

Introduction to Harmony and Formal Organization

A Vibrating String and Harmony

Examining the sound made by a vibrating string (or the column of air in a wind instrument) reveals that there are different modes of vibration that make up the sound waves that reach our ears. The string vibrates as a whole between the points at which it is held at tension. The length and tension of the string determines the pitch that we hear as the string vibrates, producing its fundamental pitch. Whether plucked or set continuously in motion by a bow, the string also vibrates in secondary modes, in halves, thirds, quarters, fifths, and smaller proportions, “partial vibrations,” each in lesser amplitude than the more fundamental segments. The result gives the string or the column of air its characteristic timbre, so that we can distinguish a violin from a flute or an oboe. The difference arises from partial vibrations, or ‘overtones,’ that are unevenly loud in different instruments, but which all share the overriding sound of the fundamental pitch. The partial vibrations can be examined and identified very precisely through their mathematical vibration ratios, a process aided by computers. It is possible to mimic the timbre of any instrument by computer generated sound, once the computation has been made.

The ‘overtone series’ discloses something about what we hear as harmony in musical sounds. The first overtone, vibrating at the ratio of $2/1$, is the octave of the fundamental, the second overtone, ratio $3/2$, is a fifth higher, the third overtone, ratio $4/3$, is a fourth, and the fourth overtone, ratio $5/4$, sounds a major third. If we stop our examination at this point (and why not, since the overtones are more faint as we try to distinguish them further), we have a physical explanation of the harmony of a four-note major chord, a chord with root (fundamental sound), fifth, and major third (in between the root and fifth) with an octave of the root on top of the chord, at the interval of a fourth above. This harmony is derived from nature, rather than invented by imagination.

Examining partial vibrations beyond the fifth overtone becomes complicated because the vibration ratios do not give us sounds that make useful musical pitches, although some of these higher frequencies do coincide with notes in our musical system. One more overtone, the fifth in the series, is already dodgy, because it is not quite in tune with the minor third that we use in our diatonic scale, and further distant overtones are at closer vibration ratios than the tones and semitones that make up our scales.

Tonality

The Major Mode

The word ‘tonality’ denotes a relationship between pitches, specifically the relationship that defines a key in which one pitch is the center of melodies and chordal relationships. Tonality is represented by a diatonic scale, made up of whole tones and semitones. C major, for instance, has the notes C, D, E, F, G, A, B, and C again, one octave higher. In an ascending scale each is one tone higher in pitch except between E and F, B and C, which are semitones apart.

The main harmonic functions in tonality are those between the dominant, a chord generated on the fifth degree of the scale (V), the subdominant, a chord on the fourth degree (IV), and the tonic, on the first (I). The tonic of C major is C, the subdominant F, and the dominant G. Each major chord is built on three notes, a ‘triad,’ acoustically generated in the overtone series of the fundamental note. The C tonic triad is C, E, G; the subdominant’s is F, A, C; the dominant’s is G, B, D. In the major mode all three of these chords are major triads, a major third from root to third, and a minor third from third to fifth.

Each has a particular role in the grammar of tonality, the most dynamic of which is that of the dominant chord. Its movement derives in part from the linear direction of its component notes. Listening to a succession of notes rising from the note c on the white keys of a piano through d, e, f, g, a, and stopping on b, it is usually clear to us that we have not completed the series: we want to hear the next semitone up, another C, one octave higher than the starting pitch. The vibration ratio of the two c’s is 2:1, they seem to repeat one another, and stopping on B clearly leaves our mind unsatisfied. We give the name “leading tone” to B, the seventh in the scale. A careful ear will detect tendencies in the directions of other notes in the scale, and melodies are created by stringing them together in meaningful successions. The direction in which chords ‘progress’ is a compound of the linear tendencies of notes of the scale amplified by conventions akin to linguistic grammar that allow us to string words together in sentences.

The dominant chord is dynamic, like verbs. It requires resolution, we want to hear it resolved, and its most satisfactory resolution is to the tonic. The subdominant chord is less dynamic, and rather gently resolves to the tonic (if this happens directly it will remind you of “amen” sung at the end of a hymn). It also resolves to the dominant chord and the succession of subdominant, dominant, tonic is a strong cadential resolution that definitively identifies a key.

