Dynamic Intonation for Synthesizer Performance

Benjamin Frederick Denckla

A.B., Special Concentration in
Experimental Live Electronic and Computer Music,
Harvard College, 1995

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences
at the Massachusetts Institute of Technology

September 1997

©Massachusetts Institute of Technology, 1997
All rights reserved
Dynamic Intonation for Synthesizer Performance

Benjamin Frederick Denckla

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on August 8, 1997,
in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences

Abstract

By default, all MIDI synthesizers are tuned to 12-tone equal temperament (12TET). This is the most convenient tuning because it is applicable to all Western music and can be controlled from conventional keyboards. Although it is convenient, it is not necessarily musically desirable. For example, harmonically speaking, many musicians find its major thirds less consonant than they would like. Fortunately, MIDI synthesizers are only tuned to 12TET by default, for convenience’s sake. They can in fact be used to realize pieces with each note tuned arbitrarily. Thus, like the violin or the voice, they do not have to conform to a tuning at all. This thesis investigates how this dynamic intonation capability can be applied to the realization of pieces of Western music. It investigates dynamic intonation from two perspectives. The first perspective is theoretical, presenting and evaluating a variety of alternatives to 12TET. The second perspective is practical, presenting software that was written in order to allow dynamic intonation on MIDI synthesizers. This software can be used to create MIDI files with dynamic intonation. The software can also allow a conventional MIDI keyboard to be used to perform a piece using dynamic intonation. It does so by following along in a score of the piece that has been annotated with intonation information, transmitting retuned versions of the notes it receives from the keyboard.

Thesis Supervisor: Tod Machover, Associate Professor of Music and Media

This research was sponsored by Microsoft and the Things That Think Consortium.
Dynamic Intonation for Synthesizer Performance

Benjamin Frederick Denckla

The following people served as readers for this thesis:

Michael Hawley
Assistant Professor of Media Arts and Sciences
Program in Media Arts and Sciences

Allen Strange
Professor of Music
San Jose State University
Acknowledgments

Thanks to Will “Billy-Dee” Oliver, for his scrutiny of drafts of this work.

Thanks to Teresa Marrin, for being a supportive and fun office-mate with whom I had many stimulating discussions about this work.

Thanks to Bernd Schoner, for reading a draft of this work and providing stimulating opposition to some of the ideas within.

Thanks to my advisor, Tod Machover, for allowing me to pursue a topic that is only tangentially related to our group’s work.

Thanks to Graeme Boone for being a generally inspiring musicological force and, specifically, for the pre-publication draft of his article on intonation in Dufay.

Thanks to Bill Sethares for pointers to important references.

Thanks to my first theory teacher, the composer Sotireos Vlahapolous, who introduced me to the wonderfully odd subject of tuning.

Colophon / Software Acknowledgments

This thesis was typeset using Christian Schenk’s MiKTeX, a port of Leslie Lamport’s LATEX and related utilities. LATEX is a set of extensions to Donald Knuth’s TeX. The font is Computer Modern, designed by Donald Knuth. Musical examples were typeset using Daniel Taupin’s MusicTeX extensions to TeX. The musical font is musikn, designed by Angelika Schofer and Andrea Steinbach.

Contacting the Author

The author can be contacted by email at bdenckla@media.mit.edu.


## Contents

1 **Introduction** 9
  1.1 The Role of Intonation in Music 9
  1.2 Intonation Theory 10
  1.3 Digital Instruments and Intonation 10
  1.4 Overview 11
  1.5 Document Conventions 12
  1.6 Review of Key Background Topics 12
    1.6.1 Pitch Names 12
    1.6.2 Interval Names and Abbreviations 12
    1.6.3 The Harmonic Model of Tones 14
    1.6.4 Frequency Ratios as Doublings 14
    1.6.5 Modulo and Integer Division 15

2 **Related Work** 16
  2.1 Tuning Theory 16
    2.1.1 Helmholtz and Ellis 16
    2.1.2 Blackwood 16
    2.1.3 Lindley and Turner-Smith 17
    2.1.4 Regener 17
  2.2 Dynamic Intonation Systems 18
    2.2.1 Sethares 18
    2.2.2 Waage 18
    2.2.3 McCoskey (FasTrak) 18
    2.2.4 Justonic, Inc. 19

3 **Tuning Theory** 20
  3.1 Models of Tunings 20
    3.1.1 Absolute Tunings 20
    3.1.2 Relative Tunings 20
    3.1.3 Transposable Tunings 21
    3.1.4 Register-Doubling Tunings 22
  3.2 The Choice of a Model for Diatonic Tunings 22
  3.3 Diatonic Pitches as Fifth-Register Vectors 23
  3.4 Converting Pitches to FRV 24
Chapter 1

Introduction

1.1 The Role of Intonation in Music

Intonation is the way frequencies are assigned to individual notes in the performance of a piece of music. Intonation must operate within a loose set of bounds that guarantee that the basic meaning of a piece will be preserved [3, 193–215], [14, 37], [20, 136–49]. For example, a minor third must be smaller than a major third. Within these bounds, intonation provides considerable expressive possibilities. The role of intonation in performing music is analogous to the role of pronunciation in reading poetry. Pronunciation operates within bounds that guarantee that the basic meaning of the poem will be preserved: for example, “mole” must be pronounced distinctly from “male.” Within these bounds, there is room for the speaker’s accent, vocal quality, and emotional expression.

What are the expressive possibilities of intonation? Here are a few. Intonation can be used to “color” or give mood to music. Historically, there has been much interest in the way intonation can give different moods to music in different keys [24], for instance through the use of a class of tunings called well temperaments. Although the mood that intonation imparts does not “happen” at any given time, intonation does influence the meaning of specific musical events. An important melodic example concerns upward steps from pitches that function as leading tones. The higher the intonation of this leading tone, the more a pull towards the pitch above it is felt. Since the pitch above it typically functions as some kind of local tonic, this pull can help enhance cadences and establish tonality [17, 63]. Low leading tones tend to give a “dull” sound to the music. Intonation also has a profound effect on harmonic events. It can determine just how dissonant a dissonant harmony is, and just how consonant a consonant harmony is. Extreme consonances can be useful to promote a feeling of tranquility and resolution in music, and can serve to heighten contrasts with dissonances [12, 331].
1.2 Intonation Theory

Intonation has only been theorized extensively with respect to the role that it can play in maximizing harmonic consonance. The role of intonation in harmonic consonance is indeed important, but the degree of theoretical attention that has been focused on it is also due to the fact that intonation’s other roles (key coloration, melodic meaning) are far more difficult to theorize about.

In addition to being focused on harmonic consonance, intonation theory has largely focused on static intonation. In static intonation, all of the notes in a piece that have the same pitch have the same frequency. Therefore static intonation conforms to a tuning, an assignment of frequencies to pitches. For example, “do you like to play your middle C in measure 97 higher than in other measures?” is a question about intonation whereas “what frequency does middle C have?” is a question about tuning. Instruments that must conform to a tuning, like the piano and organ, have static intonation. Thus it would not make sense to ask a pianist “do you like play your middle C in measure 97 higher than in other measures?” Although most pianists would not be sympathetic to the request, it does make sense to ask, “can we tune your middle C higher, so that it sounds better in the context of measure 97, but not so high that has an adverse effect in other measures?” Instruments like the violin and voice do not have to conform to a tuning and therefore can have dynamic intonation.

Static intonation (tuning) theory has had extensive practical application. The tuning and/or construction of all instruments with static intonation is based on the results of this theory. In contrast, dynamic intonation theory has not been put to practical use [17], [2], [11]. In part this is because few performers have the skill to make such explicit intonation decisions worthwhile. For most modern performers, it more pragmatic to try to adhere approximately to static 12TET intonation most of the time, making on-the-fly adjustments by ear where necessary. Another reason why theories of dynamic intonation have not been put to use is that many theorists have focused on just intonation, a type of intonation that has maximally consonant harmonies but cannot accomodate the rich harmonic vocabulary of most Western music.

