Preface

> I have never been certain whether the moral of the Icarus story should only be – as is generally accepted – "don't try to fly too high," or whether it might also be thought of as: "forget the wax and feathers, and do a better job on the wings."
>
> *acceptance speech for the Directors Guild of America D.W. Griffith Lifetime Achievement Award* (1997)
> STANLEY KUBRICK

This book, *Advanced Music for Beginners,* will teach you how to make music effectively and will give you tools that automate most of the work for you.

Advanced Music

By "advanced" I mean that the goal is to make music that is extremely beautiful, refined, subtle, varied, and sophisticated. The purpose of applying science, mathematics, and technology to music is to achieve these qualities to a degree impossible by conventional means. Yes, our project is to do better than the old masters, the great craftsmen, and the millions of highly-trained practitioners who have devoted their lives to the art.

By "music," I mean, for goodness' sake, that it will sound like music, and not like some boring noise, or artless passion, or pointless instrumental technique, or dull nature recording, or audification of meaningless data. It will please the ear and interest the mind.

How to Be a Beginner

By "for beginners" I mean that I will be explaining how music works from first principles, not that you have to be a beginner. To understand this book, musical experience and training will hurt you a little and help you a little, but not significantly.

You don't need to be a performer (or to have instruments other than a computer) – I've included software that can do the performance and synthesis for you. Manual dexterity is not, and should not be necessary – music should not depend on sleight-of-hand like some cheap magic trick. Don't worry about the performance sounding mechanical – the included software has been carefully designed to play expressively, like good human performers. Now you've got a band of your own.

This is the point at which people protest that they don't understand anything because a computer is involved. You don't need to use the software to understand the book, and you don't need any special technical aptitude to use the software. You don't need to do any programming.

To use the software, you will need to handle written scores that are in a text format, but they contain the same information as traditional notation and are as straightforward as can be. Most or all of the actual score writing will be done by other programs that you will control with graphical user interfaces. Technically, you will only be doing two things in the whole operation:

1. You will conceive of the forms, textures, flavors, and colors of the music and then tell the composing programs what you want.
2. You will assemble their output in a text editor and press a button to render the score as a sound file.

Obviously, the first operation is creative and exciting, while the second is mundane but undemanding and necessary (and lets you adjust fine points of the performance, if you wish). I'll talk more about the overall process, giving examples and suggestions later. The book exists mainly to acquaint you with the creative, exciting part of the operation – the knowledge and ideas by which composers do their work – and only then, secondarily but necessarily, with the quotidian details of using the software. The programming details are relegated to the appendices as they do not form the main argument of the book.

I'll digress a little here. My heart is with engineers. This book and system isn't specifically for engineers, but I think they would make great composers, and perhaps don't know it. In engineering, one gets a set of constraints and criteria and then creates a structure which satisfies these while attempting to maximize or minimize certain qualities of the structure. This is exactly the situation composers are in when confronted with the problem of creating music, except that the constraints, criteria, and optimizations are generally their free choices, not dictated by practical concerns like budgets. I'd like to introduce the idea of musical composition as an engineering discipline where:

1. Creative choices are made.
2. The implications of those choices are worked out, resulting in music.

This book is advice on making the creative choices wisely. Its associated software automates the task of working out the implications.

This is the point where an author is supposed to promise that there will be no mathematics in the book (presumably because his target audience is *no fun*). I promise this much: that you will not need to know any arithmetic to use the software. I won't ask you to add or subtract two numbers. You will need to be able to count (probably to twelve or so). It would be beneficial if you had some sense of when one number is larger than another, but this isn't required – for frequently used parameters I'll provide sliders. If you had some kind of notion of proportion – if you sensed that 6 was about twice 3 – then you would be ahead of your peers, but this much mathematical savvy is not required. I wrote the software to do the math for you. So, counting is all you need to *use* it.

To *understand* the software is a different matter, but I don't expect many people to attempt this. That is why I have placed the descriptions of the algorithms in appendices. It's my hope that, apart from curiosity, some people will want to delve into these matters and improve upon my software (being sure to send me the results!). To understand these appendices, at minimum, you would need to be comfortable with algebra, equations, and procedural computer programming.

I am also going to put a little mathematics in the body of the book, for which I refuse to apologize. It will help to illustrate some concepts. Some brave readers will understand it and, for the rest, I will draw pictures.

You don't need to know any jargon, especially musical jargon. Some is good to know, and if anything is important, I'll explain it. You won't, for example, need to learn the arcane terminology of all the chords in the zoo. We'll bypass most trivia like that and focus on a higher level of information, which is where composers' minds should be anyway.

Above all, I want you to understand that composing music is not a mystery. It is not something unknowable that goes on in the black box of the unconscious mind. It consists of the creation, manipulation, and fitting together of knowable, well-defined mental objects (things like rhythms, melodies, harmonies, scales, modes, meters, textures, and forms). It is not difficult to do. It is interesting to understand. It is profitable to apply conscious thought to music. Music deserves this attention. The results will be deeply rewarding to you.

How Not to Begin

In order to acquaint you with what composition is like, I need to dispel several popular misconceptions about how it is done. These are bad ideas, dead ends, and drains on your time and energy that I must protect you from. There are countless ways to compose music. I'm not saying that any are *invalid* – it is possible to arrive at good results from outlandish means. I'm saying that most methods are *unfruitful* – most of the results are very poor and it can take an extremely long time to produce anything acceptable with them. If you are tempted to protest that this or that example among your favorites was the result of something I call a bad idea, consider the millions of musicians putting millions of hours into them without producing anything like your shining example. (Also consider that you may have too much respect for your favorites.) The one fault that all these bad ideas share reveals a key theme of this book: your music will be successful to the extent that it intelligently addresses the human auditory system (by which we hear sound and cognize it as music). Unfruitful methods of music making fail to do this. Bad ideas often sound like good ideas. Beware of the following.

Improvisation is bad.

Deciding what notes to play and deciding how to play them are two very different things. This is why the traditional division of labor between composers and performers is so successful. The knowledge, thoughts, and ideas that go into producing a score (with all of its complex interrelations) are nothing like those that go into deciding moment-to-moment how to shape the notes, impart timbral shadings to them, trace the line from one to another, group them into phrases by subtle deviations, and so on. People who attempt to do both at once neglect both.

It's a tragedy that most people think that composition is nothing more than improvisation. They think this because they have no idea that there is anything *to* composition – they are unaware that there is knowledge behind it, and that there are methods of conscious thought designed to lead to good results. The standard practice has become to noodle about for a long time without thinking, and to stop and record the occasional bits that appeal. This is like setting a million chimpanzees to typing and waiting for Hamlet to emerge. There is an intense and persistent desire behind this, one that we will often encounter: the desire to avoid all mental effort, to not think about anything, to rely on some mysterious genius in the subconscious or elsewhere to solve the problems that the conscious mind has not learned to solve. This merely pushes the operation into lower, less interesting parts of the nervous system. Muscle memory dominates improvisation, reinforces familiar habits, deepens ruts, and leaves us with nothing original. The results are weak, bland, and uncreative.

Imagining the sound is bad.
The second most common misconception is that composers imagine the piece they want in their mind's ear and then simply write it down. The ruts of auditory memory being not much shallower than those of muscle memory, one will imagine what one already knows, or something not far from it. This is how plagiarism suits begin. Worse, it comes at the problem from the wrong direction: beginning with the surface level of the music and not the functional structures beneath the surface. This comes from an ignorance that there *is* anything beneath the surface. On the other hand, if you understand how music works, the underlying structures, and the techniques for creating and manipulating them, it becomes a fascinating game. Composing at this deeper level not just easier and more fun, it is far more *effective.* You will produce more novel work from the deep level than you could possibly imagine at the surface level, and, if it is based on valid theoretical principles, it is guaranteed to be aesthetically and artistically pleasing. Rather than imagining a familiar musical surface and enduring the drudgery of writing it down, you play the game of assembling theoretical structures, (let some friendly software work them out to their conclusions), and then enjoy the pleasant thrill of hearing what your ideas sound like.

Imagining the sound is not a bad practice in tiny amounts applied occasionally ("This needs a couple of notes right here."), but is a bad habit, and a bad way to approach anything larger. Again, it is the desire to get something for nothing, to pull the answers out of nowhere. It's an attempt to improvise without actually playing. Bypassing the muscle memory is a good thing, but imagining the sound doesn't go far enough and is still wrong-headed. Imagining the deep structure is far more effective than imagining the surface.

Noise is bad.

"Noise" used to mean *undesirable* sound, but, within living memory, it has come to be embraced unironically as music. How did this happen? The answers are many, complex, and extramusical, but here is a sketch. In most cultures, including the West up to the late 18th century, no one wanted to be a professional musician. It was a low-status position, given to orphans and blind people as charity. With Romanticism, this changed and the cult of the artist began; the social position of artists shifted from lowly servants to a sort of secular priesthood. Claiming to be an artist became an *upward* social move for most people. At first, this still required actual artistic skill, but, beginning in the 19th century, a wave of amateurism swept through the music world. In Germany, upper- and middle-class musical amateurs were offered the first courses in composition. By the 20th century, in popular music and the popular imagination, composition and performance had become completely conflated (leading to the fallacy that improvisation is composition). Gradually, the skill required to claim the high-status title "musician" declined. Humans are, after all, social animals, and it is polite to applaud at the end of the performance no matter what. When it is possible to claim a higher social status for oneself without cultivating any special knowledge, it is a short step for people to become ignorant that there *is* any special knowledge to be cultivated.

But what is wrong with noise itself? It's sound that is rejected by the human auditory system at a very deep level. Evolution has made this system urgently concerned with clarifying a mental picture of one's surroundings. How it does this and how musicians have hijacked the system for the sake of art is the fundamental theme of music theory and of this book. "Noise" is what we call auditory experiences that the system cannot resolve into useful information. For good evolutionary reasons, this makes the brain very nervous – the noise may well indicate a serious threat – and the listener's experience is as viscerally unpleasant as any auditory experience can be. Noise induces anxiety. This is why music has until recently been the practice of cultivating the opposite of noise. But, since it requires no skill or knowledge to produce noise, it is easy to elevate one's social position by claiming to be a "noise artist." The benefit is entirely to the producer, and the detriment entirely to the audience.

The artistic prospects of noise are worse than this suggests. Because noise (including dissonance) is the sensation of *missing* information, the sensation does not carry any information. All such sensations are alike. They are the brain telling us "does not compute – nothing to see here." There is no meaningful difference between one noise and another. Noise is a cold, grey wall of bland monotony. Noise is anxiety and boredom at the same time.

I offer an analogy. Suppose I told you I were a painter, but I don't want to bother to learn how to draw. My strategy is to add a lot of grey paint to each of my colors so that the audience cannot detect my inability to draw. Enough of my friends do this to form a cozy social "scene." If the scene becomes unsupportably large, we can split it into factions by dressing differently and affecting different attitudes. We sell some paintings and hold our heads high. In this way, noise is socially useful, physiologically unpleasant, and artistically barren.

Distortion is bad.

This is more of an orchestration fault than a compositional dead end, but it's a corollary to the previous warning, and leads to a lot of obstacles in composition. As we'll see later, the ear likes to be filled up with a lot of information. There are many interesting ways of doing this – attention to tuning, the construction of complex chords, and so on – that do not lead to the problems of dissonance and noise. Most people dismiss these methods because they require some knowledge and conscious thought. (E.g., in order to understand tuning, you need to learn how to multiply two fractions. Few musicians will accept this burden.) Distortion is a cheap way of filling the ear without the need for knowledge or mental effort. It quickly increases the spectral complexity of sounds without the musician having to think about anything. It may be reasonable to use a little occasionally, but large swathes of popular music rely on this special effect as all or most of their system of harmony. As soon as there is a little too much distortion, sounds get in each other's way, dissonance mushrooms, and we are back at the cold, grey wall of boring old noise. Think of distortion as a greasy, clogged sink, and you will have the proper attitude towards it. But then, how can we avoid the mental burden of creating spectrally interesting music? In the software included with this book you'll find programs that solve these problems better than we could with pencil, paper, and unlimited time. Dismiss greyness and welcome color. Dismiss distortion and welcome harmony.

Drone music is bad.
Drones – one or more long, held tones – have a legitimate place in music. In many traditions (e.g., in the Indian, Persian, Arab, and Turkish worlds), they are used as a background, helping melodies to imply a harmony and to illustrate melodic modes. In these traditions though, the drone plays a supporting role to other elements, mainly to melody. What I mean by drone music is something completely different, something akin to "noise music," whole pieces comprised of single notes or chords, varying, if at all, mainly in timbre. As we shall see, this doesn't present the audience with enough information to be really interesting. Its only advantage is that it requires minimal cognitive effort to produce. Don't abuse drones this way. The real adventure begins when you write that second note.

"Experimental" music is bad.
In the world at large, "experimental" suggests the scientific method – an hypothesis is formed, an experiment is designed, data is collected and analyzed, and a conclusion is drawn. Nothing like that is going on behind concerts and records labeled with the word. It is used as a shield against criticism, an insurance that everyone must applaud no matter how bad the show, a claim to secure the social benefits of artisthood no matter how bad the art. The label "experiment" invariably means that the product is unpleasant, uninteresting, and above criticism. This is not to say that legitimate experiment does not occur in music (as we shall see) or that experimentation in a loose sense should be absent from your music-making. You should try many techniques and structures to see which of them you like, but there is no point in doing this while ignoring *what is already known* and, it would be rude to present your audience with these trails without making sure it's good music first. Don't pretend to be an artist by dumping garbage on the stage and demanding applause. You may be a social success this way, but your music will be an artistic failure. Inform your experiments with existing theory and use what you learn from them to craft pieces worthy of presenting to the public.

Technical sophistry is bad.
This is something that goes on in the rarefied world of computer music. Many times, I've read the manual of some new system and stopped a hundred pages in to check whether I'd gotten the right document because

not one musical word had been used. It's obvious what has happened in cases like this. Some technical person, a computer scientist perhaps, has a few favorite mathematical or computing techniques. He decides to make a weekend project of figuring out how music works and creating a program or language for it. A few minutes of reflection reveals the secret of the old masters: they must have been using the technique of which he is already an expert. He writes something that generates data using his technique and looks for a way to map it onto something musical. He thinks MIDI is musical, so he uses that. It lives forever online, with a few people trying to use it seriously. No good music results. It's the same old story: the inventor, just like a naïve teenager with a guitar, is not familiar enough with musical tradition to suspect that there are reasons behind it. He is unaware that the subject of music has any content. The results are bad because the program does not address the human auditory system intelligently. All appropriate mathematical and computing techniques should be put in the service of music, but if their application is not musical, if they do not take into account the human cognition of music, they are useless.

Academic sophistry is bad.

Under the headline, "Call for Works" you will find paragraphs of abstract words, vaguely connected. They may speak of the redefinition of concepts and the blurring of the boundaries of sound, image, data, society, architecture, with environmental and urban factors, of deconstruction and contextualization, of problem posing and fertile grounds for solutions. The eye slides straight across paragraphs like this, and that is fine – they have no real content. The sophistry is drawn from the humanities and fine arts departments, tinged with business English, but the same fault – lack of musicality – is behind it. It exists to give people a niche to fit in, not to think about or to make real music. If you know how to make music, you won't have to resort to nonsense.

Blind negation of successful traditions is bad.

Suppose you were so ignorant of music history and organology that you thought the instruments of the Western late-Romantic orchestra were handed down by God at the beginning of time. Suppose you were so ignorant of musical acoustics that you didn't know new instruments were possible, or that superior and more varied software instruments were possible. Suppose you were so ignorant of music theory that you thought Western art music had exhausted its resources by the early twentieth century. If you were such a perfect ignoramus, then it would seem to you that the few possibilities of doing something new in music all center on doing the opposite of what is known to work, for example, misusing the instruments to make comical sounds, or avoiding pleasing sonorities and progressions. If you were such a fool, you would think serialism and extended playing techniques are cool, and you would think you were very clever. Don't be an ignorant fool. If you wish to negate something, do it in an informed way, by making something better.

Imitation is bad.

I'm not saying that it's bad to set a folk tune or to use an occasional quotation. I'm certainly not saying it's bad to use an existing idea, like polyrhythm, or a traditional mode. What I want you to avoid is the very large world where *everything* is imitation. Most people don't create; they copy a style or, at most, mix two. This is what you must do if you don't know how your craft works and you aren't able to invent anything of your own. This is why popular music is the most conservative thing in the world. In particular, for the past few generations it's been dominated by rock'n'roll – a few children's records from the mid-1950s that have been imitated and re-imitated millions of times. An utter poverty of imagination, it is the resort of the most uncreative people in the world. It isn't the only glacier in the music world, but it is the largest. Don't be dull. Learn how to create so that you don't have to copy. The less respect you have for the establishment, any establishment, the better.

Narrowness is bad.

A teacher related to me a story we have all experienced many times. On making a new acquaintance, we ask him what kind of music he likes. He replies, talking for a bit about the mind-boggling breadth of his musical taste and then gives examples that he thinks are far apart. His reply sounds to us something like, "I like everything from Metallic-*A* to Metallic-*O*." – pathetically parochial. This is a result of the culture of pervasive imitation. A world of millions of records that sound alike seems big to the people stuck in it, and they magnify its tiny variety by making hair-splitting distinctions. Marketing and the media industry locks most people in this tiny prison from birth. Break out and become cultured. Listening widely is good, but far more than this, it is important to learn all the structures and techniques that are viable, not just from exotic and historical traditions, but from first principles. Learn the full breadth of what is possible and explore it in your own work. This book is a means to that end.

Cultural appropriation is good.
Terrence wrote in the 2nd century BC, "I count nothing human alien to me.", expressing the good sense of educated people to consider all human culture their birthright. In the 18th century invention of intellectual property, capitalism found the weapon that most injures culture. Lately, this has been taken to an extreme, possibly the most right-wing extreme of any idea in history: the notion of "cultural appropriation." The idea is to divide humanity into the smallest possible compartments by every possible demographic border, and forbid the exchange of ideas between them – it would be theft, according to the hypercapitalist world view. Be like Terrence and defy such evil.

Culture is not intellectual property rights. It's our common heritage and business. Like mathematics, useful musical elements are precious discoveries that should be freely exchanged. We (those of us with human auditory systems) are all in this together. Existing musical traditions have learned to address themselves to the human auditory system intelligently. You'd be a fool not to sit at their feet and learn all you can from them, no matter who your parents were. You don't need a genetic test to use a musical idea. All men are brothers. Be cosmopolitan. Use good ideas no matter where they come from. If you wish to call me out on this point, don't use the Latin alphabet unless you are Roman, or ASCII unless you are American.

Mapping data is bad.
For decades, it has been popular to take some data from whatever source (the DNA of the AIDS virus, NASA telemetry, the patterns of pigeons sitting on a telephone line, etc.) and clumsily map it to musical data (the people who do this invariably think MIDI is musical, so they use that). We have seen this already under "technical sophistry." These, however,

generate a lot of public interest at the level of clickbait. The headline reads: "Scientists are making music from *XYZ*". Millions click, listen for half a minute, and then, next month, click again, forgetting this has been done before. It has not developed into a musical genre because, ignorant of music cognition, the results are bound to be boring. They are based on a fundamental ignorance that musical material has a structure of its own and is best generated by musical methods for musical reasons.

If you don't know this, you will assume either that musical data is essentially random so that any source of data will do, or that musicality is some unknowable essence that also occurs in natural phenomena. In the first case, you might as well use a generic pseudorandom source, but in either case, for the sake of popularity, you'll choose a data source of public interest. An insensitive, ham-fisted, unmusical mapping will complete the project and earn you your pageviews. The sonification of data is a legitimate field, but it's much trickier than the visualization of data. The mapping of data to sound is not at all simple, particularly because the relationship of sound to musical experience is so complex, as we shall see.

The software in this book is overwhelmingly deterministic, because there are almost always good musical reasons for things to happen one way or another – there is little room for randomness in serious composition Concentrate on interesting, creative ideas, not on mapping data.

Gear is bad.