Chords come in strong colors when they are in “root position”, that is when the generating pitch of the triad is the lowest sounding note, and the intervals above it are a third and a fifth. They may be attenuated by changing the place of notes relative to one another. If the third of the chord (e in the c chord) is in the bass, with g and c above it, the result is still a c chord, but it is “in first inversion,” a milder version of the tonic. “Second inversion” places the fifth of the chord, g in the

lowest voice, with c and e above it, at the intervals respectively a fourth and a sixth. This is a still more mild version than the first inversion. A mixture of root position, first, and second inversion chords add to the repertory of colors from the primary functional chords.

Chords based on the other notes of the scale are secondary in function to these three. The supertonic chord (II), one tone higher than the tonic, acts as a version of the subdominant chord in progressions to the dominant, usually in its first inversion. The submediant, (VI) halfway down from the tonic to the subdominant on the sixth degree of the scale, is a substitute in some circumstances for the tonic. Both supertonic and submediant triads are minor chords, the third of each is a minor third above the root and a major third from the third to the fifth. They contrast in the color of their sounds to the major chords for which they substitute. The mediant (III, also minor) is on the third degree of the scale. It is halfway up to the dominant, is a compound of notes from the tonic and dominant chords, and is weak in its function, but sometimes colorful. The last remaining triad is on the seventh degree of the scale and is composed of a minor third from its root, and another minor third to a note one semitone less than a perfect fifth (counting from the seventh tone of the scale). This interval, called a ‘diminished fifth’ or ‘tritone’ has a unique and peculiar sound, and is considered a dissonance.

In medieval music theory the tritone was called ‘the devil in music’ and given very special treatment. In tonal music it is a powerful tool both for confirming tonality as well as for escaping to another key because it can be spelled differently for quite different resolutions of the dissonance it represents. The tritone can help to confirm a key if it is considered a part of the dominant chord instead of a triad formed on the seventh degree of the scale. The dominant then includes another (minor) third beyond the root to F, and becomes a seventh chord (G up to F is the interval of a seventh). The power of the dominant is increased by adding dissonance (both tritone and a seventh) to linear direction. The tritone B to F resolves by two semitone steps to C and E.

A diminished fifth and an augmented fourth (in equal temperament, our modern tuning system the same sound, but called an augmented fourth) are the same sound. If we were to spell it C-flat to G-double flat instead of B-natural to F-natural (the same sound but called an augmented fourth) the dissonance could resolve by moving two semitones to B-flat and G-flat, a perfectly good resolution to a far-different key. Spelling indicates thinking, in this case, and the result explains the ambiguity of sound that resides in the interval of “the devil in music.”

The Minor Mode

The minor mode changes the sound and emotional range of music considerably from that of the major. The scale of the minor uses the same series of tones and semitones as the major but in

a different sequence. Each major key has its relative minor mode of which the key note is a minor third interval below that of the relative major. Thus A minor is the relative of C major, and the tonic (a), subdominant (d) and dominant (e) chords are minor triads on their respective scale degrees.

There are three scales for the minor mode. The ‘natural’ minor scale, unaltered diatonic notes from a to a, is antique and belongs to modal music of the sixteenth century. The ‘harmonic’ minor, the second minor scale, is required to provide a workable dominant chord, the sound of which must be major because the leading tone of the scale must be a semitone from the tonic, causing an augmented second (a rather oriental sound) from the 6th to the (raised) 7th degree. A third version of the scale, ‘melodic minor,’ is recognized when, for melodic reasons, the augmented interval is avoided by raising both the sixth and seventh notes of the ascending scale by a semitone, and these notes resume the ‘natural’ scale in descending.. The grammatical functions of the tonic, subdominant and dominant harmonies are unchanged from the major to minor modes, but their sounds and harmonic power are enhanced, the subdominant more strongly resolves to the (major) dominant chord, and the major dominant contrasts vividly with the minor tonic to which it resolves.

For formal structures in the minor mode the most significant contrast to our ears is between the minor tonic chord and its relative major (the interval of a third higher). Consequently this becomes basic to the form rather than that of tonic to dominant. Recapitulations repeat both first (minor mode) and second key (major mode) melodies and harmonies in the minor mode, a more dramatic resolution than that of the two major-mode contrasts of sonata-form.

This is then the repertory of harmonies and their proper grammatical function in defining and remaining in a key. A major scale can be built on any musical pitch if the tones and semitones of the C major scale are repeated in the same order. Chromaticism (sharps or flats altering pitches) can be used if the alterations temporarily heighten the importance of functional diatonic chords. For instance, a composer may accentuate the importance of the dominant by altering the supertonic chord from minor mode to major by replacing F-natural with F-sharp when the supertonic resolves to the dominant that (in turn) resolves to the tonic. This serves as “dominant of the dominant”, which increases in force as it approaches (eventually) the tonic. All dominant chords are major.