1.3 Digital Instruments and Intonation

Digital electronic instruments allow the expressive possibilities of intonation to be explored in a manner more accurate and more convenient than ever before. Like the violin or voice, they are capable of producing a continuous range of frequencies and therefore do not have to conform to a tuning. Unlike acoustic instruments with continuous frequency, their frequency can be controlled extremely accurately and consistently. Perhaps most importantly, digital electronic instruments can be controlled by a computer. This enables them to have at least two modes of operation that would be impossible with acoustic instruments. In the first mode of operation, an entire piece, along with intonation annotations, can be specified in a digital score and then played by the computer.
on an instrument with a MIDI interface.

In the second mode of operation, the computer does not replace the performer but rather extends his intonation capabilities. The computer "listens" to the performance data from a MIDI keyboard, retunes it according to some algorithm or score, and re-transmits it to a MIDI instrument, all in real-time. This creates a new instrument in which the performer is not burdened with the responsibility of controlling intonation, but his performance can still have arbitrarily complex intonation. Not only does this make dynamic intonation possible, it also makes any static intonation possible. Normally, a keyboard can only be used to realize the most restricted subset of static intonations, those that tune all pitches separated by a diminished second (like G♯ and A♭) to the same frequency.

This thesis presents a software suite called Helm that unleashes the intonation capabilities of MIDI instruments. These tools allow intonation-annotated scores to be played by the computer or used as instructions for the real-time retuning of a performer's keyboard input. In addition, this thesis presents a theory of intonation that can be applied to the use of Helm. This theory, like most intonation theories, focuses on the role of intonation in maximizing harmonic consonance and treats dynamic intonation mainly with respect to just intonation. Within this conventionally narrow focus, it overcomes many other theories' weaknesses by introducing some new ideas and synthesizing the ideas of various theorists in new ways.

1.4 Overview

This section will give a brief overview of the rest of the thesis.

After introducing some typographical conventions of this document, the rest of this introductory chapter will be devoted to a review of key background topics in Western music theory, acoustics, and mathematics.

Chapter 2 reviews work that is related to this thesis. The related work falls into two major categories: related theories of tuning and related dynamic intonation software or instruments.

Chapter 3 presents a theory of tuning. The theory is of general applicability but is oriented towards application to dynamic intonation. This theory begins with the presentation of several mathematical models of tunings, one of which is chosen to be the primary model used in this work. A mathematical formalization of pitch notation is presented. The criteria of just perfect fifths and major thirds are introduced as a way to evaluate tunings. Several types of diatonic tunings are discussed in the context of the formalisms previously developed. The chapter concludes with a discussion of triadic tunings and their theory, which are particularly applicable to just intonation.

Chapter 4 describes Helm, the dynamic intonation software suite developed as part of this thesis research. It presents the language in which scores are described, which includes facilities for intonation annotation. It then presents the two main things that can be done with such score descriptions. The first is the creation of Standard MIDI File realizations of the score. The second is the
creation of an instrument that follows along in the score, retuning notes it receives from a MIDI keyboard in real-time.

Chapter 5 presents some conclusions about the theory and software of this work. It also presents some experiments in just intonation that were conducted using the theory and software developed in this work.

Chapter 6 suggests some future directions that this work could take. These include extensions to the theory, extensions to the software, and new synthesis techniques related to tuning.

1.5 Document Conventions

Terms appear in boldface where they are defined. Sometimes new terms are introduced without definition or before their definition; in these cases, they appear in italics. In general, mathematical functions are notated using Greek letters and variables are notated with Roman letters. The plural form of an acronym will be the same as its singular form, like the word “moose.” This avoids the appendage of ‘s’ or ‘(s)’ to acronyms.

1.6 Review of Key Background Topics

It is assumed that the reader has a basic background in Western music theory, acoustics, and mathematics. A few background topics are especially important to this work and/or are treated in slightly idiosyncratic ways, so they merit review before we proceed with the main body of this work. These topics are pitch and interval naming, the harmonic model of tones, frequency ratio units, and modulo and integer division.

1.6.1 Pitch Names

Traditional staff-based pitch notation is difficult to incorporate into text. In this work, a completely equivalent textual representation, as suggested by the Acoustical Society of America, will be used. The following explanation is adapted from [13, xiii]. Pitches will be identified by class and register. Class will be symbolized by a letter in the set {A,B,C,D,E,F,G} followed by any number of sharp (♯) or flat (♭) signs. Register will be symbolized by an integer following the class. A register refers to all pitches that can be notated on the area of the staff from a given C to the next B above it. Defined in this way, any B♯ has the same register as the B that is its natural version; thus in 12TET B♯3 sounds the same as C4. Likewise, any C♭ has the same register as the C that is its natural version; thus in 12TET C♭4 sounds the same as B3. Middle C is C4.

1.6.2 Interval Names and Abbreviations

A knowledge of the traditional interval naming system is important to this work. This system names intervals by quality and step-type; for example, the interval
“major third up” has the quality “major” and the step-type “third up.”

The basic qualities are major, minor, diminished, and augmented and are abbreviated ‘M,’ ‘m,’ ‘d,’ and ‘A’ respectively. (“Augmented” is abbreviated ‘A’ because the more traditional abbreviation, ‘A,’ can lead to confusion because it is a pitch class letter.) Diminution and augmentation can have degrees associated with them, for example “doubly diminished.” These degrees will be abbreviated as superscripts, for example ‘d²’ for doubly diminished. The degree of one is implicit.

Large step-types such as “tenth up” are valid. Intervals with step-type “prime” are valid, but there are no such step types as “prime up” or “prime down” since their existence would lead to two names for the same interval. (Some readers may be more familiar with the term “unison,” which means the same thing as “prime.”) Step-types are abbreviated by their associated integer supplemented with a sign for direction. For example, a third up is indicated by +3. As with numbers, the positive sign is often left implicit, especially when discussing interval magnitudes, which are always positive. When one says “these two pitches form a major third,” one is discussing an interval magnitude; when one says “the interval from this pitch to that one is a major third up,” one is discussing an interval. In other words, intervals other than primes are directed.

When the abbreviations for quality and step-type are put together, the sign floats to the beginning; for example, a major third up would be abbreviated as ‘+M3.’

Be aware that the direction of an interval cannot be relied upon to determine the direction of the frequency ratio it is tuned to. (A frequency ratio greater than one is directed upwards, a ratio less than one is downwards, and a ratio of 1 has no direction.) It is the direction of the frequency ratio an interval is tuned to that determines whether that interval will be heard as ascending, descending, or neither.

The following are examples of situations in which an interval’s direction may not correspond to its heard direction, the direction of the frequency ratio to which it is tuned. Primes do not have direction, yet some of them need to be tuned to ascend and descend. For example, A1 should ascend and d1 should descend. Also, there are upward intervals that should be tuned to descend and downward intervals that should ascend. For example, +d² (a doubly diminished second up, such as C♯ to D♭♭) should descend. Finally, there are intervals that can be tuned to ascend, descend, or do neither, according to the particulars of the tuning. The best example of this is +d2. In 12TET, +d2 (a diminished second up, like C♯3 to D♭♭3) neither ascends nor descends, but there are tunings that make it ascend, and tunings that make it descend.

There is no such interval as a “half step” or “whole step” since these phrases are ambiguous. Unambiguous descriptions such as A1 or +m2 for half step and M2 for whole step should be used.

Interval expressions can be formed using addition and negation. (Formally, intervals under addition form an Abelian group [14, 78].) For example, the expression “P₅ - M₂” equals P₄, and “M₃ + 2M₂” equals A₅. (Integer scaling as in the expression “2M₂” is allowed since it is simply a shorthand for
“M₂ + M₂.”) The rules for evaluating such expressions will not be given here since many readers can evaluate such expressions intuitively and, in addition, the most convenient formalization of these rules is for intervals in fifth-register vector form, to be presented later in this work. One aspect of such expressions does deserve a brief mention, though. Although primes do not have direction, they can be negated in expressions. For example, “−A₁” is not an interval; it is an expression that equals the interval +d₁. The situation is analogous to that with the number zero; “−0” is not a number; it is an expression that equals the number zero. The point here is merely to distinguish between intervals and interval expressions.