By "gear," I mean musical instruments and associated equipment, mostly electronic hardware. Gear is designed primarily for making money, not music. That is, it is designed to take your money away from you. This is supported by whole cultures of gear magazines, retail stores, and associated fetishism that conspires to make gear appear central to music making. There are two ways in which gear is ineffective: it is inferior to software and it demands tactile input.

For decades now, programmable computers have been able to perform the same services as gear equally well, with infinitely more flexibility. Software can handle complexity much more easily than hardware can, so its output can be much more sophisticated. If you learn how to make musical programs (you needn't to use existing ones productively), you will soon get into the habit of looking at a piece of gear and estimating how many lines of code would do the same thing for free. Often the answer is on the order of ten. (I recommend *Csound* for this, but there are many other systems.) Save your money. When you want a new synthesizer, open a text file and write one (or use one that someone else has already written).

In order to hook you, gear manufacturers need a bottleneck, something to hook your physical presence to the world of hardware. This is almost always done by insisting that your physical gestures must be part of the music-making process. This is the fiction that makes the world of gear commerce plausible. Consider the absurdity of this premise: every other physical motion of musical instruments and spaces can be automated by expensive hardware, but the relatively simple motions of your hand, foot, and breath cannot. It's a confidence trick that plays on your vanity (and propagates the harmful conflation of composition and performance). Free software can automate your gestures just as easily as it can automate the behavior of musical instruments and spaces. Let go of your vanity. Accept that your-low level input (musical performance) can be automated and concentrate on your high-level input (directing the composition process). Reject expensive hardware and embrace free software.

This book includes a software orchestra which has several automated performers with different personalities (and, if you want to dig a little, you can create more of them). If you wish, you can control every detail of their performance, but you won't need to. They already know how to make the right decisions to play expressively, with soul and feeling. Let go and let them play. I don't claim that they are better than all human performers, or that my automatic performance system is the ultimate. I hope and expect others to create better systems someday. But, the system I'm giving you is good enough for professional work. You will find that working at the high level of meta-composition is much more *fun* than working at the low level of physical performance. Try it and see.

Real-time operation is bad.

In "real-time" operation, the surface level of the sound (the final product) is produced by the machinery as quickly as its own time progresses. This allows no more time than the duration of the music for all of its computation to be completed. Systems of this kind are usually centered on a user's physical input, so that the auditory feedback is immediate. We must distinguish between three contexts for real-time operation:

Pre-performance: sounds are tested before the music-making begins.
Performance-only: the notes to be played have been decided beforehand, and performers are deciding how to play them on the spot.
Compositional: either machines or humans or both are deciding what notes to play on the spot.

In pre-performance, real-time operation can be a handy thing if one wishes to try several possible details quickly, e.g., synthesis parameters or the swing of a particular phrase. The alternative is to mull over parameter changes, re-render, re-eveluate the results, and repeat. In practice, I have not found this too burdensome, in fact, the extra bit of time to think can be helpful. I would not call real-time operation bad in this context.

The performance-only context is rarely used in computer music but is the usual one for Western art music. Computer music merely replaces physical instruments with synthesizers. There's nothing particularly wrong with this, but if one uses computers for instruments, why not use them as performers also? Our orchestra does this, replacing the motion of both the instruments and performers with computation.

The compositional context is common in computer music. It implies composition-by-improvisation (details as coarse as notes being decided at the spur of the moment, whether by human or machine), the drawbacks of which have already been discussed. Real-time computer music introduces new disadvantages of its own.

Real-time composition has two fatal flaws: limitations of information and time. In real-time, future information is not available. A system operating in real-time cannot know the future inputs, and so has only half the data. This cripples its decision-making. Also, real-time composition forces decisions to be made in milliseconds, severely limiting the amount of thought or computation that can be applied to them. This inflates gesture fetishism and comforts people who are afraid to apply conscious thought to music – since it is practically impossible at that speed, it can't be asked of them – but it is disastrous for the quality of composition.

Is there some sense in which immediate feedback might be useful? I offer this analogy. An adult engineer designs a real bridge by applying knowledge and careful planning. A child designs a bridge improvisationally with blocks, using instant feedback – he can feel when the structure is about to topple and, with quick reflexes, he might save it. Compare the resulting objects. One can carry real traffic and can stand for centuries.

The other may be sentimentally pleasing to its maker and his parents. The engineer doesn't need immediate feedback because he knows what he is doing. I want to impress upon you that knowing what you are doing at a conscious level isn't just more effective. It is more *fun.*

Midi is bad.

Midi was invented to sell gear. It is a system for controlling synthesizers – those of 1981. It was not the first proposed standard for synthesizer control. It is a cheaper simplification of earlier systems that the industry agreed upon in order to avoid competition. People have come to think of it as a way to record scores or the details of musical performances only because it has long been pressed into service for these functions. The only tradition that informed its design was twelve-tone equal temperament common practice era keyboard music, and even from this very narrow perspective it records scores and performances poorly. It has all the failings of an unmusical engineering solution as described under "technical sophistry." Even simple things like timing, pitch, and control signals are coarse and difficult to control. Unusual rhythms or tunings are hopelessly awkward in it. For many years, prospective improvements to midi were the most popular research topic for Computer Music masters theses. This is because every aspect of it has so much room for improvement. It is not suitable for serious artistic work. Avoid it.

Instead of midi, the software in this book uses a simple, powerful macro language within the *Csound* score language. It is straightforward and human-readable. With a few parameters, you can control the timing, pitch, loudness, and phrasing of notes in much more sophisticated ways than midi provides for (e.g., meter, swing, mode, and harmony can be kept independent of the arrangements). With a few more parameters, all the details of performance (control signals), orchestration (synthesizer features), and even sound location and motion are available. You do not need to dig deep, though. Even if you provide only the basic information from a traditional score, the defaults and the virtual performers will produce excellent output. You will want to learn this format to work with the system, but the more advanced programs will write chunks of the score for you and allow you to think mainly about higher levels of composition – texture, color, flavor, form, structure, etc. – that is the point of all this automation, not technical details.

Musique concrète *and sampling are bad.*

These words describe four activities, three of which are bad:

1. "Sampling" can mean the basic technique of digital recording – measuring the pressure incident on a microphone thousands of times per second and recording these "samples" digitally. There is nothing wrong with this.
2. "Sampling" can also mean recording instruments playing individual notes. These "samples" can be assembled into music later. This can work, but usually it's bad or less than ideal, for technical reasons: large libraries of recorded notes are needed and they are not flexible – it is difficult or impossible to control all their parameters, to add performance nuances, to make them fit into a natural line, and so on. Ultimately, this fails because it is a substitute for synthesizing sounds intelligently. Synthesis can allow all of the control of natural sound and more, but to get good results with it one must understand acoustics, mathematics, and perception. The sampling technique is a shortcut that avoids all that thinking. It is useful in some very particular contexts (mimicking traditional fixed-pitch instruments with few expressive parameters or interacting parts, like pipe organs and some percussion), but usually it is inadequate. Most of the time, this kind of sampling is bad.

The remaining two activities, sometimes called "painting with sound," attempt to make music by assembling segments of pre-recorded sound into a finished piece. Both of these are bad:

3. "Sampling" usually implies that the recorded segments ("samples") are of real music made by someone else (someone who knows how).
4. "*Musique concrète*" usually implies that the recorded segments are of natural, but non-musical sounds (i.e., not of musical instruments or singing voices).

These last two techniques have one good thing going for them: because the ear is adapted to hear natural sounds (including those of musical instruments), they are especially salient to listeners, as long as they haven't been modified very much. But, both miss a fundamental point: music is not simply a sound. Between acoustic waves entering the ear and the mind having an experience based on them, many layers of processing occur. Broadly, we can consider them at two levels: low and high. Low-level processing gives us the experience of sound itself, resolved into a picture of something happening in our environment. This is practical, and not very artful. Art comes with higher modes of processing and more interesting experiences, of which there are at least two: speech and music. If the sound can be decoded into a known language, the listener can experience layers of linguistic and semantic meaning, and poetry and literature flow from this. If the sound can be decoded as music, the listener can have a musical experience, with its layers of sensation and meaning. If, however, the sound is not of a form that can be processed at one of these higher levels, the listener can't extract anything interesting from it. Someone "painting with sound" is very unlikely to create meaningful English sentences and equally unlikely to create music. They are probably ignorant that the musical level of cognition even exists.

Like language, music is a mode of listening that does not work with just any sound. The experience of *Musique concrète* that doesn't have a musical form (as is the usual case), is like that of trying to understand a story told in a foreign language – nothing gets through. Natural sounds might have some use as atmospheric background or as unpitched percussion, but these cases are marginal and run into all the difficulties of inflexibility described above (under sense 2).

"Sampling" in the third sense, assembling other people's music, has more chance of working, because the plundered elements probably have some musical value in themselves. But, the assemblage is usually only rudimentary layering and repetition, possibly with an simple treatment of the samples – time shifting (stretching or shrinking the duration) or pitch shifting (transposing) the samples, which is trivial compared with real composition. This is a technique used by people who don't want to learn how to make music, so they have to use other people's. At best, the result is like a clumsy collage executed by someone who hasn't cultivated the skill to draw, or to paint, or to design graphics of his own. This is what someone does if he wants to be a "composer" without taking a week or two to learn how real composition works, that is, how the human auditory system processes mere sound into musical experience, and how to create sounds that exploit that system. These are techniques for people who don't care about their audiences. Don't be one of them. Read on and learn how to compose.

If you have the urge to sample, or to imitate, or even to use your own habitual musical gestures, here is a better course of action to take:

1. Study a brief sample, a short figure. Listen to it carefully, look at the sheet music if available, and look at it in a sound editor.
2. Understand its structure, what makes it work. Find what it is about the sample that you like.
3. Do this to such a degree that you can describe it in words.
4. Now you own it. Now it's not a clip of someone else's music, or a one-off twitch of your own fingers. Now it's a tool in your toolbox, and you can deploy what you've learned from it, adapting it to whatever situation you are in, far more flexibly than merely changing its speed or pitch.

LoFi is bad.

"LoFi" stands for "low fidelity," low sound quality in the most basic technical sense – small frequency range, small dynamic range, background noise, distortion from analog equipment, and "glitches" from digital equipment. Good sound quality used to be difficult to achieve, requiring knowledge, skill, and money. Technology has progress so far that it has long been cheap and easy (this will become obvious soon, when the basics of digital sound are covered and advice on these aspects of production are given). In fact, technology is so good now that LoFi no longer requires any special effort to avoid but, like all bad things, it is actively pursued as a sort of affectation, and "LoFi" has become a genre.

There are at least five reasons behind this:

1. Fetishism for inferior technology.
2. Obsolete rebellion against the elitism of HiFi, an elitism that ceased to exist a generation ago.
3. Putting on a romantic show of wildness, or at least, of nonchalance.
4. Emphasizing the passion of the artist by stripping away all desirable qualities, like beauty and skill. If he plays so badly, if he hasn't planned the composition skillfully, if he hasn't bothered to learn his craft, if he doesn't care how he is recorded, then passion must be what is left. I will have more to say of passion later.
5. True ignorance of a few basic facts about recording and sound.

The problem is that bad sound quality is unpleasant for the audience, it distracts them from what music is present, and it severely restricts the options and resources available to the artist – subtlety, nuance, complexity, and richness can't reach the listener through the narrow channel of LoFi.

It is ugly, and ugly art is weak and ineffectual. LoFi is a strategy for musicians who don't care about the audience. Don't be one of them.

Excessive HiFi is bad.
It is possible to go too far in the other direction. The result is delivering wastefully large files as your finished product or, paradoxically, less than HiFi physical media. The cause is always ignorance about human hearing, and the science of sound reproduction. There are powerful forces invested in keeping you ignorant. Under the name "audiophile," there exists a whole gear industry with its own trade organizations and magazines devoted to selling absurdly overpriced equipment for home music consumption. They will sell you anything, as long as it is at a high price. They will sell you magic rocks. They will tell you anything in order to do this. To keep their industry going, they promulgate some of the most amazing pseudoscience the world has ever seen. Their strategy is to make audio technology seem much more complex than it is. Nonsense from this world has seeped into the world of music production.

Take some natural ignorance, the cultural influence of the audiophile world, add a few logical fallacies, and excessive HiFi will blossom. The fallacies include the idea that more expensive products must be better and appeals to your ear as the final arbiter of quality. Experienced musicians know not to trust their ear. Your ear fools you most of the time (that's part of its job). What you think you hear can very easily be suggested by circumstances (e.g., after about twenty minutes of mixing or tweaking, *any* change you make sounds like an improvement). This is why science relies on careful double-blind testing. Well-established investigations into the limits of human hearing and the efficacy of various formats has lead to some simple guidelines about reasonable specifications for production and delivery formats. These details, simple to understand, will be covered later. Don't be lead by ignorance or advertising. Your ears are not better than the other humans'. Don't waste time and money on excessive HiFi.

Analog technology is bad.
To a *naïve* understanding, analog technology implies theoretically infinite accuracy. In practice, of course, it is particularly prone to many types of noise, distortion, poor bandwidth, and limited dynamic range. In other words, it is inherently LoFi. Digital technology, on the other hand, is

accurate and clear. An analog signal can be digitally encoded and then, if the parameters of the encoding (sampling rate and bit depth) are sufficient, can be perfectly reconstructed as an analog signal to within an arbitrary tolerance. Yes, perfectly. This is the Nyquist–Shannon sampling theorem, proven and published in 1948, but which had been well understood by communication engineers for a generation before. Analog is championed only by people who cannot understand this.

Sometimes they praise sonic qualities of analog gear as having "warmth." This "warmth" is distortion and poor frequency range (I think "muddiness" is a better description). If you want this quality, it is easily achieved with digital filters acting on a digital signal. Also, there is a subculture of gear collectors who focus on analog synthesizers and praise them for their unpredictability (unreliability). This randomness is an important substitute for creativity among people who have none.

Physical formats are bad.

All are bulky petroleum products (excluding shellac, steel wire, and some kinds of wax) which require more petroleum products to be burnt in order to truck them around the world. The analog ones (e.g., vinyl phonograph records and cassette tape) are pathetically LoFi and quickly degrade with time and use. These are championed by people who don't care about either the environment or music. Ignore them. They are unimportant.

There is also a philosophical problem with physical media which must be pointed out: the real work, the thing recorded, a novel or a movie or a piece of music, is a bundle of information – words, images, sounds – being communicated from one intelligence to another. The lump of matter that carries it is an inert distraction, a heavy obstacle that gets between you and the communication. In this way, non-physical media (well, *nearly* non-physical computer files) are more authentic. Avoid physical media where you can. Oh, and one more thing: your fetishism for obsolete technology is squaresville, dad.

Identity is bad.

Most people come to know music through popular music. ("Popular music" doesn't mean music that a lot of people like; it's the category that isn't art or folk – commodified, mass-produced music.) Let's compare how that works with how music listening works in general.

People have no identity, but they desperately want to believe they have one in order to make themselves seem real. Popular culture, especially popular music, is a showroom of ready-made identities being sold for this purpose. Imagine a typical music consumer sampling recordings, concerts, and videos. The musical aspects of the products presented to him – the things we could talk about in a technical way – are all very similar and make no difference to him. What is significant is the hairstyle, attitude, and clothing of the performer – what slogan he shouts, what character he portrays on the stage. The vast majority of music fans don't actually like music. They like the lyrics, but not as poetry. They like it as a slogan, an identity badge they can take up. There are many subgenres, each an identity available for purchase. For example, if the consumer wants to be hip hop, he'll pretend confusion in a world where misspelling is creativity; if he wants to be country and western, he'll affect the air of a rustic simpleton; if he wants to be some variety of rock, he will punch his fist in the air and rebel against nothing in particular.

This whole industry works according to plan, efficiently keeping consumers happy. From our point of view, the trouble comes when one of them wants to take the playacting further and become a musician, aping his heroes. This involves some harmless dressing up and acting out. The tragedy begins when it is time to punch some helpless guitar. You cannot blame him for not knowing that there is anything "to" music. His society has taught him that music itself is a subject without content. When consumers, ordinary people, say the word "music," they never talk about musical structures and techniques. All they talk about are extra-musical associations (hairstyles and attitudes – identities) that are bundled together around the unimportant sonic object. They are ignorant that the structures and techniques even exist. At some point, they learn the bare minimum of theory – three or so chords and how to finger a scale. Pop theory is the most condensed over-simplification possible of the theory behind Western art music's common practice period. This is one of the world's great traditions, and it isn't bad, but, like all of them, it operates in relative ignorance of the others. The people punching the guitars are working under several layers of darkness. The product is easily predicted.

How does music listening work in general, outside the identity market? Roughly, it proceeds in three stages:

1. A human auditory system transforms acoustic (physical) events into perceptual objects – sounds.
2. *If* these data are in suitable forms and patterns, they are further processed by other neural circuitry into musical objects – experiences of rhythm, melody, harmony, and several other types.
3. For reasons that are probably arbitrary, idiosyncratic, and capricious, further mental processing projects aesthetic associations onto the musical objects – mood, emotion, and so on, including a judgment about whether the music is liked or not.

Don't despair: the suitability requirement of stage 2 is not at all narrow, and the unpredictability of stage 3 allows you a lot of flexibility. Notice how different natural listening is from the identity market.

When I say, "identity is bad" I don't mean that I want you to avoid presenting your music along with some extra-musical material, some theatrical personality. I am simply advising you of two things:

1. Recognize that the fictional, theatrical setting in which you present your music is completely independent of the music itself. They are two different compositions, two different creative activities, two different games.
2. Direct your attention to the italicized *"if"* at stage number 2. Make sure that your sounds are of a suitable form to be taken up by your listeners' higher processing. That way, you have a chance of causing positive aesthetic judgments in the end.

I'm here to advise you on creating music that can plausibly become the basis of aesthetic experiences. I leave it to you to discover whether it is possible to cause particular, predictable aesthetic experiences in stage 3 by using musical means alone or along with extra-musical suggestions. Just know that, whatever identity you choose to play at, music itself has a mechanism that is well worth understanding.

The identity market is for suckers and I don't want you to be a sucker, so there are three items in it I must warn you about especially: passion, rebellion, and respect.

Passion is bad.

Most human cultures in most periods of history have regarded passion as a childish, idiotic fault unworthy of wise adults, the most common cause of blunders, a fatal flaw, the mark of a fool. Then, in the 1980s, American capitalism decided that it would make employees work harder for less, and began to promote this species of foolishness as a virtue. Society adopted it wholesale, as a new religion of maximum effort, so that now, to most Westerners, or to Americans, at least, it sounds like heresy to say that passion is bad. Inspirational memes about loving your job quickly spread out of the workplace, hijacking the cult of the romantic artist. Now everything you do must be an obsession for it to be respectable. People admire nothing more than folly.

There is a small subgenre of documentary film that I enjoy. It reports on people who have allowed their stupid hobbies to ruin their lives. Some of the most delightfully pathetic are the air hockey guys. Men who live in mobile homes because they have lost their careers, families, and property due to this obsession raise their fists in the air and declare that it will all be worth it when their chosen arcade game is recognized as an Olympic sport. It is clear what is expected of the audience: we are supposed to say, "but you have to admire their passion." No, you do not. Feel free to laugh at passion. Feel free to ridicule it. Passion won't lead you anywhere good.

Music can be enthralling and absorb your whole life. Don't be a thrall. Don't waste your life on this one pleasant hobby. Pursue it, if you wish, but calmly, and coolly, along with whatever else interests you. You'll be a much better musician for it.

Sanctioned rebellion is bad.

There is only one marketing department. It dictates the protocols of mainstream conformism and, for the large segment of the population that wishes to rebel, a few alternative conformisms. Don't pick one (even if these options comprise all that you know). Don't be one of the people who are different, who are all different in the same way. Don't mistake any rebellion, old or new, as being genuine. All genuine rebellions are digested, repackaged, and commodified.

Punk is a famous example – soon after its emergence, beginning with frontline entrepreneurs like Malcolm McLaren, and ending with the anonymous executives who invented the "alternative" marketing category, it was quickly absorbed and harnessed by the system, by the identity market. Its quick mutation into many forms of post-punk did not even slow down the process. The system now assimilates authenticity so quickly that there is no point in attempting it.