Although chromaticism adds color and variety to diatonic tonal harmonies, it can also disrupt tonality and open paths to modulation, an essential process in extending musical ideas in large-scale formal structures. Good grammar in chord progressions is essential to remaining in a key, but “bad grammar” is another means of disrupting tonality and changing from one key to another. In the larger perspective of a formal structure, changes from one key to another are

essential, and we expect to move away from one key to another by modulation. We'll now move to definitions of form in the 'classic' period (Mozart, Haydn, and Beethoven), in which modulations to different keys bring contrasts that define the architecture of large forms.

The Classic Two-Reprise Form

This is the basic form of much dance music of the 18th century, particularly of minuets. There are two sections, "reprises" (both of which are repeated), in the form, hence its name. The harmonic structure contrasts tonic (I) and dominant (V). The simplest version of this form presents a first phrase (A) beginning and ending in the tonic, and a slightly different version has a phrase ending on the tonic and a second phrase ending on the dominant. The performers see this section graphically set off in their notation by a "double bars", $||: A :||$, which indicate that the music should be repeated. Then another section (B), also to be repeated, contains a few phrases that begin on the dominant but end by resolving the dominant chord in a return to the A section with a rhythmic and harmonic resolution on the tonic. $||: B + A :||$. The second of the two reprises is often twice the length or more of the first.

$$||: A :||: B + A :||$$
$$I---I \quad V---V \quad --I$$

Or $||: A :||: B + A :||$

$$I---V \quad V---V \quad --I$$

The two-reprise form may be extended by enlarging the phrases or by including more phrases while keeping the same over-all harmonic structure. In complex two-reprise forms, there may be some variety in the harmonic structure of the first A, or in the B section, that venture to more remote chords, so long as the B section ends on the dominant.

The extension of a two-reprise form usually gives greater emphasis to the dominant phrase or phrases of the B section, even temporarily modulating to the dominant as a key center, this requires a modulation back to the tonic, another extension, to complete the form. But as long as we hear the dominant as a chord of tension, the key will be unchanged and phrases only seem to *hesitate on the dominant*, stopped short of resolution. A modulation has taken place when the dominant chord has been emphasized enough to seem a satisfactory point of resolution. The recognition of a modulation is an important technical point in musical analysis, but for a listener the difference between a phrase that ends on the dominant or a section that modulates to the key of the dominant is a question of emphasis. The two-reprise form is essentially small in scale, but composers love to stretch it out, teasing the listener with suggestions of a larger perspective. Often

a composer, when enlarging a simple form, may decorate as well as extend the phrases or increase their number, even though the basic structure remains that of the two-reprise form. If you understand and hear the simple form easily, you will be better equipped to recognize it in an inflated version.

The Classic Rondo

This form developed from a simple plan used in French music of the late 17th and early 18th centuries, called *rondeau*. The large scale plan of a *rondeau* uses an easily recognizable musical segment, A, alternating with contrasting sections: A B A C A . . . A. The "A" section is usually a complete two-reprise form based on a dance rhythm. Sections contrasting to the *rondeau* and to one another, called *couplets*, offer different melodic material and harmonies but usually are constructed in a two-reprise form just like that of the *rondeau*. An economical notation was devised for the form, instructing the performers to return to the *rondeau* after each couplet. The *rondeau* defines the tonic key and the first *couplet* is in the key of the dominant (V). Subsequent *couplets* (C, D, etc.) are in keys somewhat more distantly related to the tonic. The consequence of this early version of the rondo is that the *rondeau* will be repeated exactly, except there may not be interior repeats of both reprises after the *rondeau* is first played. The *couplets* are similar in structure and length to the *rondeau* itself, and the *rondeau* will begin and end the form of the piece.

The rondo of the later 18th century is closely related to the earlier and simpler form. The changes from the *rondeau* are most notable in the contrasting sections that replaced the *couplets*, now called episodes. They have become a much stronger contrast to the rondo theme in texture, dynamics, melodic material, and formal design. Usually an episode will begin with a transition, therefore with irregular phrases and surprising harmonies. The episode may then present a new theme, or something catchy and recognizable to contrast to the rondo, before returning to the rondo itself. However, the transition and modulation may not evolve into a structured new theme, but (like a development section of the sonata-allegro form) may continue to wander and surprise the listener until harmonic tension builds to effect a return to the tonic key and the rondo theme. In both cases, the episode will present a contrast of key and the rondo will return in the tonic key.