1.6.3 The Harmonic Model of Tones

We define a tone to be a sound that is perceived to have a single frequency. This is sometimes called a pitched sound, or a sound that has a pitch, but we will scrupulously observe the following distinction between pitch and frequency. A pitch is a discrete element of a musical theory or system (like ‘A⁴’). A frequency, on the other hand, is a continuous physical quantity indicating the rate of a periodic phenomenon such as a tone.

Tones are herein modeled as being composed of a sum of harmonically related sinusoids. The amplitude and phase of these sinusoids may vary slowly over time. Tones are assumed to be perceived as having a frequency equal to the fundamental, the lowest of these sinusoids. For example, a sound composed of sinusoids at 100, 200, and 300 Hz would be identified as having a frequency of 100 Hz. Phenomenon such as missing fundamentals are therefore not accommodated by this model. In addition, this model does not accommodate frequency fluctuations such as vibrato.

1.6.4 Frequency Ratios as Doublings

The topic of tuning deals extensively with ratios of frequencies. There are often advantages to expressing these ratios logarithmically. For example, log frequency ratios are more perceptually meaningful, and equations involving them are often simpler than with plain ratios. The base two logarithm is particularly useful because of the special perceptual status of a doubling of frequency.

For these reasons, this work will often use the doubling, the base-two logarithm of frequency ratio, instead of frequency ratio itself. The frequency ratio \( r \) is equivalent to \( \log_2(r) \) doublings. For notational brevity, \( \log_2(r) \) will be notated as \( r \), i.e. the log operation will be indicated by underlining.

The term “doubling” is used instead of “octave” because an octave is an interval between pitches in Western music, not a ratio between frequencies. We will scrupulously observe the distinction between interval and frequency ratio as analogous to the distinction between pitch and frequency. An interval is a discrete element of a musical theory or system (like ‘P₈’). A frequency ratio, on the other hand, is a continuous quantity indicating the ratio between two frequencies. Thus, although in practice an octave usually corresponds to a dou-
bling, an octave is an interval in Western music whereas a “doubling” is a logarithmic frequency ratio. One slight problem with “doubling” is that it already has a separate musical meaning, as in “voice doubling,” but these two meanings can usually be disambiguated from context.

In the tuning literature, a frequency ratio \( r \) is often expressed as \( 1200r \) cents. This unit is useful because it facilitates comparisons with 12TET, a tuning that many people are familiar with. Nonetheless, it will not be used in this work since it imparts a normative status to 12TET that is culturally biased. Where a cent-sized unit is called for, the millidoubling, or mil, (one thousandth of a doubling) works quite nicely since it equals 1.2 cents.

### 1.6.5 Modulo and Integer Division

Modulo and integer division are infix binary operators indicated by ‘mod’ and ‘÷’ respectively. Though their operands are usually integers, they can be real as well. (Integer division is “integer” only in the sense that its result is an integer.) The ‘mod’ and ‘÷’ operators are defined as follows.

\[
x \div y = \lfloor x/y \rfloor
\]
\[
x \mod y = x - y \lfloor x/y \rfloor
\]

(The “floor” of \( x, \lfloor x \rfloor \), is the maximum integer \( i \) such that \( i < x \).)

Here we will present some examples of these operators applied to integer operands. For positive \( n \), if \( n/k \) is represented in mixed-fraction form as \( i + j/k \), \( n \div k = i \) and \( n \mod k = j \). For example, \( 10/7 = 1 + 3/7 \), which means that \( 10 \div 7 = 1 \) and \( 10 \mod 7 = 3 \). A slightly tricky thing about ‘mod’ and ‘÷’ is how they behave when \( n \) is negative. For example, one might think of \((-3)/7\) as \( 0 + -3/7 \), but, for the purposes of modulo and integer division, it should be thought of as \(-1 + 4/7\). In other words, the integer part, not the remainder, must carry the sign. So \((-3) \div 7 = -1 \) and \((-3) \mod 7 = 4 \).
Chapter 2

Related Work

2.1 Tuning Theory

This work’s tuning theory builds upon that of various authors, borrowing their strong points, reacting against their weaknesses, increasing their formality in some areas and eliminating unnecessary formal complexity in others. Fundamentally, it is a synthesis of the tuning theory of Lindley and Turner-Smith, the pitch representation of Regener, and various original contributions, particularly with respect to just intonation.

2.1.1 Helmholtz and Ellis

Helmholtz and Ellis established the basis of modern tuning theory, and provide an interesting late-19th century perspective on tuning. On the subject of just intonation, Ellis seems far more cognizant of its limitations than Helmholtz, who glosses over them in his great zeal for pure consonances. Ellis’ theory of duodenes [12, 457-66] is suggestive of a two-dimensional representation of pitch class in terms of P5 and M3. This idea partially inspired the notion of triadic pitch class used in this thesis. A duodene is a simple slip-free triad set, an important concept in this thesis’ coverage of just intonation.

2.1.2 Blackwood

Blackwood [3] provides detailed intonation annotations to musical examples, illustrating various issues associated with just intonation. Blackwood’s tuning theory is hampered by the fact that his “interval size convention” restricts him to frequency ratios in the range [1, 2). This approach works fine as long as any two pitches in the same register are tuned to less than a doubling apart. Unfortunately this is patently not the case for useful pitch sets, since they need to contain pitches such as C3 and B3. In his own words, “the interval size convention, although helpful in revealing interconnections among diatonic intervals,
cannot be expected to produce similar insights in the case of the chromatic intervals" [3, 58]. (In Blackwood’s vocabulary, diatonic intervals are those that are present in the pitches of a major scale, while chromatic intervals are all others.) Blackwood’s inability to deal with the fact that tunings of pitches in a single register are not confined to a doubling causes errors such as tuning C♭3 near B3 rather than B2.

### 2.1.3 Lindley and Turner-Smith

This thesis is based to a large extent on the work of Lindley and Turner-Smith [14]. Their work includes the notion of pitch class as a possibly two-dimensional entity characterized by P5 and M3 relationships. This is similar to the notion of triadic pitch class used in this thesis. Their use of Cayley diagrams to describe these P5 and M3 relationships is instructive. The idea of characterizing tunings by their deviation from just P5 and M3 is used in their work and in this thesis. Their work is one of the few that uses the mil rather than the cent as the unit of logarithmic frequency ratio, although they conflate interval and frequency ratio by identifying a mil as a millioctave, not a millidoubling. Their work makes use of an interval size convention similar to Blackwood’s and therefore suffers from similar inability to deal with register. Their use of group theory to describe pitch class relationships is powerful but in general adds unnecessary complexity compared to the simple vector representation of pitch used in this thesis.

### 2.1.4 Regener

The main contribution of Regener [20] related to this thesis is the representation of diatonic pitches and intervals as two-dimensional vectors. The bases of the vector space are any two intervals whose linear combination can generate all other intervals. A particularly useful pair of bases is P5 and P8. The vector representation of pitch was discovered independently by the author of this thesis and later confirmed by his discovery of Regener’s work, which, oddly, is not well-known in the tuning community.

Although fancy group theory formalisms seem popular among many mathematically inclined theorists [14][13], Regener’s contribution was that he correctly identified the somewhat more pedestrian topic of linear algebra as having much to contribute to the study of pitches and intervals. Unfortunately, Regener arrives at the vector representation of pitches via an unwieldy intermediate representation in terms of diatone and quint, whereas it is possible to go directly from conventional pitch notation to vector representation, as we do in this thesis. Regener’s work is confined to the study of regular tunings, although the vector representation can be used to study all diatonic tunings, as will be seen in this thesis. Regener’s work is confined to a two-dimensional representation of pitch, whereas this thesis will show how a three-dimensional representation is a powerful theoretical tool for some types of tunings.
2.2 Dynamic Intonation Systems

Dynamic intonation systems fall into two major categories: real-time and non-real-time. Non-real-time systems are typically software that converts some kind of score file with intonation information into a MIDI or audio file version of the piece. Such systems include the modified MIDI sequencers of Siegel [22][23] and Frantz [9]. In addition, most non-real-time music synthesis languages such as Csound [26] are capable of output with dynamic intonation. Real-time systems are typically software that re-maps keyboard input to MIDI output according to some set of intonation rules. Three such systems are described below. There is a rich history of experimentation in acoustic keyboard instruments with many more than 12 keys per doubling. Some of these can have dynamic intonation, but they will not be discussed here since, unlike digital electronic instruments, they are still limited to a fixed number of frequencies per doubling.