Don't think that the phenomenon is limited to the world of pop culture. Folk music in the modern world is just as much an identity-commodity as any of its plastic folk revivals. Hipsters who learned fiddle from their baby-boomer grandpa are not really different from Americana hipsters who learned it from records and books.

In the world of art music, there are also phony rebellions. Most of the "contemporary classical" movement is based on ill-advised rebellions. Western art music, like all musical traditions, has moved at a slow pace through its development, never by progression to a higher, better form, but always by the random winds of fashion, this way and that, sometimes choosing interesting paths (e.g., the increasingly extended harmony of late romanticism), sometimes abandoning interesting paths (e.g., modality and counterpoint), and sometimes choosing dull paths (e.g., 12-tone equal temperament). Like all musical traditions, its history is a slow random walk through a forest of possibilities. Musicians with no imagination interpret this as "Whig history" – a rapid evolution to perfection (in their own generation, whenever they live) which could not have proceeded any other way.

The blow dealt to Culture by World War I lead to a spirit of justifiable rebellion best embodied by the Dadas and their anti-art, but which, like all rebellions, long outlived its reasons for existing. At the end of World War II, Western art music was dominated by men not clever enough to understand its music theory and unable to see the many interesting and unexplored paths that it might follow. They concluded that all logical evolution of their tradition had reached a perfect decadence, a dead end, and, since no further development was possible, it was their mission to tear it apart by doing the opposite of everything that had worked in the past.

An era of anti-music began. The stars of the Western art music tradition, like petulant children, sought to destroy their nursery. Pierre Boulez said, "It is not enough to deface the Mona Lisa because that does not kill the Mona Lisa. All the art of the past must be destroyed." Armed with a superficial knowledge of music theory, they did this by a number of perceptually invalid – and, therefore, ineffective – means: e.g., serialism, 12-tone music, atonality, aleatoric techniques, and so on. (We will avoid composing ineffective music by placing human perception at the heart of *our* system.) Since their music doesn't work on the cognitive level, it has no aesthetic value and remains irrelevant outside their cultural bubble.

However, the stories *about* them, their rebel image, their cultural bubble itself, had an aesthetic value, so it became another costume for sale in the identity market. This is a little different than the identities that allow you to play the rock'n'roll rebel in dive bars. It's a way for rich kids to play at being composers. For example, for the price of a cruise or a stay at a resort hotel, you can become one of the exciting young composers who participate in the Bang on a Can festival (the institution which rebels against the institution of the Tanglewood festival). If you pay up and respect the *clichés* of their genre, you can be one of the brave, young musicians rebelling against the classical music world of 1945–50. You will be called "adventurous" and "innovative" and you can enjoy the warm feeling of camaraderie in joining other rich young people in rebelling against constraints that had been forgotten before you were born.

All these rebellions or rebel stances, whether comically obsolete or not, are *sanctioned* rebellions. They do the opposite of what they purport to do: they bolster whatever they claim to rebel against. At minimum, they provide a foil which defines their enemy. Enthusiastic fools are willing to become whatever straw man their enemy complains about. For example, if conservative Christian elements claim that rock is ungodly, unimaginative identity shoppers will prove their point by forming satanic rock bands. Tithes increase as a result. At the worst, rebellions are absorbed by the system and used for its own ends – and so rock'n'roll becomes the sort of music that is played at rallies for authoritarian political candidates.

Rebellions inevitably work against themselves and ultimately become sanctioned and silly. If it is impossible to sustain genuine rebellion, what course should one take? Practice creativity that is entirely constructive and avoid references to past music. Theatrically: simply ignore rebellion. Avoid existing identities from the market – the shopkeeper has supplied each with an opposite to serve as its rebel. Don't shop there. Don't play the game at all. Perhaps you can't be so austere – you must have some extra-musical material, after all: titles, cover art, even a stage show. Embrace these with ambiguity and detachment – send mixed signals, change frequently, be detached from your identity like an actor. Be the Bowie you want to see in the world. Musically: use elements that interest you (you'll find many here) to build new genres by the dozen. Maybe we can overload the identity market's systems of classification and assimilation. Let's rebel by out-competing the system with a better product.

Respect is bad.

This is a book and software bundle to give individuals non-traditional means to make music on their own. Many readers will have come to it through punk. Nowadays, to many people, "punk" seems to mean "doing things badly." It's become yet another excuse for low effort and low expectations. This is not what it originally meant. Its true meaning is a studied disrespect for everything (people, institutions, ideas, etc.), resulting in a desire to *do it oneself*. The point of doing it oneself is not to do it badly – it is to do it better than what you are being sold. A parsimony of respect leads directly to striving for excellence in art.

The trouble with respect for musicians is that you will try to imitate them, you will come to think of their achievement as a limit you cannot surpass, and you will ascribe their success to some ineffable genius you lack rather than to some graspable working method you could learn. I don't mean that you should have no respect at all, but spend respect like a miser spending a gold coin. Always see the limitations of everything, and especially of those things you respect. Study the limitations of your heroes so that you can beat them at their own game.

Musicians have wasted themselves far too much on hero worship. Musicology has squandered its efforts in the same way, becoming a collection of biographies rather than a science of music. It's a natural impulse for us to revere, say, Bach, but it's counter-productive. It would be better for us to revere his methods, counterpoint techniques and such. Best of all would be for us to respect the human auditory system, the cognitive faculties of the listener that make Bach's counterpoint techniques valid. Respect the listener's cognition above all, respect methods some, respect structures a little, but do not respect musicians. No matter how much you admire your heroes, regard them all as bums who either got lucky or who knew some tricks you need to learn, nothing more.

If both respect and rebellion are bad – if both imitation and inverted imitation are bad – what course of action should we take? My advice is this: Learn all that you can about how humans process and respond to acoustic phenomena, distilled to first principles as much as possible. Use this knowledge to make excellent music of your own, without reference to any previously marketed. Repeat this frequently, with great variety. It will help if you don't like any existing thing – it is a good attitude to think everything is bad.

1

Introduction

> ...we should continually be striving to transform *every* art into a science: in the process, we advance the art.
>
> *Computer Programming as an Art* (1974)
> Donald Knuth

There is one guiding principle we must keep always before us: our music-making must address the human auditory system intelligently. Unless we keep the listener at the center of our thoughts, we cannot hope to make good music. By "human auditory system," I mean the whole mechanism of sensation and cognition that relates to our experience of music. Let us abandon all attitudes and schemes that do not place this foremost.

What do we know about the way humans experience music? We have only two sources of information:

1. The findings of modern science – Music Cognition and a few related disciplines like Perceptual Psychology and Musical Acoustics.
2. The knowledge inherent in existing musical traditions, particularly in their explicit theories, whether literate or oral.

Science is our surest information. It has the potential to be complete, general, and extremely accurate. With science, we can be certain. With science, we are on firm ground. Unfortunately, we haven't enough science. Music Cognition is a new field, and many aspects of musical sensation have yet to be investigated thoroughly. The bulk of our knowledge still comes from studying traditional music-making.

Established musical traditions (often associated with particular cultures, time periods, and genres) are descended from millennia of practice. Though not scientific, they are empirical. Each musical tradition is a worked-out solution to the problem of creating human musical experiences, and by studying and comparing many of them, we can get an idea of the underlying universals, the cognitive system that they all address.

There are many problems with this approach. All traditions are insular and idiosyncratic – they speak only of themselves and only in their own terms, as if they were the one natural way of music-making. Most are poorly documented. When we investigate them via literate traditions of music theory, we find that they leave much unsaid.

If we turn to Ethnomusicology for information (and it is often the only source) we find a field more concerned with extra-musical phenomena – the Anthropology that is going on in places where music happens to be playing. Most ethnomusicologists would be more interested in the brand of beer the audience is drinking than in how the musicians tune the scale. Also, as a point of ideology, Ethnomusicology has forbidden any comparison of traditions because this might result in someone having the notion that one tradition is better than another. (Imagine how Engineering would have developed if it were forbidden to suggest that some techniques are more effective than others.) So, Ethnomusicology and literate traditions of music theory have collected some data but left its analysis to us. Where science hasn't yet provided us with definitive models, we must make the most we can out of traditional knowledge, however we can obtain it.

What we want is a map of the territory, drawn from aerial photography. What we have is a collection of traveler's tales, each in its own language. Nevertheless, it is possible to collate the data we get from tradition, find the underlying structures it points to, and formulate first principles from which to build up a general music theory. For example, all traditions seem to have some sort of implicit or explicit notion of melodic mode, e.g., *echos, pathet, raga, maqam, dastgah,* or even folklorists' "tune-family." By examining many of these we can work out what sort of thing a mode is, a general format for their description, what constitutes a good one, methods for generating them and for using them to compose melodies, and finally, software for executing these methods automatically.

We do have some significant results from science, and we will have more in the future. As Music Cognition progresses, it will augment and then supersede what we have learned from tradition, which is to say that everything in this book is subject to change as our musical knowledge increases and becomes clearer.

We need a model for the way humans experience music. I offer you the best one I can, derived from science and tradition. Future science may

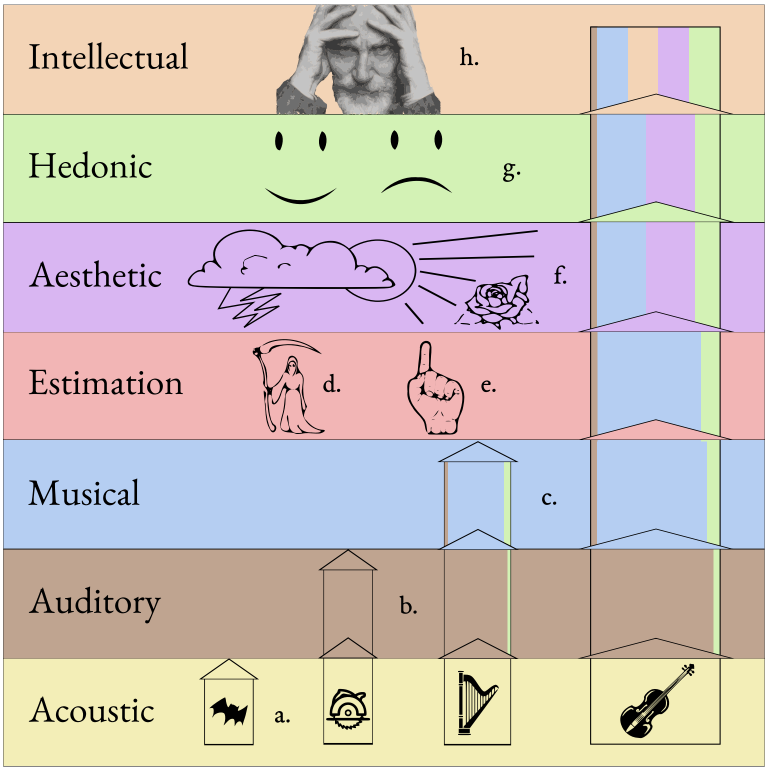

Figure 1.1: The columns represent listening experiences, only the last of which is fully successful. Each layer enriches the experience with a new kind of perception. Acoustic phenomena occur in the listener's environment. (a) Some are inaudible, but (b) some result in sounds. Not all dispositions of sounds interest the ear, but (c) some result in the musical experiences of rhythm, melody, and harmony arranged in intelligible textures and forms. (d) The unconscious gatekeeper, who may be overridden by (e) the conscious will of the listener, decides whether to pay further attention. (f) If so, the mind, for its own reasons, projects aesthetic experiences onto the music and (g) is stimulated by the hedonic charges that it has acquired in its passage through the layers. (h) The conscious mind may choose to reflect upon these experiences.

alter it, but for now it is the best framework I have for describing what is artistically useful in the practice of music-making. I have already mentioned its three main layers – auditory, musical, and aesthetic. Its fuller description involves seven layers of phenomena, the first physical, the rest mental – physical phenomena are perceived and there is a mental journey from perception to intellectual appreciation. Phenomena at earlier, lower layers cause phenomena at higher, later layers (or if they fail to, you have a failed composition). In general, each layer has access to all previous mental layers, but its principal input comes from the layer immediately before. In order, from earliest to latest, lowest to highest, they are: 1. Acoustic, 2. Auditory, 3. Musical, 4. Estimation, 5. Aesthetic, 6. Hedonic, and 7. Intellectual. The remainder of the chapter is an overview of each layer, the mental objects they produce, and what to do about them.

The seven layer model of music perception: fig.1.1, above

1.1 The Acoustic Layer

This layer consists of the physical phenomena going on around listeners, some of which they perceive as auditory phenomena – sounds. The key point is that we distinguish between physical events (as "acoustic") and their perception (as "auditory"). The physical phenomena are mechanical waves originating in musical instruments and traveling usually in the medium of air. This layer is the province of Musical Acoustics and is largely beyond the scope of this book.

Musical Acoustics, understanding how natural acoustic events (loosely speaking, “natural sounds”) and performance spaces behave, is useful for informing synthesis and sound processing. There is nothing magical about natural sounds, but because our hearing system is heavily adapted to them (and especially to human speech) they are especially salient to us. So, sounds that have been synthesized with characteristics typical of natural sounds are especially useful artistically.

All remaining layers are mental, existing only in the mind of the listener.

1.2 The Auditory Layer

For very practical reasons, animals need to know what is going on around them. Evolution has provided us with sensory systems that do this, and we humans have managed to hijack them for other purposes, including art. The auditory system does not merely detect acoustic phenomena. It also analyzes them, estimates the reliability of its analysis, and integrates this into a mental model of the listening environment.

It does so in the following way:

1. Two simultaneous acoustic signals (one measured at each ear) are resolved into sequences of discrete sounds which may overlap in time; i.e., sounds are events which come and go in time, and we can perceive several at once.
2. The system measures six characteristics of every sound:
   Timing: awareness of its starting and ending points relative to those of other sounds in the immediate texture.
   Duration: awareness of the length of a sound in time.
   Loudness: a sense of the energy in a sound.
   Pitch: the sensation of low or high corresponding to large or small sources – lions and tubas make low sounds; piccolos and birds make high sounds.
   Timbre: the sense by which we know that a sound was made by, for example, a clarinet or a piano, with all other characteristics being the same.
   Location: a sense of the direction which a sound comes from, and an estimate of our distance from the source.

3. It continuously segregates the incoming sounds into streams (it can keep track of several at once) based on these measurements.
4. It hypothesizes each stream to come from a different source, and it attempts to identify the nature of each, particularly its location.
5. These analyses form a picture, the "auditory scene."
6. The mind attempts to integrate it with the other senses into a model of the environment.
7. It continually assesses the certainty of its own analyses.

The auditory scene informs the overall mental model of the environment, but its information may be overridden by the model, e.g., if you hear a trombone on the left but see one on the right, you may end up perceiving it on the right; or, you may perceive drumbeats when you *expected* them rather than when they occurred. The auditory scene is an

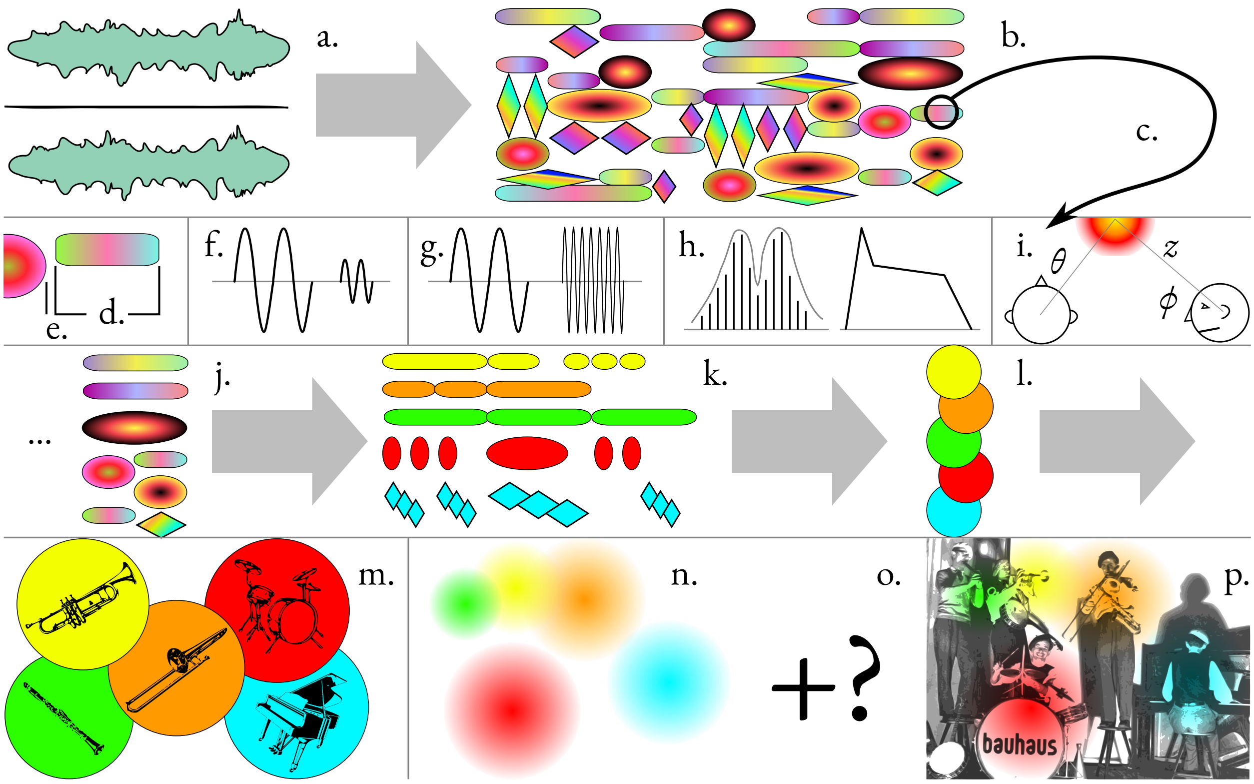

Figure 1.2: The Auditory Layer proceeds as follows: (a) Acoustic signals at each ear are interpreted as (b) a torrent of discrete sounds in time. (c) For each, the ear measures its: (d) duration and (e) timing relative to nearby sounds; (f) loudness (mostly related to amplitude); (g) pitch (mostly related to frequency); (h) timbre (mostly related to spectral and temporal envelope); and (i) position – the direction and distance from the listener, (θ, ϕ, z). Based on these measurements, (j) the sounds are segregated into (k) streams posited to come from (l) sources. The ear (m) identifies them and (n) locates them spatially. All of these analyses come with (o) an assessment of their uncertainty. Finally, the auditory sense is integrated with others into (p) a mental model of the environment.

hypothesis which may change fluidly – sources my change their character; streams may fuse, split, or move. The system estimates the accuracy of its analyses and model. It has a sense of its own uncertainty, and uncertainty makes the brain nervous – unresolvable information in the input signal could indicate danger in the environment – and may cause more attention to be devoted to some aspect of listening. The auditory system also interacts with other mental faculties, like memory and learning, and we can consciously direct our attention to it, putting more effort into certain analyses, e.g., to pick out every voice in a fugue.

The auditory process: fig.1.2, on the previous page

Musicians have exploited various aspects of the auditory system for their art. Soon we shall see how, through further processing, aspects of sound lead to musical experiences like rhythm, melody and harmony, but certain products of the auditory layer itself are exploitable. The bundling of sound sequences into streams forms the basis for melody. The capacity to track several streams simultaneously has lead to counterpoint, the art of composing simultaneous melodies (polyphony). Auditory scenes are experienced as "textures," and their artful setting is the business of arranging and orchestration. Considerations of auditory phenomena in our music-making will lead us to practical questions of performance nuance, synthesis, localization, and spatialization. The basis of the auditory layer is physiological and neural, and it is studied mainly by Perceptual Psychology and Music Cognition, particularly under the theory of "auditory scene analysis." The auditory layer serves up its objects (sounds, streams, sources, and the scene) to the next layer.