In the late 18th century, the usual rondo theme is still a two-reprise form in a popular dance rhythm. When this repeats, especially in the middle of the form, it may be shortened or varied, but the final return is usually more complete than the versions heard in the middle of the piece. The rondo theme identifies a return to the home key whether it is heard completely or not. When the theme is a particularly catchy one, a shortened return tantalizes the listener for the whole melody. Haydn loves to play with his audience this way, and he is also willing to prolong the

harmonic tension leading to the return, in effect holding the candy in front of you, yet keeping you from tasting it, until he finally gives in, resolves his dominant chord, and swings into the rondo. But then, he might just take it away from just as you think you've settled into a good chunk of happy listening...

The rondo became a very popular compositional form, in two senses: It used catchy, simple tunes and dance rhythms, and it had a wide vogue and appeal to audiences. It is found as the last movement of sonatas, symphonies, serenades, and (most important of all) concertos.

The Classic Sonata-Allegro Form

The sonata-allegro form in the late 18th and early 19th century is based on harmonic structure, like the two-reprise form, but on a much larger scale. The *exposition* of the sonata form contrasts the home key, the tonic, and its melodies and textures with the key area of the dominant and its contrasting melodies and textures. Usually the exposition is repeated. The *development* section that follows employs surprising relationships of harmonies, fragments of melodies from the exposition, changing textures, and sequences that surprise and engage the fantasy of the listener, and then it ends with a strong cadence to return to the tonic home key of the sonata. The next section, the *recapitulation* restates the melodies and textures of the exposition, cleverly remaining in the tonic key throughout, thus resolving the harmonic tension between tonic and dominant keys. This represents the structure of the major-key sonata-allegro form that in the minor mode becomes a contrast between the tonic minor and its relative major, an interval of a third higher. Harmonic tension in the minor-mode sonata is subordinated to the contrast between the dramatic minor mode with the sunny major mode, a difference that is easier to distinguish on first listening to the form.

When music is firmly lodged in a key, musical phrases have a degree of regularity. If the music seems to ask a question, an answer follows, phrases may rhyme with one another, motives evolve into cohesive phrases. The listener is engaged with a sense of logical order. *Modulation*, the process of changing from one key to another, makes a transition between sections of settled and logical order. In such transitions motives are haphazardly related, textures change rapidly and surprises replace predictable regularity.

The exposition, development and recapitulation of the form all have subdivisions detailing smaller functions of the architecture. Prior to the exposition there may be an *introduction*, easily recognized by being in a slow tempo. Its function is to create a harmonic tension that when resolved emphatically establishes the tonality of the work. Between the end of the first key area, the tonic, in the exposition and the beginning of the second key, the dominant, there is a *transition*, a

section of unsettled tonality, irregular phrases and sudden shifts. At the end of the second key area, there is often a *codetta*, basically repeated cadences to emphasize the key of the dominant.

The development section will have subdivisions, changes of key, shifts of texture, and reassembled fragments of melodies heard earlier. Ideally it is unpredictable and fantastical, a fleeting, mysterious, jumble of bits and pieces that you have already heard. It will end with strong harmonic tension, prolonged and tantalizing, before resolving to the tonic key that brings the recapitulation.

The recapitulation repeats the beginning of the exposition, but the *retransition* that follows tricks the listener by using the unsettled tonality and irregular phrases of the *transition* to return to the key of the tonic in which the melodies and textures of the second key, the dominant, were originally heard. Finally, the form may end with a *coda*. It emphatically proclaims the tonic key with cadences, but may at first seem to be a diversion with new ideas, sometimes quite delightful, which gracefully and fairly quickly return to business, closing the doors and settling down to the last cadence.

If we are listening to a composition in the key of C major, the C chord is the tonic (I), and the G chord (at the interval of a fifth higher, symbolized by V) is the dominant. After the first key area in which V predictably resolves to I, the transition brings chordal resolutions that end on V, which is now equipped with a dominant chord of its own, D major, the V of V, as it were, on its way to G. The key of the music is then said to be G major, and a modulation has taken place. If the *tonic key* C can represent repose to our ears, as the tonic *chord* did, the *dominant key* G may assume the tension of the dominant chord. But this happens during a generous time span, not immediately, so the listener must keep the pitch of both tonic and dominant chords in mind.