2.2.1 Sethares

Sethares [21] implemented software that automatically retunes a MIDI stream (in a file or in real-time) based on knowledge of the spectra produced by the MIDI device and a mathematical model of human dissonance perception. The goal was to minimize the dissonance of the output audio, where dissonance was defined according to the model of Plomp and Levelt [16]. This algorithm does not avoid any of the melodic artifacts associated with just intonation, but it automates intonation decisions. In addition, it can handle sounds with inharmonic spectra, such as bells, which cannot be tuned using conventional theories of just intonation. As Sethares admits, a shortcoming of his approach is that consonance is not a globally desirable property in music. This algorithm would probably be best applied in a selective fashion to chords for which consonance is most appropriate.

2.2.2 Waage

Waage [27][29] has proposed a real-time dynamic just intonation keyboard system. This system tunes the note produced by a key according to what other keys are currently held. As Doty points out [7] and Waage concedes [28], a major limitation of this system is that it only considers what keys are currently pressed down, whereas information about recently-depressed keys could cause it to make “smarter” intonation decisions.

2.2.3 McCoskey (FasTrak)

Marion McCoskey’s FasTrak is a dynamic intonation MIDI keyboard retuning program. It works only with the Creative Labs Sound Blaster card. It is intended for just intonation in the extended sense, that is, intonation that uses only ratios of whole numbers but these whole numbers are not limited to products of 2, 3, and 5. Like Waage’s proposal, it tunes notes according to what keys
are currently depressed. The rules for tuning are stated in a chord logic file and a scale tuning file that can be customized by the user. It includes a split keyboard mode in which one section of the keyboard is used to select the tuning of another. It capabilities are similar to the Justonic Pitch Palette discussed below.

2.2.4 Justonic, Inc.

The Pitch Palette software from Justonic, Inc. includes a dynamic intonation MIDI keyboard retuning program. The program tunes chords according to their root and the key they belong to. Root and key can be guessed by the software, or they can be specified via another MIDI device (an organ pedalboard, for instance). The Pitch Palette also includes a sequencer in which MIDI files can be annotated with root and key information and played back using their tuning algorithms. It is not clear whether the root and key guessing is done only on the basis of currently depressed keys, as in Waage's proposal [27][29], or uses information about recently-depressed keys as well.

A feature of this system that certainly goes beyond Waage's proposal is that the user can redesign the tuning of a chord so that when the system recognizes that it is being played, the user's tuning, not the default tuning, is used. Combined with the fact that each chord can be assigned a separate tuning depending on what the current root is, a wide variety of versions of the same chord can be played if another MIDI device is used to do root selection. This approaches the flexibility of the dynamic intonation software presented in this thesis, at the cost of added complexity for the performer. The advantage of such a system is that the score is not needed beforehand, so improvisation is possible, as well as freedom from the inevitable shortcomings of a score follower.
Chapter 3

Tuning Theory

3.1 Models of Tunings

This section will discuss several different models of tunings. Since these models apply to far more than just diatonic (traditional Western) tunings, care has been taken to present them in a general fashion. While general tuning theory is interesting, the main goal of this section is to provide the background necessary to understand the model of diatonic tunings that will be used in this work, the register-zero transposable tuning.

3.1.1 Absolute Tunings

The most basic model of a tuning is an absolute tuning. An absolute tuning maps a set of pitches to frequencies. This definition may sound odd because colloquially, “pitch” can be used interchangeably with “frequency.” As introduced in Section 1.6.3, in this work we use “pitch” in a restricted sense, where a pitch is an element of a culture’s music theory, whereas frequency is a physical quantity. A pitch can be thought of as an abstract representation of a frequency that can be made concrete only through the application of a tuning. An example absolute tuning is one that maps the set of pitches \{C3, E3, G3\} to \{261 Hz, 329 Hz, 391 Hz\}. Figure 3.1 is a block diagram of an absolute tuning.

![Figure 3.1: Block diagram of an absolute tuning](image)

3.1.2 Relative Tunings

For theoretical purposes, the absolute model is usually too specific. Perceptually, most of a tuning’s identity comes from the ratios between its frequencies
rather than the exact values of the frequencies themselves. For example, even if the frequencies of one piano's strings are 1.01 times higher than those of another piano, they can both be said to be tuned in 12TET. This motivates the definition of a relative tuning as a map from a set of pitches to frequency ratios that are understood to share an implicit undefined reference. For example, a relative tuning of the pitches \{C3, E3, G3\} might map them to the ratios \{1, 1.26, 1.5\}.

A relative tuning's frequency ratios can be interpreted as frequencies with undefined units, or relative frequencies. Note that a frequency ratio does not have to be a ratio of integers: it can be any real number. Figure 3.2 is a block diagram of a relative tuning.

Figure 3.2: Block diagram of a relative tuning

In order to understand the idea of a relative tuning, it may be helpful to think about how one could be used to build an absolute tuning. To build an absolute tuning from the example relative tuning above, a reference frequency \(x\) must be chosen and each element of the ratio set multiplied by it to yield the frequency set \(\{x, 1.26x, 1.5x\}\). If \(x\) is chosen to be 261 Hz, the resulting absolute tuning is \{261 Hz, 329 Hz, 391 Hz\}. Figure 3.3 is a block diagram of an absolute tuning built using a relative tuning, a reference frequency, and a multiplier (symbolized by ‘\(\Pi\)’).

Figure 3.3: An absolute tuning built from a relative tuning

### 3.1.3 Transposable Tunings

A **transposable tuning** is a relative tuning of a transposable pitch set. Mathematically, a transposable pitch set is one that can be made to form an Abelian group under addition. Musically, this means that the pitch set must have the following two properties. The first is that the interval from any pitch \(x_1\) to any other pitch \(x_2, x_2 - x_1\), must be well-defined. The second is that the transposition of any pitch \(x\) by any interval \(y, x + y\), must be a well-defined pitch. For example, the diatonic pitch set (all pitches representable in traditional Western notation) is transposable.

The thing that is special about a relative tuning of a transposable pitch set is that it isn’t just one tuning: it can be thought of as summarizing an entire class of tunings. Mathematically, a transposable tuning \(\tau(x)\) can be thought of
as summarizing an entire class of tunings of the form \( \tau_y(x) = \tau(x+y) \). For example, some diatonic tunings are playable only in keys closely related to C major. We will refer to such a tuning as being “defined about C.” Because diatonic pitches are transposable, this special relationship to C is totally incidental. For theoretical purposes, it is still the same tuning even if it is transposed to be defined about B♭, or any other pitch. For example, if \( \tau(x) \) is the tuning defined about C, its transposed version about B♭ is simply \( \tau_{B^2}(x) = \tau(x+M2) \).

### 3.1.4 Register-Doubling Tunings

Many pitch sets can be represented as a set of *registraly represented pitches*. A *registraly represented pitch* is an ordered pair \((c, r)\) where \(r\) (the register) is an integer and all pitches with the same \(c\) (class) are considered equivalent in some sense. For example, in this work, diatonic pitches are represented by strings such as “C♯4,” indicating a class of C♯ and a register of four. Most pitch sets that can be represented registraly are tuned such that an increment of register by one corresponds to a doubling of frequency. This type of tuning is called a *register-doubling tuning*. Western music and many other musics of the world use register-doubling tunings.

Since all such tunings deal with register in the same way, it is useful to think of them as being built from a tuning of all pitches in register zero (all pitches of the form \((c,0)\)). To form a tuning of all registers, all that is necessary is to add some simple, generic “machinery” around the register-zero tuning. The register-zero tuning takes only class information as input since the input’s register is known to be zero by definition. Figure 3.4 is a block diagram of a register-doubling tuning built from a register-zero tuning, \(\phi(c)\), surrounded by appropriate machinery. The block labeled “\(2^x\)” exponentiates its input.