1.3 The Musical Layer

We share the acoustic layer, in some form, with most vertebrates. In the musical layer, anomalous neural processes, possibly unique to humans, lead to the experiences we call, "musical." When the incoming auditory objects are of the right form and in the right sort of patterns, they engage neural circuitry that is meant for other tasks. For example, when a sound stream has well-defined pitches and meets other criteria, it engages the circuits devoted to speech processing, causing the sensation we know as "melody." When the incoming sounds fall within certain parameters and in certain types of pattern, they engage the circuits devoted to motor control, causing the sensation we know as "rhythm" (and, incidentally, because these are the circuits that *plan* our motions, we feel compelled to dance or tap our feet). I can describe about twenty such varieties of musical sensation, falling into five categories: 1. Rhythm, 2. Melody, 3. Harmony, 4. Texture, and 5. Form. I don't know if these are all that there are – science may discover more – but these are the ones most important to the art of music.

The categories of musical sensation: fig.1.3, on the next page

I do not know whether they are all artifacts of non-musical neural systems. Some theories speculate that music has an adaptive function, so that we have evolved neural systems devoted purely to creating musical sensations. The other possibilities are that musical senses are "spandrels," non-adaptive by-products of our neural evolution, or that we are in the early stages of developing our musical spandrels into advantageous traits. In any case, we are concerned with their use, not their origins. Our game is to exploit all that is salient.

In the remainder of this section, I give brief descriptions of the five categories, mainly to show that they are not mysterious, but definite and easily understandable. A fuller discussion of each, including their various parts, forms the bulk of this book. We can regard them as the different types of mental object derived from the auditory objects and passed on to the next layer. The sensations always come to us intertwined and the business of composing each kind is interrelated with the others, so there is no perfect order in which to present them. This is the clearest presentation I can come up with – it might require some repeated reading. I will be loose with some terminology to be made precise later. From now on, when referring to the faculties of the auditory and musical layers together, I will avoid loquacious constructions like "the human auditory system," or "the human cognition of music," calling it simply "the ear," even though the equipment is mostly neural.

Introductions to the five categories of musical sensation: §§ 1.3.1–5, pp.7–30

The objects of the Musical Layer: ??–??, ??–??

1.3.1 Rhythm and Its Antecedents

Recall the beginning of the auditory experience: an incoming acoustic signal is divided up into (possibly overlapping) discrete events – sounds. We hear sounds as events in time; to compose music we plan and schedule them. Let's call planned sounds *notes* and their schedule a *score.* The score represents the surface level of music, what the musicians will play and the listeners will hear.

Rhythm in detail: ??, ??

Along with their start times and durations, notes typically have some information about how they are to be played or used, e.g.: loudness, pitch,

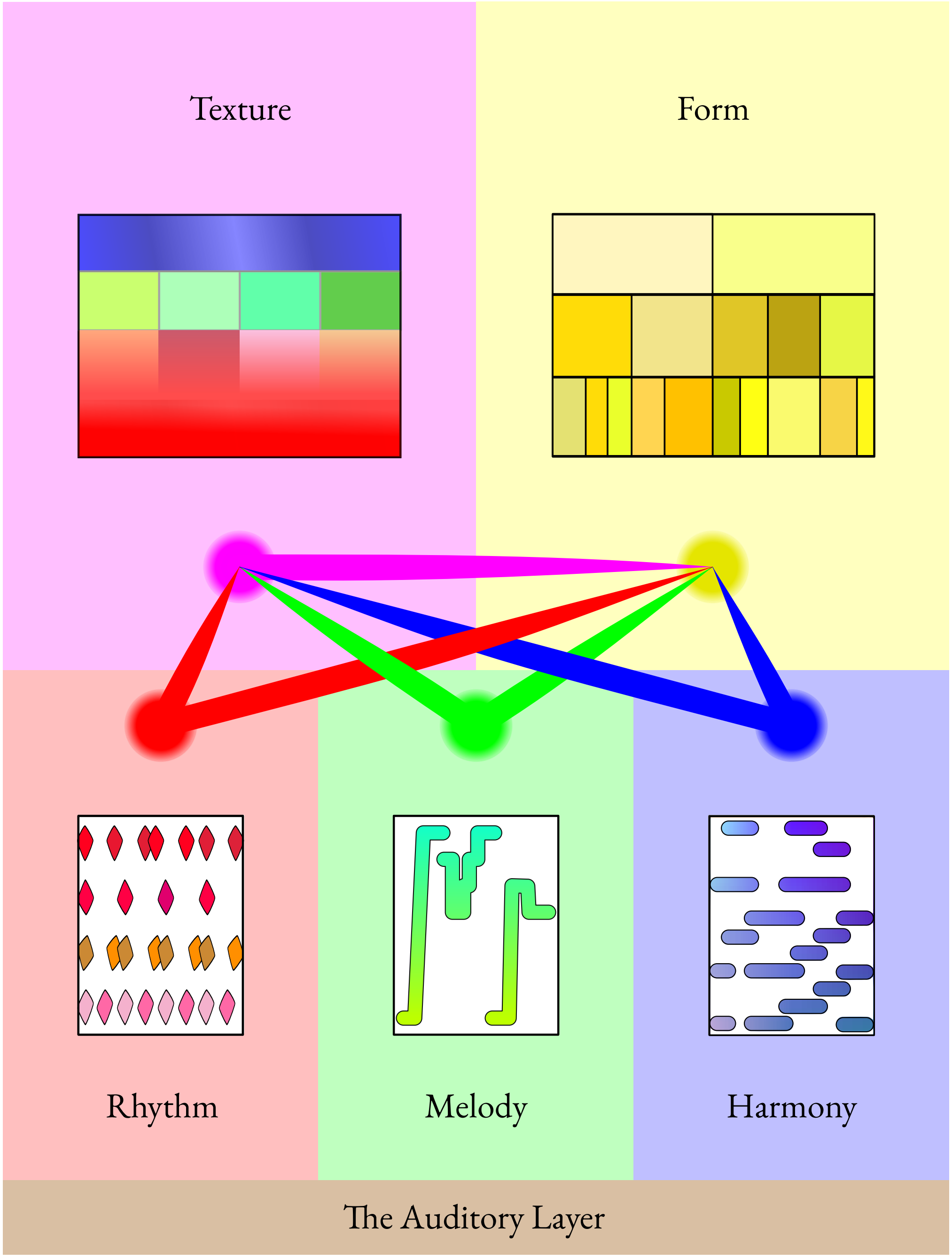

Figure 1.3: The Musical Layer arises from the Auditory Layer, beginning with the primary sensations: Rhythm, Melody, and Harmony. Regular recurrences of sounds in time cause Rhythm. Sequences of tones that seem to be connected cause Melody. Simultaneous tones – sonorities – appearing in succession cause Harmony. At any moment, these three fit together into a mental picture, a Texture. Probably, one or more melodies contribute to the harmony and, through their individual rhythms, to the overall rhythm. In a piece of any length, both the texture and the particulars of the three primary experiences – the rhythmic, melodic, and harmonic material – will likely change over time, contributing to the experience of Form, a mental division of the piece into a hierarchy of similar and contrasting sections.

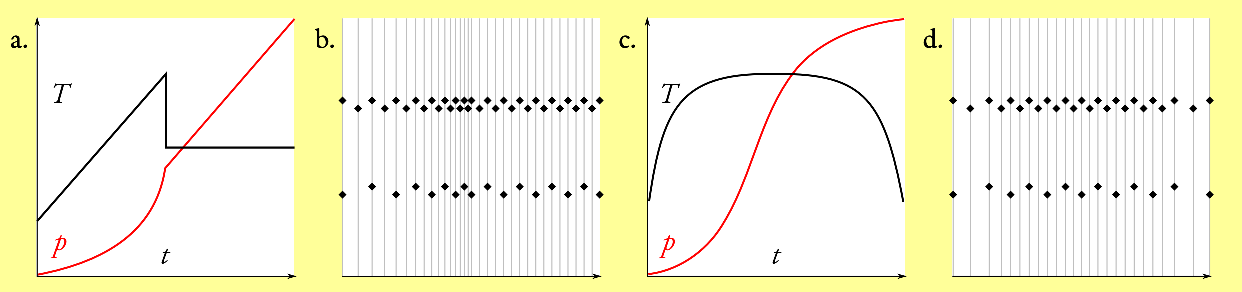

Figure 1.4: (a & c) To control the speed of the music, the composer sets the tempo T (the black line) in units of pulses per second; it may vary over time t. This creates a new timescale, p (the red line), in units of pulse. (At any time t, the value of $T(t)$ is the slope of $p(t)$, so $p(t) = \int T(t)\,dt$.) (b & d) A series of events spaced regularly in pulse time, when heard in physical time, reveals the speed the composer intended. Some common cases: (a & b) Different tempi help to demarcate sections. Continuously varying tempi (as in the first half) can provide dramatic effects. (c & d) Accelerating into and decelerating out of a phrase is a typical performance nuance.

timbral parameters, harmonic "strength" (whether it should reinforce the harmony or be dissonant against it), or melodic "strength" (whether it is important to the melodic structure or is merely ornamental), etc. For reasons having to do with melodic composition that will become clear later, I'll use a non-standard term, *shade,* for the non-pitch information in the notes.

If a pattern of sounds in time (i.e., a pattern of relative onset times) that is not too long repeats with enough regularity, neither too fast nor too slow, the listener will perceive a *rhythm.* Loosely, we may call the timing of a set of notes its "rhythm"; but strictly, a rhythm is this musical sensation we are trying to induce in the listener. Even if it is not especially prominent, most music has an underlying rhythmic structure. Rhythm is one of the few channels by which listeners can be sent musical experiences, so composers only deprive themselves of the resource occasionally, for dramatic effect. Rhythms are typically built up from deeper structures that the composer plans and that listeners sense as parts of the rhythmic experience: 1. tempo, 2. meter, 3. groove, and 4. rhythmic mode.

We begin with the physical level of real clock-time measured in seconds. *Tempo* abstracts this to a warped timescale measured in *pulses.* It is simply a quantity of pulses per second that gives us control over the perceived speed of the music. Tempi may be static or change in time, either abruptly or continuously. Different layers of warping may be applied at once to create a rich terrain (e.g., a "terraced" plan of fast and slow movements, along with a continuous change over the course of sections, and accelerations and decelerations to mark the beginnings and endings of phrases).

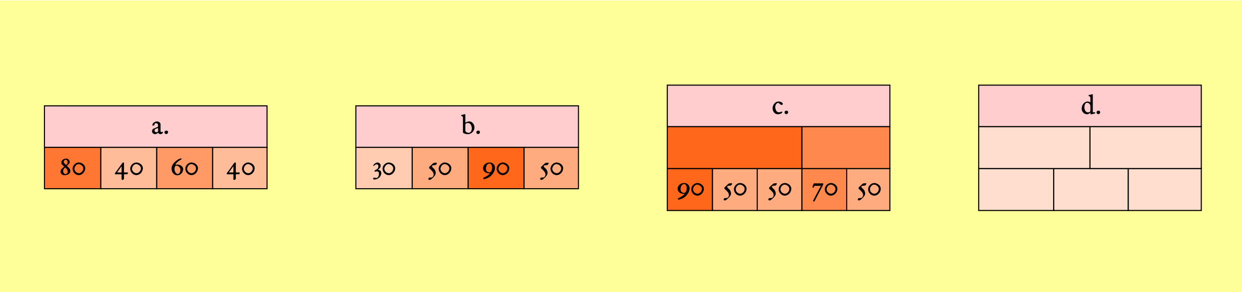

Figure 1.5: Some examples of meter. The pink bars represent the undivided measure. The orange bars represent the beats. Some have shades filled in – in this case, simply loudness on a scale of [0, 100]. (a) Common $\frac{4}{4}$. (b) Reggae has the same division (four equal beats), but different shades produce a very different feel. (c & d) More complex meters may have any number of concurrent schedules of beats, represented here by courses of orange blocks under the measures. (c) An "asymmetric" meter, "**3+2**," with beats of unequal length (upper course) which may be interpreted as unequal groupings of equal-length beats (lower course). (d) A polyrhythmic meter, "**2** against **3**," simultaneous equal divisions of the same measure. Note how the beat onsets needn't coincide except at the start of the measure. When meters are repeated, this provides the minimal regularity for the sensation of rhythm to occur.

Tempo: fig.1.4, on the previous page

So, just by scheduling notes in units of pulse, you can incorporate all your decisions about speed without having to think about them explicitly.

Meter further abstracts pulse-time into a structured timescale of recurring *measures* holding a recurring sequence of (possibly unequal) *beats*, both anchored to pulse time. Complex meters (such as "polyrhythms") may have several simultaneous sequences of beats, each producing a parallel timescale (or we might think of them as subchannels of one complex timescale). It is to the beats that shades are typically assigned, creating the familiar repeated patterns of stress. Meter contributes the regularity needed for the sensation of rhythm, and the pattern of its beats gives a rhythm much of its character. Tradition teaches us a rich variety of paradigms for composing meters. So, simply by addressing time in units of measure and beat, you can incorporate all your decisions about rhythmic structure without having to reconsider them for every note.

Meter: fig.1.5, above

Upon meters are scheduled (usually subtle) modifications – deviations of timing or inflections of shade, called *groove*. This may change by a pattern applied to several measures to enhance formal structure, or continuously over time, to achieve subtle effects of tension and relaxation. Groove adds character and a human feel to rhythm, making it more savory to the ear.

Groove: fig.1.6, on the facing page

Finally, to produce a rhythm from a meter, small patterns of notes called *rhythmic primitives* are assigned to the beats, according to a set of rules. These are typically single notes that relay the shade of the beat,

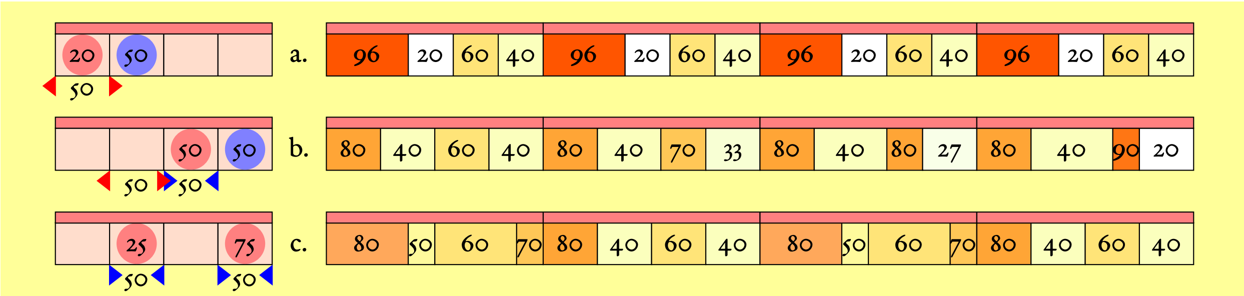

Figure 1.6: On the left: three grooves, patterns of percentage changes to beat durations and shades (once again, just loudnesses). On the right: each groove applied to four measures of the common $\frac{4}{4}$ meter from the previous figure. This is done in three ways: (a) The same to every measure, a steady "swing." (b) Increasingly, until full-strength in the last measure – this can be done for a subtle dramatic effect. (c) By pattern, to every other measure – this can be done to emphasize phrase structures. Only m. 1 of (b) and m. 2 and 4 of (c) are "played straight," without any groove. I have chosen larger than typical percentages for the sake of clarity.

but they may be rests or small constellations of notes that can adapt to their surroundings and carry modified beat shades. The rules are a sort of grammar which may allow for random choices and the influence of shades. Placing notes on beats connects the score level – through meter, groove, and tempo – back to physical clock time, so that the rhythm can be performed. Collections of rhythmic primitives and methods for assigning them to beats – manners of realizing meters – are called *rhythmic modes*. They contribute to the character of the resulting rhythms, and can create "family resemblances" between them, even when the meter differs.

Rhythmic mode: fig. 1.7, below

Any of these four parts of rhythm – tempo, meter, groove, and rhythmic mode – may vary over time. They are objects that can describe a listener's experience and, by creating and arranging them, we compose that

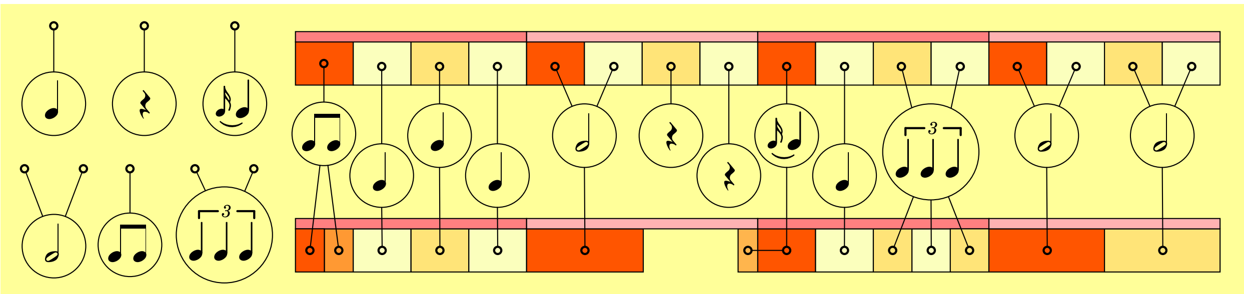

Figure 1.7: On the left: rhythmic primitives – small clusters of notes that attach to beats, copying their shades to the notes (possibly with modifications). In order: a note of one beat; a rest (no note) of one beat; a note of one beat with a grace note (a small note before); one note in the time of two beats; two in the time of one; three in the time of two. On the right: Some of these are hung on the beats of four measures of common $\frac{4}{4}$ meter (above), creating a playable score (below). The rules about how to apply them are a sort of grammar which may be probabilistic and depend on certain shades. The rhythmic mode is the set of primitives and the rules for applying them.

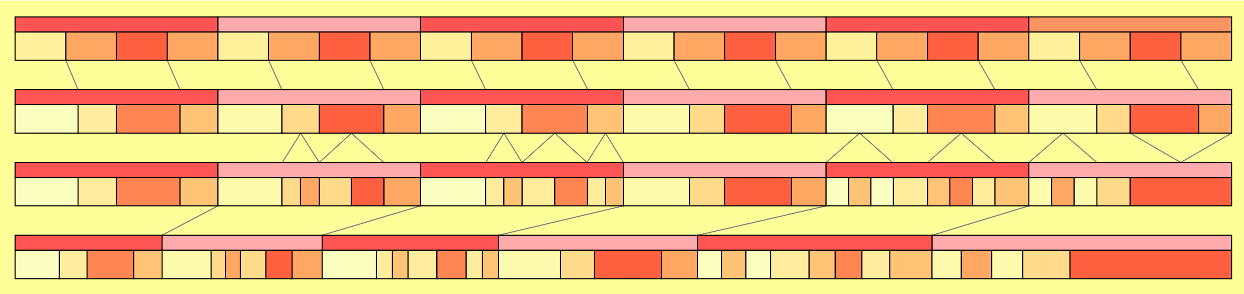

Figure 1.8: A rhythm is built, from top to bottom. First row: a meter is selected – the version of reggae presented earlier. Second row: a groove is applied to this, increasingly over time. Third row: the beats are replaced with notes by a rhythmic mode, creating a score in pulse time. Fourth row: finally, this is warped into physical time by a constantly decelerating tempo. A composer combines a few simple ideas and automated computation realizes them as a rich, organic musical structure.

An example of rhythmic construction: fig.1.8, above

experience. The results are surface-level rhythms that we can send to the orchestra to play or that we can use to build melodies or harmonies.

I want to pause a moment to point out a theme which will recur throughout the musical layer. Consider the construction method I've suggested. The composer chooses from stock objects or creates new ones from scratch:

1. A tempo curve that may outline the formal structure of sections and phrases.
2. A meter to impart a rhythmic feel (perhaps including shades that influence rhythmic, melodic, and harmonic composition).
3. A grove or patterned schedules of grooves to enhance the meter with human drama and nuance.
4. A rhythmic mode to realize the meter as a rudimentary score with another layer of character.

