And, in each prediction, it applies only one force only one time.

In light of these limitations, the performance of What Next seems striking. Consider the seventy-five patterns that do not end on the tonic pitch. These patterns led subjects to sing up to four different third notes (excluding continuations sung less than twelve per cent of the time) for a total of 165 patterns. What Next predicts 132 of these (80%). The more that subjects agreed on a continuation, the more they agreed with What Next. For patterns in which subjects sang only one third note, the predictions of What Next included that continuation every time (100%). For patterns in which subjects sang two different third notes, What Next still did well (88%).

Another look at Lake [1987]

Krumhansl, Carol L. 1990. *Cognitive Foundations of Musical Pitch*. New York: Oxford

... levels requires further analysis. In the end, such counting might prove more about the intellectual theory behind the analysis than about the aural and emotive experience of musical forces. Furthermore, we must not confuse frequency

Another look at Lake [1987]

However, this comparison is based on an examination of only the first response tones (Lake felt that as we got further from the stimulus, it would exert less influence on the response tones). But What Next generates complete patterns of varying lengths. To better test the theory advanced here, we must compare the entire responses of What Next to the entire responses of Lake's subjects. So, with the help of Erick Carballo, I have compiled the entire responses of Lake's subjects into computer-readable form.

Before comparing those entire responses to the responses of What Next, I asked some simple questions about the responses as a group. Which notes did they *end* on? Not surprisingly, the distribution of the pitches for the final response tones bears a striking resemblance to Krumhansl's major-key profile. How *consistent* was each subject? Because each subject had two chances to respond to the same stimulus, we can determine how consistent each subject was by asking how often that subject gave a single response twice. This also allows us to rank the subjects in terms of consistency. How representative (or conformant) is each subject? We can also determine how *agreeable* each subject was by asking how many times that subject's responses match another subject's response. This allows us to rank the subjects in terms of agreeability. The rankings (in terms of consistency and in terms of agreeability) are almost identical (yet they diverge strongly from a ranking based solely on response length.)

I then asked the same questions of the responses given by two "artificial subjects". The first artificial subject, called "random-next", generates two continuations for each stimulus by adding random notes to create responses of varying lengths (while its response lengths are random, its computer code ensures that the distribution of its response lengths will match that of Lake's subject's). The second artificial subject, called "next-generation", draws two random responses for each stimulus from the pool of the continuations generated by What Next. When the artificial subjects are compared with Lake's subjects, random-next ranks at the bottom of the list (both as least consistent and least agreeable) and next-generation ranks near the top of the list (and its rank in terms of consistency tends to agree with its rank in terms of agreeability).

Acknowledgments

The Center for Research on Concepts and Cognition at Indiana University supported research on this project. My colleagues at CRCC, especially Doug Hofstadter, Gary McGraw, and James Marshall, offered thoughtful ideas on music and considerable help with computer programming. Bill Lake offered help and advice. Erick Carballo helped translate Lake's transcriptions of his subjects' responses from musical notation into integers that the computer could read.

What Next was written in MacScheme 4.0. Scheme is a dialect of LISP invented by Guy L. Steele, Jr. and Gerald Jay Sussman. Scheme for the Macintosh was developed at MIT (© 1984). MacScheme is published by Lightship Software, Inc. (© 1992).

References

Bharucha, Jamshed J. 1984. Event Hierarchies, Tonal Hierarchies, and Assimilation. *Journal of Experimental Psychology* 113/3 421-425.

Deutsch, Diana and John Feroe. 1981. The Internal Representation of Pitch Sequences in Tonal Music. *Psychological Review* 88/6 503-522.

Koehler, Wolfgang. 1947. *Gestalt Psychology*. New York: Liveright.

Krumhansl, Carol L. 1990. *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

Lake, William. 1987. Melodic Perception an Cognition: The Influence of Tonality. Ph.D. dissertation, University of Michigan.

Larson, Steve. 1987. Questions About the Ursatz: A Response to Neumeyer. *In Theory Only* **10/4** 11-31.

Larson, Steve. 1993. Modeling Melodic Expectation: Using Three 'Musical Forces' to Predict Melodic Continuations. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (Hillsdale, NJ: Lawrence Erlbaum Associates), 629-634.

Lerdahl, Fred and Ray Jackendoff. 1983. *A Generative Theory of Tonal Music*. Cambridge, MA and London: The MIT Press.

Lerdahl, Fred. 1988. Tonal Pitch Space. *Music Perception* 5/3 315-349.

Maus, Fred Everett. 1990. Humanism and Musical Experience. Ph.D. dissertation, Princeton University.

Meyer, Leonard. 1956. *Emotion and Meaning in Music*. Chicago: University of Chicago Press.

Narmour, Eugene. 1990. *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model*. Chicago and London: The University of Chicago Press.

Narmour, Eugene. 1992. *The Analysis and Cognition of Melodic Complexity: The Implication-Realization Model*. Chicago and London: The University of Chicago Press.

Shepard, R. N. 1982. Structural Representations of Musical Pitch. In Deutsch 1982, 343-390.

Westergaard, Peter. 1975. *An Introduction to Tonal Theory*. New York: W. W. Norton.