![Figure 3.4: Block diagram of a register-zero tuning](image)

### 3.2 The Choice of a Model for Diatonic Tunings

To choose an appropriate model for a tuning, we must start with its absolute model. We can then discriminate between what information in the absolute model is essential and what information is incidental or redundant. This discrimination between the essential and inessential will suggest which (if any) of the more condensed models is most appropriate. The most appropriate model
will factor out all incidental information and encode redundant information compactly.

Our absolute model of a diatonic tuning is a map from the diatonic pitch set (all pitches representable in traditional Western notation) to frequencies. We will consider the actual frequencies of a diatonic tuning to be incidental; therefore we will use at least a relative model. The diatonic pitch set is transposable, so this relative model is a transposable one. Finally, we will assume diatonic tunings to be register-doubling; therefore we will use a register-zero model in order to avoid redundant information about all registers. Briefly, what we consider to be essential about a diatonic tuning is the way it maps pitches in register zero to frequency ratios. Thus we will use a register-zero transposable tuning to model diatonic tunings.

3.3 Diatonic Pitches as Fifth-Register Vectors

In addition to the traditional representation of diatonic pitch, this work will use a mathematical representation called the fifth-register vector (FRV). The fifth-register vector \([x_r, x_f]^T\) is the pitch \(x_f P5 + x_r P8\) away from C0. For instance, G0 is \([1, 0]^T\) since it is \(1P5 + 0P8\) away from C0. Table 3.1 shows the location of some traditionally-represented pitches in FRV vector space.

<table>
<thead>
<tr>
<th>(x_r)</th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>A♭1</td>
<td>E♭2</td>
<td>B♭2</td>
<td>F3</td>
<td>C4</td>
<td>G4</td>
<td>D5</td>
<td>A5</td>
<td>E6</td>
</tr>
<tr>
<td>3</td>
<td>A♭0</td>
<td>E♭1</td>
<td>B♭1</td>
<td>F2</td>
<td>C3</td>
<td>G3</td>
<td>D4</td>
<td>A4</td>
<td>E5</td>
</tr>
</tbody>
</table>

Table 3.1: Some pitches in FRV space

Readers familiar with the notion of the circle of fifths may find it helpful to think about the “fifths” axis of FRV as being an unwound circle of fifths. (The notion of the circle of fifths only makes sense in the context of tunings in which all pitches separated by a d2 are tuned to the same frequency. Therefore, the circle of fifths has no place in a tuning-independent representation of pitch, and we must unwind it into a line (axis) of fifths.)

The FRV representation of pitch was discovered independently by the author, who then later found that Regener [20] had discovered it as well.

Like traditional pitch representation, the FRV is a registral representation. The use of “register” to describe the second component of an FRV may seem odd, as the range of pitches in any one register is very large (in fact, infinite) and overlaps with other registers. These spread-out registers are not consistent with the traditional, dense sense of the term “register,” but they are wholly consistent with the definition of register and register-doubling tunings given previously in this work.

FRV can be interpreted as pitches or intervals. For example, \([1, 0]^T\) can be interpreted as the pitch P5 away from C0 (G0) but it can also just be interpreted as P5. Pitches and intervals were actually always the same thing, but
traditionally they have been represented differently. This difference in representation can avoid confusion; for example, it would sound a bit weird to say that G3 is E♭0 above E3. But for mathematical purposes, it makes sense to use the same representation for pitches and intervals.

The familiar mathematical operations of vector addition and negation on FRV are equivalent to the corresponding operations on traditionally represented intervals and/or pitches. For example,

$$P5 - M2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = P4.$$  

An interesting fact is that the FRV is not the only vector representation of interval possible. Regener [20, 58-85] explores the idea of alternate interval spaces, although unfortunately within his unwieldy formalism of diatones and quintes. Bases can be changed from P5 and P8 to a variety of other intervals via standard matrix transformation. For instance, to find the transformation matrix to switch to a space whose bases are M2 and A1, we invert the matrix whose columns are M2 and A1 expressed as FRV.

$$M = \begin{pmatrix} M2 \\ A1 \end{pmatrix}^{-1} = \begin{bmatrix} 2 & 7 \\ -1 & -4 \end{bmatrix}^{-1} = \begin{bmatrix} 4 & 7 \\ -1 & -2 \end{bmatrix}$$

Another interpretation of the transformation matrix is a matrix whose columns are P5 and P8 expressed in the new bases of M2 and A1.

Not just any two intervals can form a basis of an interval space. In particular, two FRV $x$ and $y$ are valid bases if and only if the matrix

$$M = \begin{bmatrix} x_f & y_f \\ x_r & y_r \end{bmatrix}^{-1}$$

exists and has only integer components.

### 3.4 Converting Pitches to FRV

This section will demonstrate that any pitch can be represented as an FRV by giving an explicit algorithm for the conversion from traditional notation to FRV. This conversion algorithm is original to this thesis, i.e. not present in Regener’s work.

For our present purposes, it will be useful to consider traditional pitch notation as having three components: letter, accidental section, and register. It is possible to represent these components as a triple of integers $(x_l, x_s, x_r)$ where $x_l$ is in $[0 \ldots 6]$ and indexes into the letter array [C,D,E,F,G,A,B], $x_s$ is the number of sharps in the accidental section (where flats count as negative sharps), and $x_r$ is the register. For example, D♯4 is $(1,1,4)$, G6 is $(4,0,6)$, and
E♭5 is (2, −1, 5). Using this triple notation, we can express the interval from C0 to any note $x = (x_{\ell}, x_s, x_r)$ as

$$\beta(x) = \gamma(x_{\ell}) + x_s \tilde{A}1 + x_r P8$$

where $\gamma(x_{\ell})$ is as follows:

<table>
<thead>
<tr>
<th>$x_{\ell}$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma(x_{\ell})$</td>
<td>P1</td>
<td>M2</td>
<td>M3</td>
<td>P4</td>
<td>P5</td>
<td>M6</td>
<td>M7</td>
</tr>
</tbody>
</table>

Note that $\gamma(x_{\ell})$ yields the intervals of an ascending major scale.

To convert pitches to FRV, we need to rewrite $\beta(x)$ with all intervals expressed in FRV rather than traditional form. This gives us

$$\beta(x) = \gamma(x_{\ell}) + x_s \begin{bmatrix} 7 \\ -4 \\ 1 \end{bmatrix} + x_r \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

where $\gamma(x_{\ell})$ is as follows:

<table>
<thead>
<tr>
<th>$x_{\ell}$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma(x_{\ell})$</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>-1</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

Reviewing the concepts above, Table 3.2 lists some pitches in traditional and FRV form.

<table>
<thead>
<tr>
<th>C0</th>
<th>D♭1</th>
<th>E♭-1</th>
<th>F0</th>
<th>G-1</th>
<th>A0</th>
<th>B0</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0]</td>
<td>[0]</td>
<td>[2]</td>
<td>[4]</td>
<td>[-1]</td>
<td>[1]</td>
<td>[3]</td>
</tr>
<tr>
<td>[0]</td>
<td>[0]</td>
<td>[-3]</td>
<td>[1]</td>
<td>[-1]</td>
<td>[-1]</td>
<td>[-2]</td>
</tr>
<tr>
<td>E♭0</td>
<td>B♭0</td>
<td>F♭0</td>
<td>C♭0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[-3]</td>
<td>[-2]</td>
<td>[6]</td>
<td>[7]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Example pitches in tradition and FRV form

### 3.5 Fifth Tunings

Previously, we decided to model diatonic tunings as register-zero transposable tunings. If diatonic pitches are represented as FRV, a diatonic register-zero transposable tuning is a **fifth tuning**, $\phi(x_f)$, that maps the pitch class $x_f$ to a frequency ratio. It is called a fifth tuning because the class of an FRV is its “fifths” component. Without loss of generality, we assume that

$$\phi(0) = 0 \quad (3.1)$$

in order to simplify calculations involving $\phi(x_f)$. (Recall that we use underlining to indicate the base-2 logarithm.) From the definition of a register-doubling
tuning, the relationship between a diatonic transposable tuning \( \tau(x) \) and its underlying fifth tuning \( \phi(x_f) \) is

\[
\tau(x) = \phi(x_f) + x, \quad \text{where} \quad x = \begin{bmatrix} x_f \\ x_r \end{bmatrix}.
\] (3.2)

At this point we have completed the development of the formal framework in which we will examine diatonic tunings. Before we discuss actual diatonic tunings, though, we need to have some reference or basis for what a “good” diatonic tuning would be. Our formal framework makes no value judgments along these lines and therefore gives us no guidance. As will be explained in the next section, our basis for such value judgments will be the beatless tunings of M3 and P5.