At each step of the process, the composer is making decisions about easily understood objects which correspond directly to aspects of the musical experience that will be induced in the listener. When these ideas are combined and worked out, the resulting rhythm, a score fragment, will have a rich, organic, subtly-textured surface, filled with human nuance and artistic significance. This may then be played by virtual performers with unpitched instruments, or used as the basis for melodies or harmonies. The application of the composer's ideas might involve a lot of tedious mathematics, but this part is automated. This is the nature of good computer composition: from the composer's side, interesting choices about sophisticated, perceptually significant objects; from the listener's side, a rich experience determined by those objects. By speaking the ear's language, the software facilitates clear communication from composer to listener, and avoids the possibility of sending the ear messages that it cannot interpret. By automating the low-level tedium, it allows the composer to concentrate on higher things. The result is more fun for everyone.

I'll provide software for composing and performing rhythm:

- The orchestra has many built-in features that handle tempo, meter, and groove independent of the score details. In the score, one typically writes timing information in terms of measures and beats and leaves dynamics blank. The orchestra works out all the details of timing and loudness automatically, according to the composer's centrally located specifications of tempo, meter, and groove.
- Apart from the orchestra, an interactive program with integrated editors for designing and scheduling:
  - Multi-layered tempi;
  - Meters (with support for many traditional and Euclidian construction techniques);
  - Grooves; and
  - Rhythmic modes.

  It features a click-track of percussion samples allowing adjustments to be heard in real time and includes a mechanism for outputting the resultant rhythms as score fragments ready for the orchestra or for melody programs.

1.3.2 Melody, Scale, and Mode

Melody is a sensation that comes from sequences of tones (pitched sounds) in time. Not just any sequence is a melody. But, if the ear has segregated the tones into the same stream (so that they appear to come from one source) and if certain conditions about them are met (e.g., well-defined pitches with no long silences between them, no large, "unprepared" leaps in pitch), then they will not be heard as simply separate sounds of differing pitch, but as a single entity *moving* in the pitch space – an illusory entity we call a "melody." Consider how a dog hears a melody played on a piano. His mind determines: "Unconnected sounds coming from a source over there. Not a threat. No further processing." For him, there is only an auditory layer. For us, the lucky illusion of melody opens the door to art.

Melody in detail: ??, ??

The conditions for melody: ??, ??

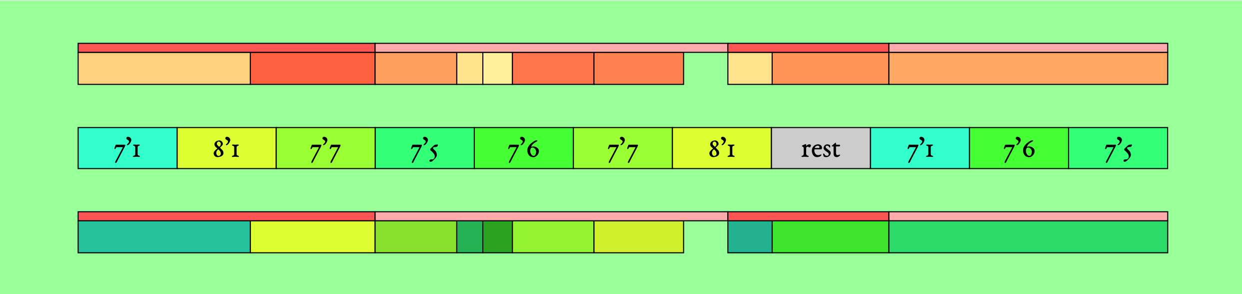

Figure 1.9: Top: a *talea* is a list of durations which may contain shades (here only the loudness is shown). Middle: a *color* is a sequence of pitches with no timing information (the pitch notation will be explained later). Bottom: the result of combining them is a melody (here hue shows the pitch and the lightness shows loudness). The tune is *Over the Rainbow* – color and talea by Harold Arlen, groove by Judy Garland.

There is a problem with conceiving of melodies. Their essence is a sequence of pitches. It would be convenient to think only about the reasons for one pitch to follow another and to compose them accordingly. This is inadequate because melodies are affected by everything in their musical environment. The reasons to choose the pitch of a note may depend on many things, including the melodic mode, the presiding chord, and the note's length, loudness, and metric position. In particular, a melody occurs in time and is bound to its own rhythm – it cannot be realized without timing information. The significance of a note's pitch depends on its place in the rhythm. So, a melody is not merely the sequence of pitches – you may keep the same pitches, but change the rhythm and you have changed the melody.

In order to disentangle this complexity and to facilitate algorithmic composition, I have revived a medieval concept. Melodies may be thought of as having two components:

Color: a sequence of pitches (or rests); and

Talea: a corresponding sequence of durations (with optional shades).

It is much easier to compose and work with the components separately, in their own terms, though with reference to each other. We may disassemble an existing melody to work on its color and talea separately. We may compose a rhythm to use as a talea, and then compose a color for it. We may even compose a color first and then come up with a talea to fit it.

Color + talea = melody: fig.1.9, above

A few formal restrictions must be observed to make these objects computable. Obviously, the two sequences must have an equal number of elements – when a talea is associated with a color of equal length, we have a realizable melodic score. Taleæ are like rhythms where the notes are guaranteed to come one after another with no gaps or overlaps between them. In practice, often, a talea will have been constructed as a rhythm, and may carry a shade with each duration.

Though we could compose melodies where each pitch is drawn arbitrarily from the audible pitch continuum, this is not done traditionally. Apparently, the ear likes to posit that melodies are drawn from a small set of pitches. It divides the pitch continuum into bins corresponding to members of the underlying set, and any tone that falls into a bin is experienced as an instance of that member – a symbolic pitch. As usual, since we seek to oblige the ear, we conform our practice to it by speaking its language. So, when composing, we don't work with literal frequencies or pitch units; we use symbolic pitches – members, called *degrees*, of predefined sets called *scales* that refer to bins on the pitch continuum.

Two other objects work with scales. A *tuning* is a mechanism that calculates the exact points within each bin that a scale refers to. A *mode* is a subset of a scale, important because melodies often do not use every member. Modes are a notational convenience, re-indexing the scale by ignoring unused degrees. Modes also have an artistic function: the choice of which degrees to exclude imparts character to the melody. The large

Symbolic vs. literal pitches: fig.1.10, below

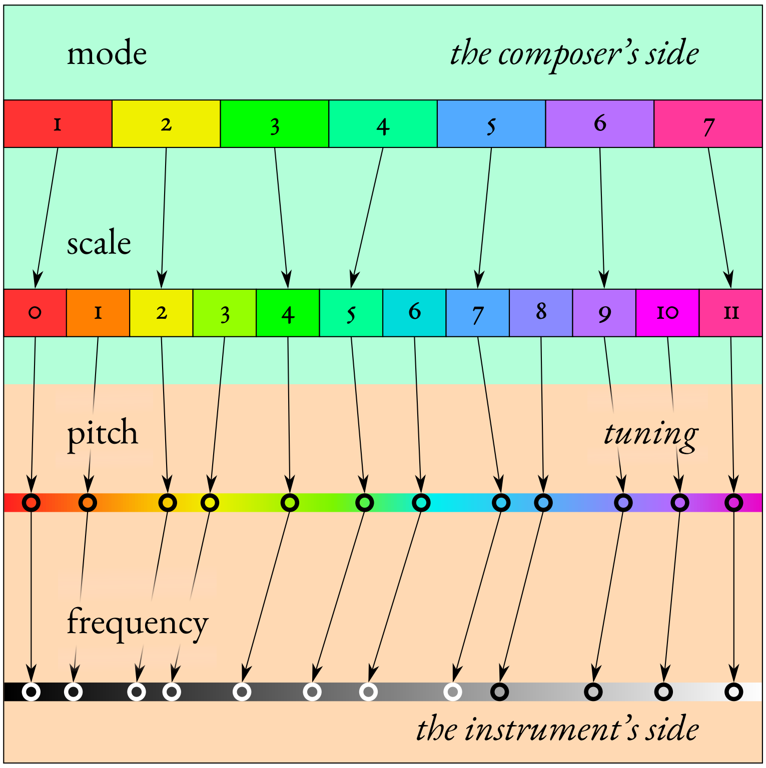

Figure 1.10: Composers work with symbolic pitches, degrees of modes and scales. Precise, literal pitches and frequencies are the business of instruments and performers (for us, the computer). Scalar and modal degrees are indicated by index numbers (but typing numbers is exciting, rather like pushing prettily-colored buttons). Modal degrees are aliases for scalar degrees. Tunings determine the precise value of the degrees, which may change according to musical circumstance, even during performance. The mapping of pitch (auditory sensation) to frequency (acoustic fact) is physiological. A composer chooses scales, modes, and tunings; the operations indicated by the arrows are automated, and don't normally require attention. Note the unevenness of the pictured tuning, a just intonation centered on a septimal minor chord.

Scale, Mode, and Tuning: ??, ??

subject of scales, modes, and tunings, will require a chapter of its own before melody and harmony can be fully explained. The pitches contained in colors (and sonorities and scores) will be recorded as degrees of scales or modes. The pitches spoken of in this book are mostly of this kind. For us, tuning is automated; its operation and the resulting literal pitches and frequencies will not require our attention.

Our system features an important innovation: it separates the design of scales and modes from the design of tunings. The reason for this has to do with the nature of human hearing. Generally, the ear doesn't care about absolute pitches – isolated tones can differ greatly in pitch without the ear noticing. The ear cares about relative pitch, the pitch differences (called *intervals*) between tones. Its sensitivity to harmonic intervals (those between simultaneous tones) is much greater than its sensitivity to melodic intervals (those between consecutive tones). So, scales needn't be chained to a precise tuning. It is better if we let them represent that hazy picture of a source set that the ear posits when it hears a melody. We can leave the precision to another object that acts later in the process, the tuning. Melodically, the important thing about scales is their gapped structure, the vague terrain of step sizes the ear appreciates, determined by modes. Harmonically, the important thing about scales is the precise intervals that occur in the music, determined by tunings. Scale and mode design is best done for melodic reasons; tuning design is best done for harmonic reasons. Separating the two gives us access to the full range of melodic and harmonic possibilities, while allowing the tedium of tuning to be automated and put behind the scenes.

The ear's view of scale and tuning: fig.1.11, on the facing page

Usually, composers select an existing mode and tuning, usually the defaults of their tradition, and get on with their work. But, like meter, modes and tunings are a great under-used resource, and I'd like to see more of the possibilities explored. Understanding them will help you to make informed choices about existing ones. Also, I'll provide tools that automate the difficult aspects of creating new ones.

Terminology can be tricky in this area. By "scale," I mean what are sometimes called "chromatic" scales, the roughly uniformly-distributed parent sets from which gapped child sets are selected. By "mode," I mean the re-indexed subsets, the gapped child scales of the chromatic parents. By "melodic mode," I mean a manner of composing melodies, expressible as a set of probabilities or other rules. *Key* names the scalar degree on

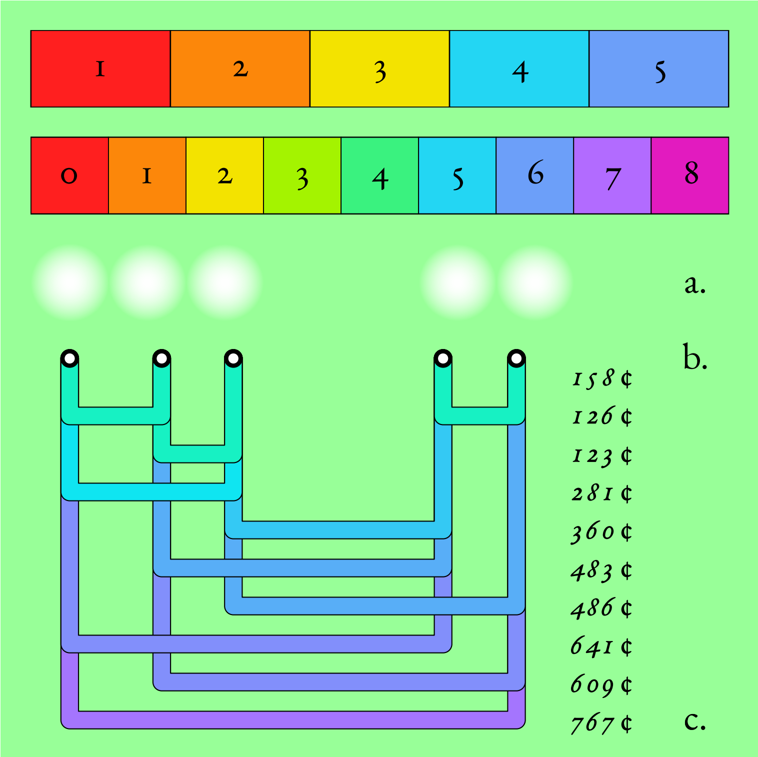

Figure 1.11: The first two rows show the composer's view of mode and scale as before – uniform indices. (a) The listener's experience of a melody includes modeling its underlying pitch collection. A bin is confirmed as a member of the mode whenever a tone falls in it. Even if the pitch is always the same, the model is as nebulous as the perception of sequential intervals. The pattern of gaps and members is distinctive and colors the melody. (b) The precise value of the pitches is set by the tuning for acoustic reasons: (c) the precise intervals between simultaneous tones cause distinctive sensations, contributing to harmony. The pictured mode and tuning is an actual realization of *selisir,* from Bali. Note the irregular positions of the pitches in (b) – the equally-sized bins do not imply equally-spaced degrees.

which a mode begins. Changing key means moving the pattern of modal selection up or down, a mere transposition of the music. The choice of key is often arbitrary, e.g., to accommodate a singer's range, but as we shall see, tonal harmony likes to play games with it. Also, when using fixed, unequal tunings, the choice of key is effectively the choice of a tuning. In this way, keys of the same mode can have individual characters, and we shall play our own games with that effect.

We are concerned with composing melodies. Our output therefore is matched pairs of color and talea, given a variety of inputs. Later, under the heading of harmony, we will deal with melodies created in reference to other melodies, or emerging from harmonies (counterpoint). First, we must consider methods for composing colors given taleæ (and for creating taleæ, hence rhythms, to fit existing colors). Although there may be a stochastic element to choosing the pitches of a color, they are not simply random. We need to understand the relation between melody and the objects which influence it, the "variety of inputs." At least six things influence the composition of colors:

1. *General melodic rules:* roughly, these favor melodies similar to speech and restrict us to tone sequences that will be experienced *as* melodies.

2. *Melodic mode:* rules and tendencies that give a melody character, and cause melodies of the same mode to have a family resemblance.
3. *Phrase structure:* an aspect of musical form that orders music into temporal units like those of language – words, phrases, sentences, paragraphs, and so on.
4. *Imitation and contrast:* if we wish past and future parts of a melody to be similar or dissimilar.
5. *Harmony:* which sometimes asks melody to support chords.
6. *Rhythm:* which supplies the color with a talea and shades that inform its composition. Shade information can include:
   a) *Melodic significance:* the relative importance of a note to the melody – whether it is structural, ornamental, or something in-between.
   b) *Melodic function:* the probability that the pitch will be a member of certain abstract classes called "functions" (e.g., "tonic," "dominant," etc.) which are expected to appear in certain places and to form sequences according to a sort of grammar.
   c) *Relationship to harmony:* the probabilites of the pitch being a member of the chord or not (and, if so, which member; or, if not, how badly it might clash).
   d) *Interpretations of other shades:* if the above shades are not given explicitly, a melodic mode might choose to interpret them from existing shades like loudness, timbre, and metric strength.

All of these inputs can be expressed as Markov chains (sophisticated representations of conditional probability) and used in melodic programs.

I'll provide software for composing and performing melody:

- The orchestra has many facilities for:
  - Declaring modes and tunings, including tunings that can adapt to the harmony.
  - Indicating pitch according to scale or mode (and in other ways).
  - performing melodic lines gracefully, including easily notated true legato and other articulations.
- An interactive editor for the design of scales and modes.
- An interactive editor for the design of melodic modes, including the grammar of melodic function.
- The composition of colors given a talea (and optionally, any of the above influences and a partially given color) using Markov chains.

- The composition of taleæ given a color, by similar methods.
- The composition of melodies by other methods, e.g., centonization and partimento.
- Other operations on melody: variation, reduction, elaboration, ornamentation, scale recasting, etc.
- Methods to analyze existing melodies or colors to Markov chains for use by the above.

1.3.3 Harmony and Tuning

Harmony is the sensation that comes from simultaneous (or closely associated) sounds of different pitch. This may be described as a

Harmony in detail: ??, ??

Sonority: a set of specific, literal pitches; or as a

Chord: a set of scalar degrees without octave information (therefore, a chord is equivalent to a class of sonorites that differ only by inversions and octave doublings).

We will deal with both, but mostly with sonorities, because it is through them that listeners' natural experiences are most easily understood.

Sonorities have perceived qualities, called *flavors,* that can be predicted given the set of pitches. Flavors have six components; four come from individual sonorities:

Sonorities and flavor: fig. 1.12, p. 21

1. *Dissonance* sums up the ear's measure of its own uncertainty in making sense of confusing audio signals. For evolutionary reasons it finds this sensation unpleasant, and so feels a pleasant relief when dissonant sonorities resolve to consonant ones.
2. *Tension* is a more savory, less objectionable sort of uncertainty which arises from tones placed symmetrically in pitch (e.g., as in diminished and augmented chords). Like dissonance, tense sonorities create a sort of instability that the ear likes to hear resolved to lax ones.
3. *Modality* describes the flavor of the laxity. Consider a perfectly tense, symmetrical triad (three-tone sonority). It has *neutral* modality. To become asymmetrical, the middle tone must move either up or down. These possibilities produce the distinctly different sensations we call *major* and *minor.*
4. *Richness* is a sensation of the complexity that comes with larger sonorities. As the number of tones in a sonority increases its dissonance tends to increase, but, pleasantly and paradoxically, richness

tends to mitigate the harshness of this – perhaps because the ear has its hands full tasting other things.

...and two come from the transition between sonorities:

5. *Progression* is a satisfying feeling of forward motion, which can be analyzed as a combination of the individual melodic motions that the transition contains. This is the one component that may not be completely universal. All or part of it may depend on expectations listeners learn from particular cultures. This allows us the convenience of defining the expectations in one place.
6. *Counterpoint* is a measure of how clearly the ear will be able to follow the individual melodic motions. A high score indicates that they are distinct, a low score that they are fused. You may wish to create textures in which some melodies are distinct and others fused. Our mathematical definition of counterpoint will be sophisticated enough to take this into account.

A *harmony* is a sequence of sonorities, and so, a sequence of flavors. *Harmonic modes* are manners of composing harmonies. I will present two paradigms for harmonic mode:

Tonal harmony: the traditional theory which posits that chords are perceived as locations (called "functions") in an abstract space, making harmony a journey through that space.

Flavor harmony: my own theory, which composes harmonies according to criteria of flavor and flavor change.

Tonal harmony is a cultural practice that depends on learning the abstract locations attributed to certain chords or groups of chords. Even within its culture, most listeners have not studied it formally. They have, at most, picked up unconscious expectations about chords appearing in typical positions and sequences, usually from the endless rote of popular music. The flavors, by contrast, are a natural part of sound cognition that come to all listeners, whatever their training (with the possible exception of the progression component). Therefore, flavor is a far stronger effect to exploit for art than is tonal expectation. By composing in terms of flavor rather than tonal function, the harmony is automatically rooted in the listening experience. Combined with automation, this frees the composer from having to learn the vast vocabularies of chord names. It is more efficient and to-the-point to simply speak of the six flavor components.