### 3.6 Beatless (Just) M3 and P5

In this section we will discuss beating and derive the frequency ratios for beatless tunings of M3 and P5. These ratios will form an important standard by which tunings will be judged.

#### 3.6.1 Beating

When tones are played simultaneously, their harmonics can interact in a variety of ways in our auditory system. One potential interaction is **beating**. It can occur when two of the harmonics of the constituent sounds fall within about 15 Hz of each other. **Beating** is a periodic pulsation of the loudness of a sound at a rate equal to the frequency difference between the nearly-coincident harmonics. It is not always easy to predict what combinations of tones will cause beating. Whether or not beating is heard depends upon the amplitude of the harmonics in question and a variety of other factors whose discussion is beyond the scope of this work. For our purposes it is enough to know that beating happens in many musical situations.

It is important to realize that beating is a perceptual, not a physical phenomenon. In other words, it is psycho-acoustic, not acoustic, in nature. Many authors explain beating by showing a graph of two superposed sinusoids whose frequencies are close. Indeed it is easy for our visual system to extract a convincingly pulsating shape from such graphs. This approach takes advantage of visual shape detection as an analog of auditory loudness perception. The problem with this approach is that it tends to give the impression that beating is a physical, not a perceptual phenomenon. Since we can both see and hear beating, it is tempting to assume that beating “exists” in some absolute sense. This is of course false. It is only by virtue of a system capable of some nonlinear processing (such as our auditory system) that two sinusoids can produce something that fluctuates at their difference frequency. In the air (a largely linear system), nothing beats.
3.6.2 Beating and Dissonance

We will assume that the phenomenon of beating is perceived as a kind of dissonance, where the rate of beating is proportional to the amount of dissonance. It is not our purpose here to justify this assumption any more than to say that it is fairly commonly held, and has always been used as a means to evaluate tunings. Discussions of beating as a form of dissonance can be found in [8], [12], and many other works on tuning and/or psycho-acoustics. Our purpose is to proceed with this assumption and see what theoretical and practical consequences it leads us to.

Since M3 and P5 are used as strong consonances in diatonic music, we will assume that when they are rendered harmonically, it is desirable to tune them so that they beat as slowly as possible, ideally not at all.

3.6.3 The Derivation of Beatless M3 and P5

The beatless tuning of M3 and P5 can be derived from the beatless tuning of a major triad, which in turn can be derived from the harmonic structure of tones. For this derivation, we introduce the notion of a **schematic spectrogram**, a series of lines that indicate the frequency components of a sound without reference to their amplitude or absolute location on a frequency scale. A tone with harmonics 1 through 20 is illustrated in Figure 3.5.

![Figure 3.5: A schematic spectrogram of a tone](image)

If we listened to harmonics 1 through 5 of this sound melodically, it would sound like a rolled major chord with the root appearing at two octave transpositions. If the root were C2, it would sound like a rolled chord of C4–C5–G5–C6–E6. Why stop at harmonic 5? Well, harmonic 6 is just an octave transposition of harmonic 3, and then harmonic 7 sounds like an m7 above harmonic 4 (albeit a very flat m7 to ears accustomed to 12TET). In terms of pitch equivalents, harmonic 6 would be a G6 and harmonic 7 would be a funny-sounding Bb6. Since our goal is to derive the major triad from the harmonic structure of tones, we should stop at harmonic 5. (Nonetheless, it is worth mentioning that many theorists have investigated the possibility of using harmonic 7 as a basis for tuning the 7th of dominant seventh chords.)

Discarding the octave transpositions of the root, if we actually form a chord from tones based on harmonics 1, 3, and 5, it has the schematic spectrogram that appears in Figure 3.6. (In this and subsequent figures, example pitches are given for the fundamentals of the tones illustrated.) Since all the harmonics of the constituent tones are aligned, this chord is beatless and extremely consonant. In fact, it could be considered a spectrally shaped version of the root, not a chord of distinct tones.

Transposing the fundamentals of the tones of this chord down to be within an octave of each other, we get a major triad. Harmonic 5 is transposed down...
two octaves to form an M3, which therefore has ratio 5/4. Harmonic 3 is transposed down one octave to form a P5, which therefore has ratio 3/2.

This major triad is illustrated in Figure 3.7, and, to be able to see the alignment of harmonics more clearly, it appears in overlapped form in Figure 3.8.

Although this chord looks like a mess compared to the untransposed version, it does not sound like a mess. In fact it is still considered beatless. There are a few reasons for this.

The harmonics that do not align would cause beating only if the root were fairly low. Let $a$ be the frequency of the root, which is also the distance between harmonics of the root. If the root were middle C, $a \approx 265$ Hz. In this situation, two harmonics would have to be within about $a/20$ of each other in order to beat. But no such close spacing exists in the major triad. The closest any two harmonics come to each other is $a/4$. The observation that a major triad is not so consonant at low frequencies is in accord with conventional musical wisdom, which would prohibit any close voicing of a chord at low frequencies.

Another reason that a major chord doesn’t sound as messy as it looks is that the combined waveform is periodic with a period two octaves below the root. In other words, if the root is C3, the combined waveform is periodic with the same period as C1. In fact, due to the perceptual phenomenon of difference tones, this C1 may actually be heard along with such a chord.

We have derived the ratios for a beatless M3 and P5 and can now proceed to study diatonic tunings with reference to how they treat these intervals.
3.7 Evaluations of Fifth Tunings

We have derived the just tuning of P5 and M3 in the previous section. Since we have decided to evaluate tunings on the basis of how close they come to the just versions of these intervals, we need to be able to derive this information from a fifth tuning $\phi(x_f)$. Let us first consider how to do this for P5. There can be as many tunings of P5 as there are pitch classes: for example, the P5 from C3 to G3 could be different from that from G3 to D3, and so on. (Tunings of P5 from register to register are always the same; for example the P5 from C3 to G3 is always the same as that from C4 to G4.) We shall develop a function, $\Phi(x_f)$, that describes the tuning of the P5 from a pitch with class $x_f$ to one with class $x_f + 1$ in terms of its deviation from just. For a fifth tuning $\phi(x_f)$,

$$\Phi(x_f) = \left( \phi(x_f + 1) - \phi(x_f) \right) - \frac{3}{2}. \quad (3.3)$$

(Recall that we use underlining to indicate the base-2 logarithm.) So $\Phi(0)$ is how far CN to GN is from just, $\Phi(1)$ is how far GN to DN is from just, $\Phi(2)$ is the same thing for DN to AN, and so on. The function $\Phi(x_f)$ is similar to Lindley and Turner-Smith’s $\text{tem}(q, V)$ [14, 58].

Let us now consider how to evaluate a fifth tuning in terms of its tunings of M3. Again, there can be as many tunings of M3 as there are pitch classes: for example, the M3 from C3 to E3 could be different from that from G3 to B3, and so on. We shall develop a function, $\Theta(x_f)$, that describes the tuning of the M3 from a pitch with class $x_f$ to one with class $x_f + 4$ in terms of its deviation from just. We begin by observing that for any pitch $x$, the tuning of the M3 from $x$ to $x + M3$ is

$$\tau(x + M3) - \tau(x) = \tau(x + [4 - 2]T) - \tau(x)$$

$$= \left( \phi(x_f + 4) + x_r - 2 \right) - \left( \phi(x_f) + x_r \right) \quad \text{by (3.2)}$$

$$= \phi(x_f + 4) - \phi(x_f) - 2.$$ 

This motivates the definition

$$\Theta(x_f) = \left( \phi(x_f + 4) - \phi(x_f) - 2 \right) - \frac{5}{4}$$

$$= \phi(x_f + 4) - \phi(x_f) - \frac{5}{4}.$$ 

So $\Theta(0)$ is how far CN to EN is from just, $\Theta(1)$ is how far GN to BN is from just, $\Theta(2)$ is the same thing for DN to FN, and so on. The function $\Theta(x_f)$ is similar to Lindley and Turner-Smith’s $\text{tem}(q, III)$ [14, 24].