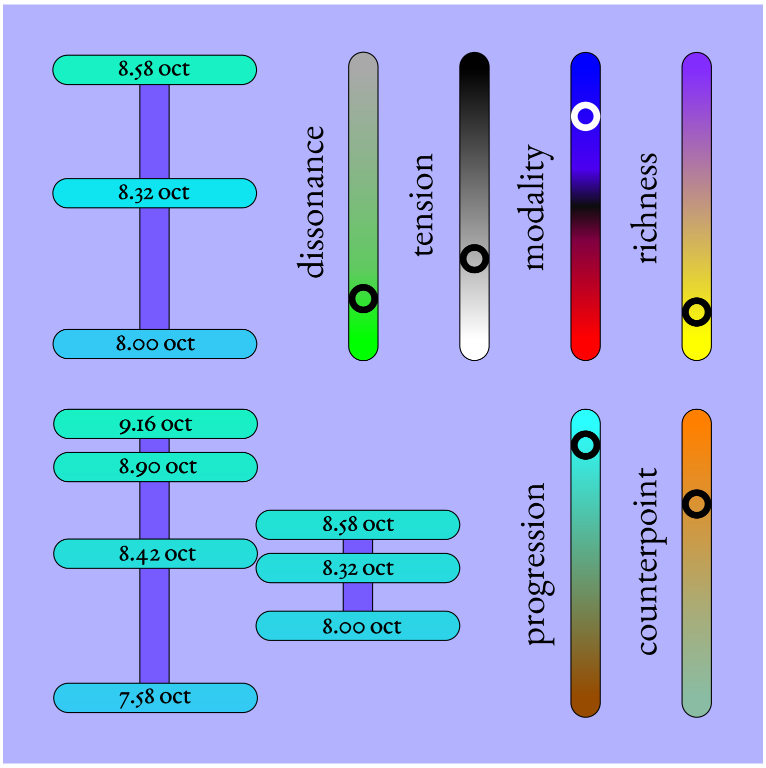

Figure 1.12: Top: A sonority has four flavor components, which are a function of its pitch set: *dissonance* is caused by the misalignment of the tones' partials; *tension* is caused by the symmetry of the intervals; *modality* (major or minor) is the direction the asymmetry skews; *richness* is caused by the number of unique pitches (disregarding octave repetitions). Bottom: A transition between sonorities has two flavor components, which are a function of the two pitch sets together: *progression* is a sensation of forward motion; *counterpoint* is a measure of how distinct (as opposed to fused) the implied melodies will be. Composing in terms of flavor makes it unnecessary to remember or even name the countless possible sonorities. By speaking in terms of the six components, one addresses the listener's experience directly.

Since a sonority is merely a list of pitches, a sequence of sonorities is equivalent to several simultaneous colors. Obviously, when the size of the sonorities varies, this brings up questions about whether some colors should play the same pitches or take rests, but this can be determined algorithmically. Any harmony can be automatically "unwoven" into several simultaneous colors. So, the composition of harmonies creates melodies as a by-product. This fortunate circumstance has a compliment. When composing a harmony, one of the most interesting constraints is *harmonization,* the requirement that the harmony can serve as accompaniment for one or more given melodies. As you can imagine, with harmony and melody generators accepting each other's products as input, you can create endless amounts of related musical material from small sources.

Harmonization and "unweaving": fig.1.13, on the next page

The flavors of sonorities depend upon the precise pitch relationships within the set, as determined by a tuning. There are countless tuning paradigms, but I will detail only the three that I think are artistically useful. They follow from three approaches to dissonance:

1. Adaptive minimal dissonance methods (including just intonation). By "adaptive" I mean that the pitches of the degrees are not fixed for the duration of the piece, but that they can shift to suit whatever

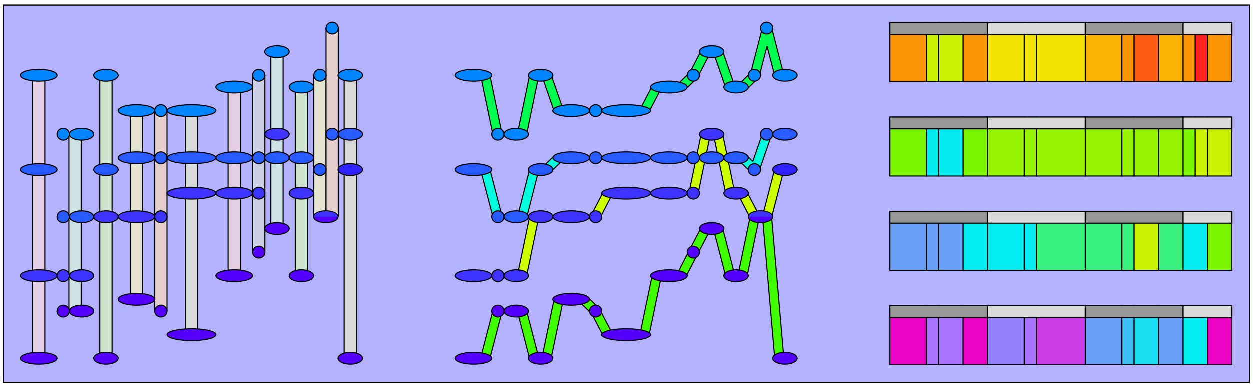

Figure 1.13: Three representations of the same piece of music, the *Gaudeamus Igitur* as set by Christian Wilhelm Kindleben (1781). The first two are in the style of a piano roll – time proceeds left to right, pitch from bottom to top. The last is as four simultaneous melodies, with pitch represented by hue. On the left, we see the music as a harmony, a sequence of sonorities, each with its own flavor. In the middle, we see the same notes with one solution to the problem of how they can be performed by four voices. On the right, these are fully unwoven and presented as simultaneous melodies. Consider the composer's method. He probably began with the folk melody (middle, top), harmonized it (left), and then distributed it to four voices (middle and right). If you've noticed that the counterpoint is poor, remember that it is, after all, a drinking song.

sonority is playing. Care is taken to minimize the shifts so that they are rarely, if ever, perceptible.

2. *Schwebungsdiaphonie* (maximal dissonance methods).
3. Well temperaments (artfully controlled dissonance) – and I will introduce a class of them, polychrome well temperaments, that can be used to maximize key characters and to create hypertonality, an optimization of tonal harmony.

The first is most suited to flavor theory, and the last to tonal harmony, while the second is a special case of limited but significant value.

I'll provide software for composing and performing harmony:

- The orchestra has facilities for:
  - Declaring chords (on several concurrent "chord channels," in support of polyharmony).
  - Indicating pitch in reference to the currently sounding chords.
  - Performing chords as "strums" on any instrument, with many naturalistic performance parameters, and with the "fingering" being determined automatically.
  - Performing in arbitrary fixed or adaptive tunings.
- A program for the composition of harmonies using flavor theory, capable of Debussy-level harmony with Bach-level counterpoint. It also calculates the tuning and is capable of harmonization.

- A program for the composition of harmonies via traditional tonal harmony, including harmonization.
- A program for the composition of polyphonic music based on free and imitative counterpoint – for canons and the like.
- Interactive editors for the design of tunings, which may be fixed or adaptive, and can take into account arbitrary instrumental spectra (not restricted to harmonic spectra). Including:
  - The design of maximally consonant tunings for given spectra.
  - The design of spectra to achieve maximal consonance given an arbitrary tuning.
  - The design of polychrome and hypertonal well temperaments.
  - The design of *Schwebungsdiaphonie* tunings.

1.3.4 Texture, Arrangement, Orchestration

At any moment during a piece of music, a set of musical objects of the above types (rhythm, melody, and harmony) function together in a kind of ecology – a deep texture. To make this structure audible, it is rendered into instrumental parts (by *arrangement*); realized with nuance (by *performance*); clothed in instrumental timbres (by *orchestration*); and set within an auditory scene (by *production*). These procedures create a surface texture with an ecology of its own, through which the listener experiences the deep texture.

Texture in detail: ??, ??

As composers, we want to give the audience a clear perception of an interesting deep texture through a beautiful surface texture. To do this we must solve a series of problems:

1. Designing the deep texture.
2. Mapping its elements to instrumental parts.
3. Realizing the parts by:
   a) Assigning them to performers,
   b) Assigning them to instruments,
   c) Placing and moving these in the listening environment, and
   d) Designing the listening environment itself.

The possible deep textures differ in the number of elements (typically few) and their relations, of which there are only a few types. So, the total number of possibilities is not vast. For example, there are probably one to five melodies, one harmony and one rhythm (possibly with layers for harmonic, melodic, and percussive parts), but these numbers may be larger or smaller. Elements may differ in relative importance or register, and melodies may be distinct from each other or part of a fused group.

Arranging, the art of mapping the deep texture to playable parts, enlarges the range of possibilities. The simplest arrangement is a one-to-one mapping – each melody gets one part; and, if they are not sufficient to support the harmony, or if a clearer rhythm is desired, parts are created for these tasks. A melody may be distributed to several parts ("doubling," done for reasons of timbre and surface texture, or for the dramatic effect of highlighting important passages), or divided between parts (e.g., by phrase, to create a call-and-response effect).

Harmonies usually require the most work. The existing melodies may be sufficient to represent the harmony (as in polyphonic music), but if not, extra material must be composed. This can be as simple as block chords or strums, or the void may be filled by arpeggios, or other figures, or even by the composition of new countermelodies. In any case, the new material should respect the rhythm.

The arranger should also judge whether the existing parts are sufficient to represent the rhythm or whether more should be added. These are usually unpitched percussion, but weakly-pitched percussion can also enrich the harmony. A rhythm may be distributed to parts according to the shades of its notes (e.g., strong beats to the bass drum) or according to its own internal structure – different combinations of metric layer, rhythmic mode, and even groove may be assigned to different instrumental parts, creating great depth and richness while still representing the unified rhythmic idea intended by the deep structure.

Our software orchestra supplies all the tools needed to realize the parts: performers, instruments, and the listening environment. It replicates the traditional division of labor whereby composers delegate to performers the rather dull task of deciding how to play each note (in our system, about a hundred decisions per note). It provides us the luxury of looking at performance nuance from a higher level – we assign parts to performers with different personalities (really, different packages of decision-making subroutines), e.g., a lead part to the flamboyant soloist, fused accompaniment melodies to restrained choristers, and polyphonic melodies to disciplined consortists (who, while not flamboyant, make different choices from each other to subtly maintain the independence of their parts). Orchestration, the art of assigning instruments to parts, is a larger issue. The main concern is to create a pleasant surface texture where the parts are balanced with each other and appropriately distinct or blended. Traditional practices must be reconsidered when working with the software orchestra. For example, it can handle the assignment of many similar instruments to one part (e.g., 16 violins), but one should recognize that in physical ensembles this is done for reasons of timbre and volume that we can more efficiently simulate with a chorus effect. There is no practical need, as there is in living ensembles, to have a fixed set of instruments and performers who are doing something with them most of the time. The variety of software instruments is great, and their individual timbres can be very flexible. Their placement, or even motion, around the listener and the acoustic qualities of the concert hall are now under our control as well. You might even say that even the listener's ears are under our control, since the orchestra provides two sophisticated output routines:

- *Transaural* output for potentially arbitrary loudspeaker arrays in a real room, with *Ambisonic* technology optimizing the accuracy with which the composer's positioning instructions are reproduced.
- *Binaural* output for headphones. This allows the system to add more psychoacoustic cues to the audio signal for greater realism.

These are issues of production, but composition is more intimately connected to arranging, so the relation of deep texture to the instrumental parts requires another look.

With a diagram, I present three examples of texture to give you a glimpse of the subject: a polyphonic consort (a string quartet); a homophonic one (a barbershop quartet); and a popular one (a basic rock band), which though simple, will nevertheless illustrate the kinds of compositional choices by which deep and surface textures are designed. The left column represents deep textures – the abstract pictures that will appear in the listener's mind. The middle columns show the *arrangement* (left) into instrumental parts (right) that "displays" the deep texture to the listener. The pipes connecting the two sides show which parts are most responsible for representing which elements of the deep texture. Rhythm and harmony generally have at least some bearing on every part, so these are omitted unless they show the main duty of the part. The right column represents the *orchestration* – the assignment of parts to instruments. To the right of this we might imagine the remaining details of performance, position, balance, and space, as *production* – the creation of the surface

Texture and arranging: fig.1.14, on the next page

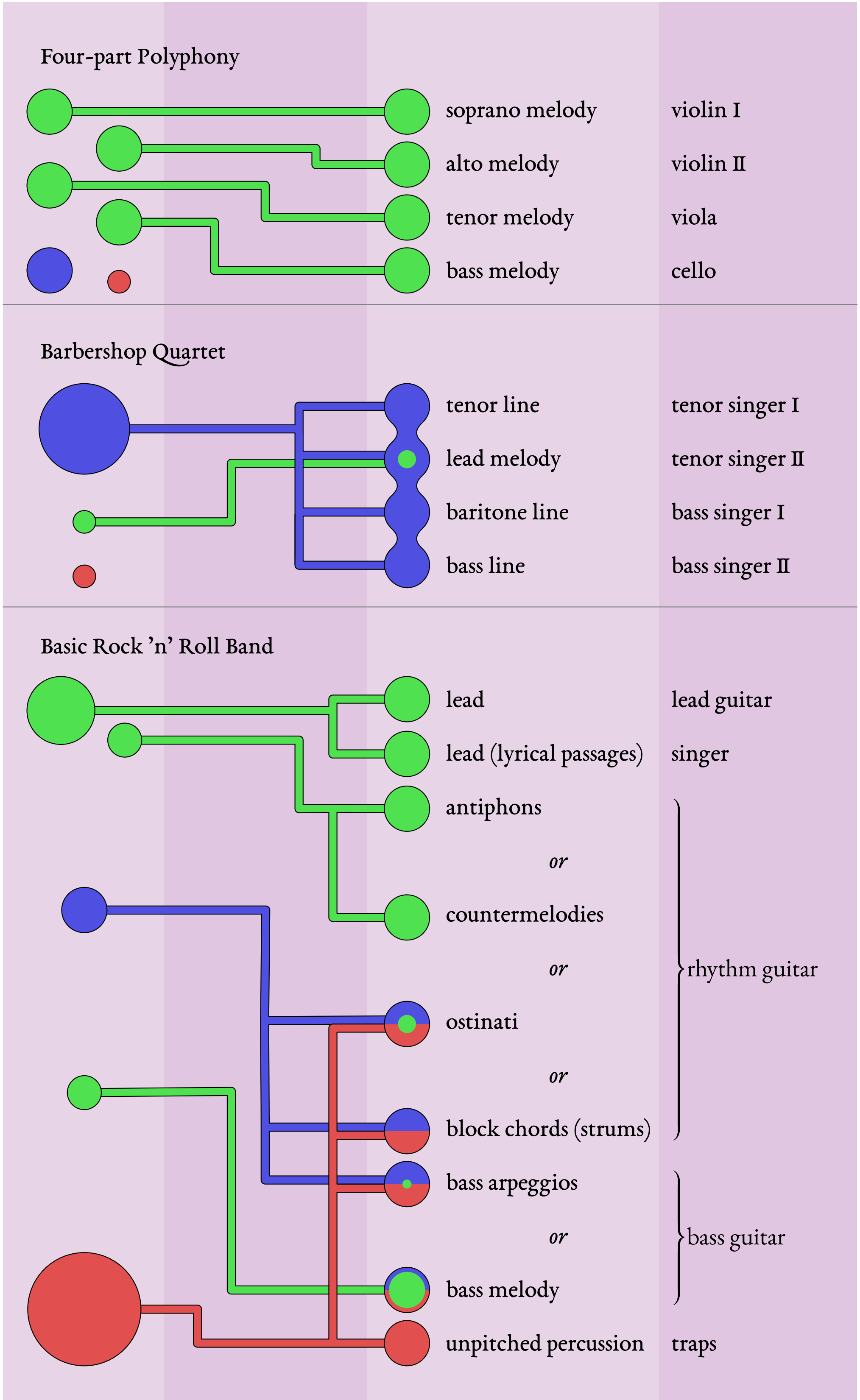

Figure 1.14: Three examples of texture and arranging (see text).

texture. When composing, opportunities and problems encountered in one column may well cause you to reconsider the others. Red, green, and blue represent rhythm, melody, and harmony respectively. The size of the dots represents their importance, and melodic parts also indicate their relative pitch level and degree of fusion. Generally, all parts have some rhythmic function and all pitched parts contribute to the perceived harmony; an instrumental part may simultaneously serve several purposes within an arrangement, and to differing degrees.

The first example is typical four-part polyphony (as may be generated by our harmony composition program with the criterion to maximize counterpoint). The four melodies are completely distinct, allotted to different pitch ranges, and of equal importance. Though listeners tend to be most aware of the top and bottom voices, the inner voices can be followed without effort, especially if they have different timbres and come from different directions in the listening space. In this example, the harmony is of equal importance to any melody, an approach that became common during the Baroque period of Western art music in the early 17th century. Before this, it would not have been unusual for a sequence of sonorities to have high counterpoint and low dissonance scores, but without any plan structuring the other flavors, or the appearance of certain chords at certain points. The large blue dot here implies that these factors are being considered. The small red dot indicates a simplicity of rhythm – the talea of the melodies might even be identical. That would tend to make the melodies more fused, but it was often done to make lyrics clear. The arrangement is trivial: the four melodies are each assigned a voice. This exactly represents the desired harmony and is sufficient for the rhythm – no other parts are needed. The choice to assign similar timbres to each voice (and, we might imagine, to sit them together on the stage) also works against counterpoint, but this is often done for reasons of balance and smooth surface texture.

"Polyphony" refers to simultaneous melodies of equal importance and "homophony" refers to textures where only one melody is important and the other parts exist mainly to support the harmony. They are often presented as the two main possibilities, but when you come to understand texture, you'll find terms like this too vague and broad to be useful. A simple example of homophony is a melody part accompanied by block chords (all the notes of a sonority simply played at once, perhaps with a guitar or keyboard instrument). Our second example, a style of vocal music called "barbershop quartet," could be called homophonic, but here the harmony is much more important than even the lead melody. The lead is really just an excuse for having a harmony. Though this music is not unmetered, rhythm plays very little part in. The arrangement consists of assigning the lead to a voice and filling out the harmony with three fused voices. Unusually, this tradition places the dominant melody second from the top and allows the third lowest voice to cross above it, though it more often sounds below (the genre's names for the parts don't exactly correspond to the traditional names of voice types). There is little effort to make the lead distinct from the other melodies, and none to make them distinct from each other, so they all end up fused. In general, any two melodies might be more-or-less fused – if we regard a group of simultaneous colors as a sequence of sonorities, this is measurable as a combination of their counterpoint scores. What can be measured can be composed, and a few genres (e.g., shoegaze) have played with the grey area of moderate counterpoint, but there are many games left to be played.

The final example demonstrates that even a small ensemble can quickly change its arrangement to represent a variety of deep textures with a variety of surface textures. Composing can involve back-and-forth considerations between the two. Let's begin with the forces of the ensemble, and work backward to their typical arranging resources and then to the possible deep textures they can represent. A typical rock'n'roll band has four pieces: "lead" and "rhythm" electric guitars (the same type of instrument but playing different roles in the arrangement), an electric bass guitar (a similar instrument, but pitched lower), and "traps" or "drum kit" (a set of drums and percussion instruments with one player); one or more of the guitarists typically doubles as a singer.

Any guitar may play a melody or strum block chords or play arpeggios or other ostinato figures (a skilled guitarist might also play polyphony, but this is rare in rock), and they may switch between these roles in an instant. An *ostinato* is a repeated figure, a pattern of notes that usually has a strong harmonic and rhythmic function, as well as some melodic interest. Its function is to play chords in a more interesting way that simply strumming block chords (e.g., the "Alberti bass": chord tones played on successive beats, in the order low–high–medium–high, over and over again, whatever the current chord happens to be). It may include simultaneous tones and, to heighten their melodic interest, some non-chord tones. If it consists of only chord tones, it is called an *arpeggio.*

How do these resources relate to possible arrangements? In the lower registers, chords tend to be dissonant, so the bass guitar rarely strums block chords – it is restricted to melody or arpeggio, underscoring the strong beats of the rhythm and supplying the bottom tones of important sonorities. The lead guitar usually sticks to playing the lead melody, though diversions like strums or other simultaneous notes may be played to accent this. The singer typically doubles the lead, at least for the passages with lyrics. The rhythm guitar is the most flexible. Its main function is to play the harmony and, secondarily, to enrich the rhythm; but, there are many ways it can do this. It might play block chords, or ostinati. It might play snatches of melody in counterpoint to the lead (*countermelody*) or it might even take over the lead for some passages to sound like a "response" (*antiphon*) to the lead's "call." The drummer's function, of course, is to realize the rhythm, but it needn't be simple or obvious (as it too often is). A rhythm can be thought of as multi-layered, with different combinations of metric level, rhythmic mode, and groove going to different instruments, particularly to the drums and the rhythm and bass guitars. The drum kit itself is capable of playing several voices, e.g., two rhythmic modes might play two renditions of the same meter: one assigned to the bass drum and toms, and the other to the snare and hi-hat.