We will now show how $\Theta(x_f)$ can be expressed in terms of $\Phi(x_f)$. First,
we will express $\phi(x_f + 4) - \phi(x_f)$ in terms of $\Phi(x_f)$ as follows:

$$\phi(x_f + 4) - \phi(x_f) = (3/2 + \Phi(x_f)) + (3/2 + \Phi(x_f + 1)) + (3/2 + \Phi(x_f + 2)) + (3/2 + \Phi(x_f + 3))$$

$$= (3/2)4 + \sum_{i=x_f}^{x_f+3} \Phi(i).$$

Now we can proceed to express $\Theta(x_f)$ in terms of $\Phi(x_f)$ as

$$\Theta(x_f) = \left(\frac{3}{2} + \sum_{i=x_f}^{x_f+3} \Phi(i)\right) - 5$$

$$= C_S + \sum_{i=x_f}^{x_f+3} \Phi(i) \text{ where } C_S = \frac{81}{80}. \quad (3.4)$$

The constant $C_S$ is an important frequency ratio called the syntonic comma, or the comma of Didymus. It is the difference between four just P5 and a just M3 plus 2 doublings. We will now proceed to discuss some specific diatonic tunings, using $\Phi(x_f)$ and $\Theta(x_f)$ as our criteria for evaluation.

### 3.8 Discussion of Specific Diatonic Tunings

In the following sections, we will examine three types of diatonic tunings: regular tunings, truncated tunings, and well temperaments.

#### 3.8.1 Regular Tunings

A tuning $\tau(x)$ is regular if and only if it is a homomorphism, i.e. for all pitches $x$ and $y$,

$$\tau(x + y) = \tau(x) + \tau(y).$$

This is a rather nice mathematical definition, but musically, what does it mean? Musically, a regular tuning is one in which an interval is tuned the same no matter where it occurs. Let us see how this flows from the mathematical definition of regularity. First of all, what does it mean to add pitches? When two pitches are added, think of the second one as an interval, so for example, $x + D0$ becomes $x + M2$. In a regular diatonic tuning $\tau(x)$,

$$\tau(x + D0) = \tau(x) + \tau(D0).$$

30
No matter what $x$ is, it has the same frequency ratio to the pitch an M2 above it. For example, the M2 from C3 to D3 is the same as that from D3 to E3. By identical means, it can be seen that all intervals of a given type are tuned the same no matter where they occur. The definition of a regular tuning above is similar to that of Lindley and Turner-Smith [14, 43] and Regener [20, 87–8].

For regular tunings, $\Phi(x_f)$ and $\Theta(x_f)$ are constants. In other words, all P5 deviate from just by the same amount, and all M3 deviate from just by the same amount. This is a natural consequence of the fact that in a regular tuning, all intervals of a given type are tuned the same no matter where they occur. Since for regular tunings $\Phi(x_f)$ and $\Theta(x_f)$ are constant, we shall often abbreviate them as $\Phi$ and $\Theta$. These are similar to Lindley and Turner-Smith’s $t_v$ and $t_M$ respectively. We shall see how to calculate these amounts below.

12TET is an example of a regular diatonic tuning. In fact, it is the only regular diatonic tuning that can be implemented on an instrument with 12 frequencies per doubling, as will be shown later.

A regular diatonic tuning $\tau(x)$ has a regular fifth tuning underlying it. Let us see why this is true. Since $\tau(x)$ is regular for all pitches $x$ and $y$, this certainly applies to all pitches in register zero. By definition, the register-zero tuning is doing all the tuning of register zero, so it, too, is regular. Since all intervals of the same type are tuned the same in a regular tuning, its fifth tuning can be characterized by a single constant, $v$, the frequency ratio of P5. Since, in addition, we know that $\phi(0) = 0$ (Eq. 3.1), we can state that regular tunings have the simple form

$$\phi(x_f) = x_f v.$$ (3.5)

The values of $\Phi$ and $\Theta$ for a regular tuning can be calculated as follows:

$$\Phi = v - 3/2$$

$$\Theta(x_f) = C_S + \sum_{i=x_f}^{x_f+3} \Phi(i) \quad \text{by (3.4)}$$

$$\Theta = C_S + 4\Phi.$$ (3.6)

We will now present some theoretically important regular tunings using the formulas developed above.

In Pythagorean tuning, all P5 are just, meaning $\Phi = 0$ and therefore

$$\Theta = C_S + 4\Phi = C_S \approx 17.9 \text{ mil.}$$

In 1/4-comma meantone (QCM) tuning, all M3 are just, meaning $\Theta = 0$ and therefore

$$\Phi = (\Theta - C_S)/4 = -C_S/4 \approx -4.5 \text{ mil}$$

We can see that the comma referred to in the name of this tuning is the syntonic comma.
In 12TET, all d2 are unisons, meaning \( \tau(d2) = 0 \). Since \( d2 = [-12 7]^T \),

\[
\begin{align*}
\tau(d2) &= 0 = -12\nu + 7 \\
\nu &= 7/12 \\
\Phi &= -1.6 \text{ mil} \\
\Theta &= 11.4 \text{ mil}.
\end{align*}
\]

In 1/5-comma meantone (FCM) tuning, P5 are flat (from just) by the same amount M3 are sharp (from just). This has an appealing egalitarianism to it.

\[
\begin{align*}
\Phi &= -\Theta \\
&= -\left( C_S + 4\Phi \right) \\
&= -C_S/5 \approx -3.6 \text{ mil} \\
\Theta &= +C_S/5 \approx +3.6 \text{ mil}.
\end{align*}
\]

Once again, we see that the comma referred to in the name of this tuning is the syntonic comma.

Figure 3.9 compares the regular tunings presented above in terms of \( \Phi \) and \( \Theta \). In addition, it shows the line \( \Theta = C_S + 4\Phi \) on which all regular tunings lie.

**3.8.2 Just P5 and M3 Impossible in Diatonic Tunings**

A fundamental problem common to all diatonic tunings is that they cannot have all P5 and all M3 just. Definitely only a regular tuning could accomplish this if it were possible, because it is the only type of tuning in which an interval is tuned the same no matter where it occurs. But since Pythagorean (just P5) and QCM (just M3) tunings are different, we know that having just P5 and M3 in the same tuning is impossible. In regular tunings, this problem leads to compromise tunings like 12TET that lie in between the extremes of Pythagorean and QCM on the line \( \Theta = C_S + 4\Phi \). 12TET, for example, is close to Pythagorean on that line, so it has P5 that are close to just and M3 that are far from just.
3.8.3 The Size of Regular Diatonic Tunings

The size of a tuning is an important factor if it is to be implemented on an instrument that can produce only a small number of frequencies per doubling, like the piano, which can only produce 12. The size of a tuning is the number of unique values \( \phi(x_f) \mod 1 \) takes on. For example, 12TET has size 12 since \( \phi(x_f) \mod 1 \) is \((7/12)x_f \mod 1\), which takes on only 12 unique values. The size of a tuning is also the number of frequency ratios per doubling that is required to implement it. Most regular tunings are very large (many even infinite!) and therefore difficult or impossible to implement on an instrument with a finite number of frequencies per doubling.

Regular tunings with rational \( v \) equal to some fraction \( n/k \) have size \( k \), assuming \( n/k \) is in reduced form. These are the equal temperaments, like 12TET and 19TET (\( v = 11/19 \)). A regular tuning with an irrational \( v \) has infinite size. Mathematically, this means that \( \tau(x) \) is a one-to-one mapping. By Eq. 3.5 and Eq. 3.2,

\[
\tau(x) = \phi(x_f) + x_r = x_f v + x_r.
\]

The function \( \tau(x) \) is one-to-one if and only if any change to \( x \) results in a different value yielded. Clearly a change to \( x_f \) or \( x_r \) alone will result in a different value. In addition, any change to \( x_f \) changes the value of \( \tau(x) \) by an irrational amount, so there is no way to cancel this change with a change to \( x_r \) since \( x_r \) is an integer.