Taking one more step leftward in the diagram, we can see how the ensemble might realize different deep textures. Usually, rhythm is prominent and many resources (traps, bass arpeggios, and rhythm guitar block chords or ostinati) are devoted to realizing it, and providing opportunities for it to be richly multi-layered. Harmony is usually carried by the rhythm guitar (as block chords or ostinati) with support by the bass. The lead melody is most often handled by the lead guitar, with a singer doubling it for the lyrical passages, though the rhythm guitar might take over for antiphons. The lead is the most prominent melody but there may be other, optional melodies of secondary importance played by the other guitars. The bass guitar may do more than play important chord tones on strong beats. It may play its own melody in counterpoint to the lead. The rhythm guitar may also take a break from its usual duties to play melody, probably but not necessarily in counterpoint with the lead (and a phrase that begins as a countermelody can easily continue as an antiphon when the lead part rests). In the diagram, I've represented these options alongside each other, and used multi-colored dots to represent parts with multiple functions. The melodic content that arises out of rhythm guitar ostinati or bass arpeggios does add melody elements to the perceived deep texture, but it is an artifact of the arrangement, so I don't show these connections to the secondary or bass melodies. Later, we will see a more rigorous language for formalizing deep texture and arrangement.

In addition to the resources of the orchestra, I'll supply software to aid in arranging and orchestration:

- A program for designing deep textures.
- A program for planning the parts to represent a texture and for assessing how well given parts represent it (whether they are insufficient or whether they have added elements to it).
- A program for composing accompaniments given a harmony and rhythm – block chords, arpeggios, ostinati, and other, non-repetitive accompaniments.
- A program for composing countermelodies, including the option of counterpoint, given the existing elements.
- A program for composing supporting rhythmic parts.
- A program for designing and balancing surface textures.

1.3.5 Form

Form in detail: ??, ??

The texture and all the other objects introduced so far may change in time over the course of a piece of music. The material itself probably changes, that is, we may hear a sequence of different melodies (or harmonies, or rhythms). This causes us continuous experiences of recognition and surprise, determined by similarities and contrasts between new elements and old. The music plays with our faculties of memory and expectation. As we hear a piece for the first time, we are constantly referring to our memories of its previous sections. By the end of the piece, our minds have constructed (probably unconsciously) a model of it. On subsequent auditions, we also refer to that model. On any audition, we are also referring to models of music we have heard before which have created expectations in us, and these are either being satisfied, thwarted, or transformed.

We call these models *forms*. Generally, they have several simultaneous logical levels, each dividing the duration differently, into blocks we can think of as containing the other musical objects. Most forms are hierarchical, like speech, with blocks similar in duration to words, phrases, sentences, and paragraphs (the musical names are cells, phrases, periods, and sections). A hierarchical structure carries two implications:

1. A time boundary on a high level is repeated on all lower levels (e.g., phrases do not cross section boundaries).
2. The qualities of a low-level block inherit, or are at least influenced by those of its higher, containing blocks (e.g., a phrase in a quiet section is likely to be quiet).

Heterarchical models are also possible – some pieces are better understood as several simultaneous divisions of the duration, neither really higher nor lower, but each applying to different musical objects (e.g., one a sequence of modes, another a sequence of meters, each independent of the others), without the boundaries needing to line up – i.e., ignoring the both implications above. Boundaries occur in models when we sense *some* change in the musical material or qualities, but it is not necessary for everything to change to create the sensation of a boundary, and the changes may be abrupt or gradual. A boundary may be as subtle as an elision between two contrasting divisions.

Anatomy of forms: fig.1.15, below

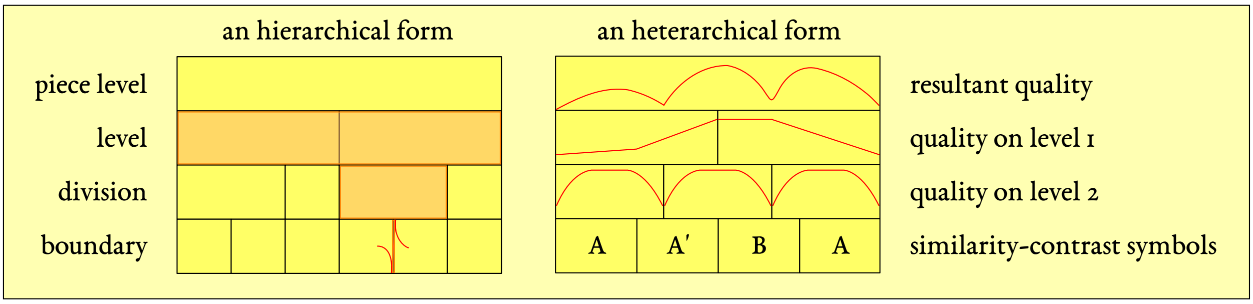

Figure 1.15: Two examples of form: time proceeds to the right, logical level downward. On the left, the structural parts of a form are highlighted. All forms have undivided "top" level for the whole piece. Simultaneous *formal division levels* (FDLS) divide the piece into *formal divisions* (FDS) separated by *formal division boundaries* (FDBS). Here, the fifth of seven boundaries on the lowest level is insisting that the change in some quality should be abrupt. On the right, some types of content are illustrated. Some quality (perhaps tempo) is defined differently on two different levels and a combination of these, a resultant, will be used in composition. *Similarity-contrast symbols* (SCSS) indicate repetition (A and A), variation (A′) and contrasting material (B) by the sorts of symbols familiar in poetry and music analysis. They may apply to the FDS in general, or to one aspect of the music (e.g., melodic content). In hierarchical forms, a boundary must be repeated in all lower levels, and every FD is contained by a higher FD. In heterarchical forms, this is not the case, and though the levels are numbered for convenient indexing, they are simply containers for different types of content.

This is how listeners experience form, but from our side, the composer's side, we would like to plan the experience using an object that mirrors it. For us, a form is a structured container that holds our plans and, eventually, their realization as instrumental parts – playable chunks of score. The form then helps us organize our intentions and, with software tools, becomes a workshop for realizing them. In the end, when all the parts have been composed, the form becomes the score.

As a container, a form can hold the following things in each FD:

1. A title and notes about the plan for the FD.
2. Duration information, which may be given in any units of time (seconds, pulses, measures, or beats).
3. A tempo, which may vary in time over the course of the FD.
4. A deep texture with an arrangement mapping its elements (DTEs) to instrumental parts (IPs).
5. Rhythms, which may be recorded as given taleæ, or as objects that generate taleæ:
    a) Metric schedules – sequences of meters, which may include loops, and may indicate definite or relative numbers of measures or repetitions.
    b) Grooves.
    c) Rhythmic modes.
    d) References to combinations of the above to indicate how a talea should be generated.
6. Melodies, which may be given as combinations of color and talea or as input parameters for use by melody generators or processors.
7. Harmonies, which may be given as taleæ, combined with either:
    a) sequences of sonorities (SOS), or
    b) sequences of chords (SOC).

    or as input parameters for use by harmony generators.
8. *Qualities* – values that may vary over the course of the FD, to be used as parameters by the orchestra or by compositional programs.
9. *Similarity-contrast symbols* (SCSs), indications of how similar this FD is to others on the same FDL, in general or for specific qualities, or types of material content.
10. The instrumental parts themselves, containers for score fragments.

...and FDBs may contain:

11. Indications for modifying one or all qualities at the boundary, to make them larger or smaller, smoother or more abrupt.

Qualities are a new idea here. They are simply any time-varying quantity that may be used later on, e.g., the relative loudness of a part (a "conducting channel" as used by the orchestra), or a parameter that might be passed on to a musical generator, like the desired tension used as input to a harmony generator. Notice that it is perfectly legal to specify conflicting instructions on two FDLs, for example, a quick passage might be indicated during a slow section. These conflicts can be solved automatically, producing a resultant object for the entire piece, e.g., a single, complex curve representing tempo. Similarly, there are easy mathematical solutions for conflicting specifications of duration (e.g., time given in seconds and pulses not agreeing with the tempo, etc.).

Form and its contents are a plan that we work with to produce a score. Some of these contents require further explanation. Deep textures are plans for a listener's experience, and their presence in the form acts as a reminder about what is to be composed. The arrangement is a plan for how to represent the deep texture elements (DTEs) with instrumental parts (IPs), a sort of to-do list. There are software tools for aiding this, but back-and-forth working methods are typical. For example, suppose your deep texture is simply one melody with a rhythm and harmony. Once you have written the melody and a bass arpeggio, a software tool may tell you that the harmony is insufficiently represented. You might then use another tool to compose an ostinato between the melody and bass to perform this function. Another tool may point out that the ostinato is in good counterpoint with the lead melody, so that, really, you have added a melody to the deep texture. If you don't like that idea, you might want to try again, composing an ostinato that is fused with the lead melody, or a different tool to compose block chords for the new part instead; and, if you liked the first ostinato itself, you might decide to use it elsewhere.

Similarity-contrast symbols are also plans reminding you how you intend to compose an element – as a repetition of, or variation on another element, or as something completely different. For example, you may want to repeat a melody from a previous FD, but with an entirely different harmony (reharmonization). After composing the new harmony, you might find (or maybe a software tool alerts you) that it's similar to the old one, prompting you to try again with different parameters.

The compositional process can proceed in all sorts of non-linear ways. You might begin without a form, simply composing or finding musical objects that intrigue you – a bit of folk melody you like, a traditional meter you read about, a harmony you generated one day – with no form in mind. You might allow a form to grow to accommodate the material you have at hand. Or, you might plan it from the top down, even with durations given literally in seconds (as you must when composing a soundtrack). But a form needn't be so specific. You might begin with one only partially filled-in, providing a sort of template for a given genre of music (e.g., a symphony or pop song), waiting for you to select some options, work out the details with compositional programs, and fill in the musical material. You may choose to use a familiar form and play along with your audience's expectations, or devise your own, effectively using patterns of contrast and similarity to teach the audience how to know the piece.

Just as notes are planned sounds, scores are planned pieces, and once all the parts are "filled in" with musical material, we have a complete score and we may move on to rehearsal. It is at this stage that final decisions can be made about orchestration and production. You can choose the instruments and performers to assign to parts. If you like you can tweak the phrasing, the timbral parameters of the instruments, and give special performance instructions (e.g., staccato bowing). If you disagree with a performer's interpretation, you can change any detail of a note's performance including the articulation between it and adjacent notes. You might adjust the balance between parts by telling the performers to play louder or softer with a conductor channel, or by placing them differently in relation to the virtual microphone, or by altering the acoustic properties of the virtual studio. You might even decide that they, or the listener might need to move around during the performance. Rehearsal might cause you to discard or change some parts, or to return to the form to reconsider the arrangement and compose new parts. For rehearsal, I recommend rendering the parts separately and assembling them together in a multi-track sound editor. This way is faster and allows them to be auditioned separately. It's also easy to try out ideas about dynamic level and timbre crudely with volume an equalization controls before re-rendering a part with different dynamics and timbral parameters. Finally, you render the parts in high quality (in your chosen format, e.g., stereo, 5.1, or binaural) and assemble them into a finished sound file. (And, of course, there is nothing to stop you from translating the score into traditional notation and giving it to fallible, difficult, human musicians.)

For working with forms, I'll supply:

- An editor for composing hierarchical or heterarchical forms and working them into finished scores, including integration with the previously mentioned composition and analysis programs.
- The orchestra itself is the main tool for rehearsal and final rendering.

1.4 The Estimation Layer

This minor layer is a sort of gatekeeper. When we hear music there is a very quick, subconscious, initial judgment as to whether we should attend to it and proceed to the next layer of processing, where enjoyment begins. If the gatekeeper judges that enjoyment is unlikely, however, it will not allocate the cognitive resources to engage with the music, and we will experience it as a more or less annoying distraction. Of course, even if our subconscious decides that music isn't worth attention, our conscious mind can override this and attempt to appreciate it. Overrides are rare. Don't count on your listeners to do this.

I think the judgment is an estimate of future pleasure that depends on two things (but there may be more that I don't know about): information flux and familiarity.

By "information flux" I mean the flow of information – in the sense of Information Theory *à la* Claude Shannon – out of the musical layer per unit time. The subconscious gatekeeper may balk at having too much to sort out but will certainly get bored if there is too little to bother with. I wish I could quantify this sort of information, but there are too many unanswered questions: How does the brain represent the musical objects? How do their interrelations affect this? Is the ear comfortable ignoring less important musical objects? How much does the listener's familiarity with an object reduce its information value? Luckily, it is not too difficult to get a feel for the limits – your experience as a listener will guide you.

Until the subject can be properly quantified, I offer some advice on information flux. It's harder to create too much rather than too little, so it's better to err on the side of high flux. Performance nuance is usually enjoyable, and is very unlikely to take you over the limit. The worst high-flux offenders are excessive complexity in rhythm and lengthy, meandering melodies that do not rest or at least pause on certain notes. Melodies that repeat some patterns are very welcome. Recall that only certain patterns of sound make it into the musical layer – when patterns barely make it in, they are a strain on the listener. The qualities and patterns of natural sounds, especially human motion and speech, are the easiest for listeners to process.

Keep in mind how objects in the deep texture are entangled: all melodies and harmonies have a rhythm and all the rhythms in a piece sum to a resultant rhythm; a harmony contains a number of melodies woven together; melodies may be relatively distinct from each other or fused together into an harmonic object. It is not difficult to distinguish three simultaneous melodies if they are kept distinct with counterpoint, but this quickly becomes difficult for larger numbers. Even if more than three melodies are distinct, they do not necessarily become taxing, as the listener will probably only attend to one or two at a time. Excessive dissonance, clashes of timbre, poor dynamic balance or spatial separation between parts can make listening more fatiguing.

Keeping the flux adequately high is easiest when you use all three main channels: an interesting rhythm, a harmony, and at least one melody. It's quite possible to get along with only two of these for a while (indeed, much of the world's music has little or no harmony). It's possible to get by with only one for a short time before the gatekeeper stops you. If you are trying to make music without any rhythm, melody, or harmony, you would be happier painting or writing poetry instead, or perhaps you should join the circus. Performance nuance and rich timbres are almost always welcome and can keep your flux from running too low (notice how frequently they are found in traditional solo genres).

As for the issue of familiarity, I'm sorry to report it. I would prefer that the gatekeeper loved novelty, but I know this is not the case. Perhaps familiar musical objects and combinations allow an ease of processing that the brain finds welcome. In any case, the gatekeeper appears to judge whether the musical elements are similar enough to those that have been enjoyed in the past. If so, they are passed on to the next layer. It may be that all the musical objects are passed, but it is entirely possible that this is done per element. For example, a listener from a culture where harmony is unfamiliar may ignore the harmony and only attend to the melody and rhythm. Or, if a listener is unfamiliar with simultaneous melodies, they may be heard as an unpleasant distraction, an overflow of information flux. If enough harmonic or contrapuntal music is auditioned, these listeners will develop familiarity and appreciation later. It's true that different musical cultures can be like different languages, but they are like languages one can pick up in an hour. If you attempt something far outside your audience's experience, the question is: will they give you an hour of attention?

In many cases, sufficient exposure to a type of musical element or pattern may make it familiar and lead to enjoyment. There is a related psychological phenomenon known as "mere familiarity." In the classic experiment, subjects are asked to rate random tone sequences. Sequences they have heard before (even if only subliminally) are rated better than new sequences. The nervous system seems to reward itself with pleasure when its pattern-matching system reports a success. (Cunning composers can take advantage of this by hiding one or more repetitions of a melody in the inner parts before it is ever heard in the more salient outer parts.) There is, however, no guarantee that exposure will cause appreciation. We have all had the experience of being exposed to some type of music for a long time without ever coming to enjoy it (e.g., pop music that inevitably fails on account of low information flux).

This is why I think the decision to attend to music depends mainly on whether it resembles past enjoyments. The pleasures of information reduction and mere recognition may be factors, but ultimately, the estimation layer is deciding whether future pleasure is probable. The judgment may be negative even if we choose culturally common features and our audience listens politely for a long time. The gatekeeper will not be bribed.

1.5 The Aesthetic Layer

If some objects from the lower layers (mainly the musical, but sometimes the auditory) are admitted to this layer, we will engage with them actively. We will project moods, emotions, imagery and other aesthetic feelings onto the objects. The projection may come from conscious efforts, or unbidden from our subconscious, so that it *seems* to be coming from the music. Either way, it is a creative activity coming from our side, not from the music itself. The musical objects act as a sort of blank screen ready to accept whatever projection is suitable for a particular listener. This is why, generally, no one disagrees about the content of the musical layer but it is common to disagree about the content of the aesthetic layer. So we may agree that such and such melody is present, but you think it means Paris in the spring while to me it evokes Chicago in winter (and we both love it). The sensation is much the same as when poetry succeeds in moving us, but because poetry is language it can hardly avoid representing specific things, while music is abstract and free to represent anything – its screens are blanker.

What are the causes of these projections? They can be divided into four categories: 1. Universal, 2. Cultural, 3. Personal, and 4. Capricious

A few aesthetic associations may be universal. For example, a crescendo (a passage increasing in loudness) is exciting because it may mean something is approaching – even animals know this. These responses may be hardwired in the brain or, surprisingly, learned identically by all healthy humans. For example, there is probably no built-in neural circuitry that identifies sounds with harmonic spectra, but because humans live on the surface of earth and are frequently exposed to human speech, they hear sounds of this description all the time and become quite adept at processing them. Whether a universal projection proceeds from congenital neurology or from a universally shared experience need not concern us. In any case, there probably aren't many universal aesthetic projections. A catalogue of them would be useful, but I don't have one yet.

Some aesthetic responses are learned, but only within a particular musical culture. This sometimes takes advantage of universals. For example, when the pitch of a speaker's voice is lowered from its usual level, this is usually interpreted as indicating a negative emotion like sadness. In Western culture, we most often hear the major scale, but sometimes the minor. Relative to the major, the minor has some prominent degrees lowered, and so we tend to hear it as indicating a negative affect. In other cultures, this is not necessarily the case. On the other hand, completely arbitrary associations between musical feature and aesthetic affect may be promulgated within a musical culture. For example, the very specific associations of particular ragas in Indian music, are probably arbitrary signs accessible only to encultured listeners, in much the same way that words in a language arbitrarily indicate objects, but only for speakers of the language. These effects may be potent to connoisseurs, but unavailable to outsiders. The size of the culture may be quite small or recent – for example, the cliches that have developed in film music, like strings indicating love, or the whole tone scale indicating unease, or simple electronic sounds indicating an alien presence. There are countless subcultures that teach countless associations, and it would be interesting to see a catalogue of them, but keep in mind their limited strength and scope: they will only be as effective as your listeners have been conditioned.

Some aesthetic responses are learned, but only by an individual. They may evoke the time or place where a musical element became familiar or the emotions the listener was feeling at that time. These associations are habits of the mind, and so vary widely in strength and stability. They can wax or wane. They can be unlearned or reinforced by time and listening habits. They can change a great deal over the life of the listener. Personal associations are generally unknowable to others, including the composer.

Finally, since aesthetic projections occur in the flux of mental activity, they are subject to the influence of whatever else is going on in the mind at the time – current, moods, emotions, circumstances, trains of thought, fleeting memories, sensations, and so on. All the mental contents interact in some unpredictable way to give us the aesthetic experience of a particular audition. This is how different auditions of the very same recording of the very same piece can give us very different experiences. At one time the music seems transcendent, at another, trite, at another, irritating. The aesthetic experience depends upon the projector, the listener, not upon the screen, the music. These capricious responses can hardly be predicted by the listener, let alone by the composer.

Can we cause listeners to project as we wish? When it comes to universal projections, yes, but the palette of affects is extremely limited. If we try to trigger a culturally-learned projection, the palette can be very wide, but the demographic scope is more or less limited, and the strength is relatively weak and unreliable. If we attempt to activate an individual's learned response, the effect can be strong but these associations are probably unknown to us unless we have taught the audience ourselves by associating musical elements with extramusical signs – as happens with leitmotifs in film scores or with the titles of programme music. To control capricious projections would be sublime but is impossible for the composer – the rare listener might attempt it as an act of will, but I would not propose it. So, generally, no, we cannot "call our shots" with music, but if we wish to try, plausible strategies are:

1. Take advantage of universal associations where they align with your aesthetic goals.

2. Take advantage of cultural associations if you think they will work with your audience.
3. Set up associations, if possible, with accompanying extramusical material – titles, program notes, cover art, theatrical presentation (especially if it is repetitious), etc. In the extreme case, you may be so foolish as to tell them what they are supposed to think of it.

1.6 The Hedonic Layer

The preceding operations may trigger the brain to experience various forms of pleasure (and displeasure) by various channels. I have already mentioned that the brain experiences anxiety when confronted with indecipherable stimuli (in the auditory layer), and rewards itself with pleasure for successful pattern matching (in the musical layer). A similar reward probably results from a successful unweaving of a polyphonic texture into simultaneous melodies. In the musical layer, many composers adopt a strategy of causing pleasure by satisfying expectations, which they often intensify with delays. Many hedonic (and anti-hedonic) effects can occur in the aesthetic layer. For example, the evocation of negative emotions can, via catharsis, induce pleasure (the brain, tricked into thinking it is suffering may release what are essentially opiates that, since one is not actually suffering, elevate the mood). Or perhaps the evocation of an unpleasant memory simply induces displeasure.

This is an area of ongoing study – scientists attempt to correlate musical experiences with reported introspections, neural imaging (e.g., music can activate pleasure centers of the brain associated with food, sex, and drugs), and measurements of biochemical messengers (i.e., hormones and neurotransmitters such as dopamine and serotonin, which are associated with various emotions). The number of relevant stimuli is large. Several mechanisms, "hedonic channels," difficult to disentangle, are at play, causing different kinds of pleasure related to different biochemical messengers or different patterns of neural activity. I don't have a clear map of how this layer works, except to say that the objects from preceding layers pass through this one acquiring hedonic charges and colorings from the various mechanisms. I don't know the full palette of possible charges, or how to quantify them, or how to cause them reliably.

In any case, the auditory and musical objects probably avoid displeasure by their clarity; the musical objects may cause some pleasure by their richness and familiarity (the rewards of decipherment, pattern matching, and satisfied expectation); and the aesthetic objects probably create pleasure according to how welcome they are psychologically. I think the bulk of hedonic experiences come from the aesthetic objects for psychological reasons, and these themselves are difficult to predict. It will be interesting to see what science tells us about this layer in the future.

1.7 The Intellectual Layer

This final layer is the conscious, intellectual understanding of the music we have experienced, a considered judgment. Here we combine the products of the previous layers into a single, complex object and, through memory, integrate it with our understanding of music as a whole – with all of our previous experiences of music. Every part of the psyche can influence this understanding, including our self-image and our attitudes concerning all other subjects. It is here that musical experience becomes welded to extra-musical ideas. This can happen so consciously that it can be put into words. This is the layer in which people function when they write music reviews. Of utmost importance to working musicians, this is also the layer that is active when people go shopping for music.

The intellectual layer is where we consciously decide what we like. Since all our other opinions may be involved, it is not simply a measure of how much we enjoyed certain pieces. Our considered judgment may depend more on who we think we are and what we think is good than on our actual experience. What we think we like isn't necessarily what we enjoy the most, but what we habitually tell ourselves we like. For most people, I think the strongest factors are social and extra-musical and that, among the purely musical factors, the strongest are objects from the aesthetic layer, colored and charged by the hedonic layer. Objects from the musical and auditory layers are only factors for real musicians, the people who know and understand those layers to such a degree that they can describe them in words. Incidentally, this explains why most music writing is so bad: most writers, ignorant of the intricacies of the lower layers, speak only about their own idiosyncratic aesthetic projections that the reader probably doesn't share – why should anyone else care?

I think that the results of this layer depend so heavily on social, cultural, and individual psychological factors that using our musical knowledge to make an excellent product will only put our foot in the door, not make the final sale. That will depend more on having a haircut the customer admires. However, if we are concerned more with artistic success than popularity, knowledge of the lower layers is essential.

1.8 Strategies

Given this seven-layered model of musical experience, what are the best strategies for music-making? Notice that the lower layers (acoustic, auditory, and musical) are universal – their operation is essentially the same for all healthy humans. The higher layers (from the estimation to the intellectual) may have some universal elements, but generally, their operation varies from one listener to another. To be effective on the higher layers, you must know your audience. All good strategies must first be effective on the lower layers, then you have a choice about how to address the higher ones.

Two goals apply to all listeners and must be common to all strategies:

1. Get all the material out of the acoustic layer and into the auditory.
2. Get all of this material out of the auditory layer and into the musical.

In other words: make every sound count. People attempting music-making without training often fail to achieve this and end up making noise or ineffective jumbles of sound. Some people have some success by merely imitating others, but then they have no real control over their art. People with training may succeed in these goals, but usually by following the rules of one narrow tradition. They may have a little control, but very little freedom. I want you to learn everything about how the auditory and musical layers work so that you have complete artistic freedom. I want to give you the power to exploit everything the ear savors.

The first problem is to make the acoustic signal beautiful in itself, to pick up whatever hedonic charge we can even at this most basic level. The most effective way to do this is to give it the qualities of human motion and speech, by performance nuance and possibly even by timbre. The orchestra will do this for you automatically.

The acoustic layer and our orchestra: ??, ??

The second problem is to make sure everything in the acoustic signal gets easily and unambiguously parsed into auditory objects (sounds, streams, sources, and scene). Unparsed material (e.g., noise, dissonance, confused auditory scenes) makes the brain anxious about possible threats in the environment, and carries a potent anti-hedonic charge. It's possible to use this briefly as a special effect that draws attention to some part of the music, but it's something you should avoid most of the time. Hedonic charge can be gained here by setting a clear and interesting auditory scene, a pleasant surface texture.

The auditory layer and surface texture: ??, ??

The third problem is to make sure that all the auditory objects contribute to the creation of your desired musical objects in the musical layer. Recall that only certain types of sounds in certain patterns can cause the sensations of rhythm, melody, harmony, and so on. Material that does not rise to higher layers cannot result in significant enjoyment and is not useful artistically. To create musical objects in listeners' minds, you must order the sounds in certain ways. The bundled software will help to keep you on track, by designing valid musical objects and realizing them as playable scores. Don't despair – the range of validity is quite large. Even without knowing the audience, some hedonic charge can be gained here by:

The musical layer: ??–??, ??–??

1. Providing rich, challenging deep textures (e.g., polyphony, multi-layered and nuanced rhythm) or other objects (e.g., scales, melodic and harmonic modes) to create the pleasure of engagement.
2. Setting up and satisfying your own expectations by repeating musical objects, creating the pleasures of pattern-matching. This can be enhanced by delaying the satisfaction (of, for example, a repeated melodic phrase) by intervening material.

Beyond this, you face a decision about whether you will write for a target audience (that you must know well), or for an unknown, general audience. Depending on this decision, I offer you two extreme and two intermediate strategies:

Generic: restrict yourself to a genre and exploit its expectations.

Original: produce music that is interesting without concern for satisfying cultural expectations.

Compromised: do a bit of both (combine some predictability with some creativity).

Mixed: use more than one genre (mix and match familiar elements from different sources to create composite genres).

The generic strategy risks several pitfalls. You are restricted to the target audience and you can't expect success outside it. It is necessary to know the audience's true preferences and expectations very, very well. This is more difficult than you might expect, and you might end up working hard to satisfy expectations that don't exist, resulting in an ineffective product. For example, it is a common assumption that all Westerners have tonal harmony deeply ingrained in their ears, but I think this is only the case for the tiny fraction who have had formal training. Finally, if you take this path, you will inevitably be producing a product similar to existing products, so you run a huge risk of being boring and clichéd. You will always be trying to squeeze between the Scylla and Charybdis of familiarity and novelty, of satisfying expectations or giving interest. And remember: the similar products that already exist are already familiar to the audience and are already preferred to yours on that account – you will always be working against this disadvantage (until you become a legendary superstar). If you choose the generic course:

1. Observe the general strategy.
2. Create only musical objects of a sort familiar to your audience so that they make it past the gatekeeper.
3. Choose musical objects that have the aesthetic associations (for your audience) that you want to evoke.
4. The hedonics in this case are all based on familiarity, so they require no special attention.
5. If you care about how your audience thinks of you within this genre (i.e, their experience of your music on the intellectual layer), try to position yourself in some niche in relation to its established composers. This means concerning yourself with tiny distinctions of similarity and contrast between you and other musicians.

I confess that the original strategy, not aiming at a particular audience, is far more interesting to me. The pitfalls here are:

1. Not making it past the gatekeeper for want of familiarity.
2. Creating variable or ambiguous aesthetic experiences.
3. Lacking the hedonic resource of familiarity.
4. Listeners not knowing what to make of you in their considered judgments (in the intellectual layer).

The first obstacle is not insurmountable. All listeners must give a piece some chance for a few seconds, and mature listeners give a bit more – there *are* some open-minded people in the world. Also, consider that your chances are better than most. If you have observed the general strategy, you won't be delivering noise or an incomprehensible mess. You'll be presenting the listener with valid musical objects, things that *can* be understood and can quickly become familiar, perhaps during the course of a single piece. The second pitfall is actually the charm of the strategy. Since musical objects are like a blank screen for the aesthetic layer to project upon, a new combination of unfamiliar objects may well receive interesting new projections – it's fun to put strange screens into the world and see what results you get. The best strategy, then, is to keep the information flux high (give them a lot of screen, for larger, richer projections), and give them a variety of different screens (to offer a variety of new aesthetic experiences – something will stick). The third pitfall is the price we pay, but we can still generate hedonic charge by setting up our own expectations and satisfying them with repetition. The last pitfall is an issue that is rarely in our hands anyway. If you choose the original course:

1. Observe the general strategy.
2. Create whatever sorts of musical objects are piquing your curiosity at the moment.
3. Don't worry about their aesthetic associations (you can still use the universal associations, if you wish).
4. Use repetition and formal similarity to create hedonic charges, if you like.
5. Don't worry about what anyone else thinks. Instead, enjoy the delight of finding out what your musical ideas sound like to you.

The compromised strategy is to use some musical elements from an established genre, but be original in other elements. Some hedonic charge and aesthetic associations will still be available. If it works, your target audience will feel that you have extended their genre into new and exciting territory. It will fail, however, if they don't recognize the generic similarities, or if they take your music as a failed attempt to perform within the genre. The key is in the balance between generic and original elements and in not choosing original elements that alienate the audience.

The mixed strategy, combining elements from more than one genre, depends on the target audience. If they are familiar with both genres, a full range of hedonic charges and aesthetic associations are available to you. If they are familiar with only one of the genres, then it is equivalent to the compromised strategy; and, if they are familiar with neither, the original strategy. You are looking for a target audience that not only knows, but likes both genres, but even then, two pitfalls remain. You may well loose listeners who have a stronger allegiance to one or the other genre – they may come away feeling that their favorite has been adulterated or done badly. You may also end up sending mixed signals by trying to use two different sets of aesthetic associations.

1.9 The Music-making Method

I have so far described the musical experience from the listener's point of view, from the lowest layer to the highest. As a composer, you will be working from the top down, starting with a strategy, considering desired effects in the higher layers, and finally delivering an acoustic signal. To summarize then, here is the outermost loop of the composer's algorithm:

1. Decide on your strategy:
   (a) generic, (b) original, (c) compromised, or (d) mixed.
2. Consider how you'd like the audience to think of you in relation to other composers and their music:
   Generic: Position yourself within the genre in relation to its artists and typical aesthetics. Perhaps you want the texture of one, the rhythm of another, but a mood of your own; perhaps you can intensify the usual emotions beyond what has been done before.
   Original: Don't worry about exactly what the audience will think, but strive to give them a lot to think about.
   Compromised: Consider the typical aesthetics and musical elements of the genre. You might contradict the aesthetics or blend them with others, extending the genre by abandoning some of its typical musical elements in favor of new ones.
   Mixed: Between the two genres, consider which aesthetics to keep and which to abandon. This will determine which musical elements you use, but for now concentrate on which elements are most typical of each, which are most attractive to you, and try to balance the mixture as you would like it.
3. Consider how to create hedonic charges:
   Generic, Mixed, & the Generic part of Compromised: Much of the pleasurable quality of your music will come from the audience's recognition of familiar elements, but you can still add pleasure through any of the other means.

Original & the Original part of Compromised: Since you will lack the pleasures of familiarity, you must pay special attention to adding hedonic charge by other means.

All strategies: Avoid noisy or confusing surface textures, but make the texture rather rich, if your genre constraints allow it. Consider using polyphony if it is within your rules. Whether or not you are leaning on generic elements, you can set up and satisfy your own expectations with a judicious use of form by repeating material, and especially by having familiar material return after an absence.

4. Consider the aesthetic experience you want to the audience to have:

   Generic: Select aesthetic associations from your genre's palette.

   Original: Allow novel aesthetic associations to arise in the audience (though you can use universal associations, if you wish).

   Compromised: Select the aesthetics you wish from your redacted palette, while allowing room for them to blend with new and ambiguous effects. Accept the air of mystery that arises.

   Mixed: Having two half-palettes to choose from allows you different kinds of play. In a composite genre, a particular musical element with meaning in only one of the genres will cause a clear aesthetic sign. If it has meanings in both genres, you can create contradiction (if they are opposite), or ambiguity (if they are different). If a given aesthetic effect can be caused by a different elements in each genre, you have a sort of musical synonym to play with – you could possibly double the efficacy of the association by using both at once.

5. Consider how familiarity will cause people to listen:

   Generic: By keeping within a genre, your acceptance by the genre's audience is assured; however, within this arena you will be competing for their attention with other artists and works they already know. You are using general familiarity to contend, but you will be outgunned by specific familiarity. A thorough understanding of the lower layers might give you an advantage.

   Original: You are depending on an open-minded audience to give you a chance. You must gain their attention quickly, so make an interesting introduction. It may take several attempts with different combinations of musical elements before you get a positive audience response. So, let your goal be the thrill of playing with new elements and combinations and finding out how *you* like them. When you strike on something that works well, you can develop it into a genre of your own. What a joy music would be if we could invent a new genre every Monday morning.

Compromised: Your half-compliance with genre should get people to listen, and your original elements may either repel or captivate them. The attitude should otherwise be the same as the *Original* strategy of this step, except that you may be seeking to extend a genre rather than create a wholly new one.

Mixed: A composite genre has the potential to attract an audience corresponding to the union of its sources, but only an intersection of them is guaranteed. The aesthetic elements you select may either enhance or contradict the two sides. Enhancements can widen the appeal to one side or both (e.g., punk may energize an already energetic folk dance for folkies while charming punks with its more sophisticated melodies). Contradictions – conflicting aesthetics – may either displease the audiences or create a new interest. Experimentation will tell.

6. Plan to get the information flux within acceptable limits:

   Generic: Familiarity may make higher fluxes possible (as a sort of data compression goes on in the audience's mind) and lower fluxes more in danger of running under the limit. If you are used to the genre, take your own ear as the guide.

   Original: It is probably good to keep the flux high to give the audience more room to project, but unfamiliarity may make the music more difficult to parse, so beware of this. Remember that, having worked with and rehearsed a piece, you are much more familiar with it than is an audience hearing it for the first time.

   Compromised: The above phenomena (under *Generic & Original*) will tend to mitigate each other in this case.

   Mixed: For the intersection audience, this case is the same as *Generic* strategy. For the rest of the audience (those familiar with only one of the genres), it is like the *Compromised* strategy.

7. Choose the musical elements you will use:

   Generic: Your genre dictates what possibilities are available, and you select from them according to your aesthetic (and, secondarily, hedonic) goals, minding the flux.

Original: Play with any musical ideas that interest you. Discover them, and the projections they suggest to *you*. Enjoy the game.
Compromised: Satisfy the constraints of your half-genre as in the *Generic* strategy, then play with the rest as in the *Original* strategy.
Mixed: You have chosen a composite set of possibilities from the two genres. Select from them as in the *Generic* strategy.
All strategies: You may select elements in any order. Your ideas may begin anywhere – with aesthetic or hedonic goals (if you have them), or with the elements themselves. E.g., the idea of a rhythm might lead to a texture, or vice versa. You might include several elements of the same type for the sake of contrast (e.g., two different harmonic modes), especially if the piece is to be long. When you have enough materials for the piece, proceed.

8. Put the elements into order using form and textures:
A piece will need at least one texture. A piece of any length, will need some sort of form. Fill out the form so that there is always a deep texture of some kind, and fill out the textures with every element they require. Where melodies are simultaneous, decide whether you want them to be relatively distinct (i.e., in polyphony via counterpoint) or fused together (into a harmonic part, as in homophonic accompaniment). Consider how elements of the texture relate to one another, whether some are more important and prominent or subordinate and subtle. You may make arranging and orchestration decisions now, but be willing to revise them later. Consider how the blocks of the form (FDs) relate to one another, whether their contrasts and similarities are as you wish. Pay special attention to the boundaries between blocks (FDBs): you may want some to be abrupt and dramatic or elided and subtle. You can use orchestration as well as material and texture to differentiate sections. Consider the hedonic charges you may achieve with the textures and form. You may fill out the form and textures in any order you wish with newly composed or existing (but not plagiarized) elements. You may wish to change existing elements to fit with what you have chosen, e.g., altering a melody to fit your rhythm or recasting it into your melodic mode. You may well begin with basic elements and choose the textures and form last. In any case, once you have a fully filled-out form, proceed.

9. Reduce this material to a score:
   Consider the arrangement. Distribute the deep texture to instrumental parts of the surface texture. You may need to compose percussion parts to fully express the rhythm, if the melodies do not express it sufficiently. You may need to compose accompaniment parts to support the harmony for the same reason. These may be unmelodic, or they may be as involved as counter-melodies which may cause you to reconsider the deep texture. Consider the orchestration. Assign the parts to instruments: maintain the relationships of prominence and subordination between the parts as you have chosen; remember the degrees of distinctness and fusion you want between the parts; keep the surface texture beautiful, balanced, and clear. Consider the performance and spatial aspects. Assign the parts to performers with respect to their role in the texture (e.g., soloists to lead parts, choristers to fused homophonic parts, consortists to distinct polyphonic parts). Finish setting the auditory scene by stationing the instruments (or moving them about) in an appropriate acoustic space. All these decisions may be made in any order. When all is ready, reduce the form to a score.
10. Rehearse the score:
   Go over the score with particular attention to articulation and phrasing – these are performance details that your ear may have strong opinions about. Render it with the orchestra. It is quickest to do this part-by-part, and it allows careful auditioning. You may want to make some adjustments like the relative loudness of parts, details of tempo and timing, position and timbres of instruments. The orchestra has mechanisms for all of this. It even has mechanisms whereby you can change the slightest detail in the performance of a single note. This will be extremely rare however – the virtual performers always make excellent or, at lest, good decisions. Repeat this step until you are satisfied with the result
11. Deliver your piece:
   Render it in high quality in your preferred output format: either binaural for headphones or Ambisonic transaural for your desired speaker array – stereo or something more exotic. The exact specifications of the rendering will depend upon the target medium.

The rest of this book is concerned mainly with describing how the auditory and musical layers work (the orchestra can handle the acoustic layer without needing your attention, but I'll provide a user's manual in case you want to dig into those details or – bless you – extend it). The other layers are beyond our scope, except as mentioned above. My task is to lead you to the gatekeeper. The rest is up to you. Now let us set about exploiting all that the ear savors.