We will sometimes refer to a tuning of size \( n \) as an \( n \)-tuning.

3.8.4 Truncated Tunings

To realize a piece using a regular tuning, a large number of frequencies are often required, either because \( v \) is irrational or because \( v \) has a large denominator, where anything greater than 12 is considered large. One way to limit the size of a regular tuning is to truncate it. A truncated tuning is the same as a regular tuning for a range of 12 or more pitch classes, but tunes all other pitch classes to be enharmonic to ones inside this range. Two pitch classes \( x_f \) and \( y_f \) are enharmonic if and only if \( \#(x_f + c) \mod 12 = \#(y_f + k) + 12 \) for some integer \( k \). In other words, two pitch classes are enharmonic if and only if they are tuned to a multiple of a doubling apart.

The most common form of enharmony tunes pitch classes separated by 12 P5 so that they are separated by 7 doublings. For example, G\# (8) and Ab (−4) can be tuned enharmonically such that \( \phi(8) = \phi(-4) + 7 \). This would make the pitches G\#N and AbN be tuned enharmonically for all \( N \).

Truncated tunings of size 12 always take the same form, since all pitch classes separated by 12 must be tuned enharmonically. A truncated 12-tuning \( \phi(x_f) \) derived from regular tuning \( \phi_R(x_f) \) can be expressed as

\[
\phi(x_f) = \phi_R((x_f + c) \mod 12) + 7((x_f + c) \div 12).
\]
The transposition of \( c \) simply allows the regular range to be shifted in phase. For example, \( c = 0 \) makes the range \( 0 \ldots 11 \), \( c = 1 \) makes the range \( -1 \ldots 10 \), and so on. Since we consider diatonic tunings to be transposable anyway, \( c \) can be dropped, forming the simpler equation

\[
\hat{\phi}(x_f) = \phi_R(x_f \mod 12) + 7(x_f \div 12). \tag{3.7}
\]

The general structure of truncated 12-tunings can be thought of as follows. They consist of chains of 12 pitch classes tuned according to the regular tuning they were derived from. These chains are connected by \( P_5 \) tuned to a wolf fifth characteristic to the truncated tuning. For example, if ‘\( ~ \)’ indicates \( v \) in the tuning they were derived from and ‘\( \sim \)’ indicates a wolf fifth, truncated 12-tunings can be pictured as follows:

\[
\ldots -Ab-Eb-Bb-F \sim C-G-D-A-E-B-F_1-C_1-G_1-D_1-A_2-E_2 \sim B_2-F_2-C_2\sim \ldots
\]

By transposition, the “phase” of the chains can be aligned arbitrarily. For instance, the chains below are also possible.

\[
\ldots -Ab-Eb \sim Bb-F-C-G-D-A-E-B-F_1-C_1-G_1-D_1 \sim A_2-E_2-B_2-F_2-C_2\sim \ldots
\]

Let us calculate \( \Phi(x_f) \) for truncated 12-tunings. If \( \Phi_W \) is the difference between the tuning of the wolf fifth and a just fifth,

\[
\Phi(x_f) = \begin{cases} 
\Phi_W & \text{if } x_f \mod 12 = 11 \\
\Phi_R & \text{otherwise.}
\end{cases}
\]

We can calculate \( \Phi_W \) as

\[
\Phi_W = \Phi(11) \\
= \hat{\phi}(12) - \hat{\phi}(11) - 3/2 & \text{by (3.3)} \\
= 7 - \phi_R(11) - 3/2 & \text{by (3.7)} \\
= 7 - 11(3/2 + \Phi_R) - 3/2 & \text{by (3.5)} \\
= 7 - 11\Phi_R - 12(3/2) \\
= -C_P - 11\Phi_R \text{ where } C_P = 12(3/2) - 7. \tag{3.8}
\]

The constant \( C_P \) is an important frequency ratio called the Pythagorean comma. It is the difference between twelve just \( P_5 \) and 7 doublings. Now that we have calculated \( \Phi_W \), we can see that

\[
\Phi(x_f) = \begin{cases} 
-C_P - 11\Phi_R & \text{if } x_f \mod 12 = 11 \\
\Phi_R & \text{otherwise.}
\end{cases}
\]

For example, a truncated Pythagorean 12-tuning (\( \Phi_R = 0 \)) has

\[
\Phi(x_f) = \begin{cases} 
-C_P - 11(0) = -C_P \approx -19.6 \text{ mil} & \text{if } x_f \mod 12 = 0 \\
0 & \text{otherwise.}
\end{cases}
\]

34
Now let us calculate $\Theta(x_f)$ for truncated 12-tunings. We know that it will take on two values: some value $\Theta_W$ for thirds whose classes span the wolf fifth, and $\Theta_R$ otherwise. In other words, it will be of the form

$$\Theta(x_f) = \begin{cases} 
\Theta_W & \text{if } x_f \mod 12 \geq 8 \\
\Theta_R & \text{otherwise.}
\end{cases}$$

We can calculate $\Theta_W$ as

$$\Theta_W = C_S + \sum_{i=x_f}^{x_f+3} \Phi(i) = C_S + 3\Phi_R + \Phi_W$$

by (3.4)

$$= C_S + 3\Phi_R + (-C_P - 11\Phi_R)$$

by (3.8)

$$= C_S - C_P - 8\Phi_R.$$

So, for truncated 12-tunings,

$$\Theta(x_f) = \begin{cases} 
C_S - C_P - 8\Phi_R & \text{if } x_f \mod 12 \geq 8 \\
C_S + 4\Phi_R & \text{otherwise (by (3.6)).}
\end{cases}$$

For example, truncated Pythagorean 12-tuning ($\Phi_R = 0$) has

$$\Theta(x_f) = \begin{cases} 
C_S - C_P \approx -1.6 \text{ mil} & \text{if } x_f \mod 12 \geq 8 \\
C_S & \text{otherwise.}
\end{cases}$$

Truncated tunings of size greater than 12 are quite interesting. Historically, truncated 13- and 14-tunings were actually used on keyboard instruments, typically through the use of a split key for $G_f/Ab$ and/or $D_f/Eb$ [14, 138–9]. Some modern meantone organs continue to have this feature. For example, the Brombaugh organ in Duke University Chapel, completed in 1997, has a single key for $G_f/Ab$ and $D_f/Eb$ but the tuning of these keys can be toggled between the two pitches through the use of a foot pedal [4].

The structure of a truncated 14-tuning, in terms of its fifth relations, appears below.

$$... \sim Gb-Dg \sim Ab-Eb-Bb-F-C-G-D-A-E-B-F^2-C^2-G^2-D^2 \sim A^2-E^2 \sim ...$$

The size of the wolf fifth ($\sim$) is the same as for a truncated 12-tuning derived from the same underlying regular tuning. Unlike truncated 12-tunings, larger truncated tunings have some ambiguity with respect to how enharmonics are assigned. For example, should $C^\#$ be tuned the same as $D_f$ or $Eb$? The same issue arises with pitch classes like $Fb$. This is a fairly esoteric theoretical issue, though, since very little (if any) music uses a range of pitch classes so extreme as to make these questions important.
3.8.5 Wolf Intervals in Truncated Tunings

The main problem with truncated tunings is that they tend to produce some awful sounding intervals. These are called wolf intervals and are those that are between pitches whose classes span a wolf fifth. For example, we have seen that in a size 12 truncated Pythagorean tuning, $w_P \approx -19.6$ mil. This means that the $P5$ between pitches like $[11\ 0]^T$ and $[12\ 0]^T$ is $-19.6$ mil flatter than just, a very large error. Such a fifth is said to “howl” and cannot be held for any appreciable amount of time without inducing a queasy feeling in listeners. Not all wolf intervals sound bad, though. For example, we have seen above that the wolf M3 in a size 12 truncated Pythagorean tuning deviates from just by only $-1.6$ mil, a very small amount. In contrast, the normal Pythagorean M3 deviates from just by $17.9$ mil! For size 12 truncated tunings, wolf intervals become less severe as $R$ approaches $\frac{C_P}{12}$, which is the value of $\Phi$ for 12TET. Of course, for 12TET, no truncation is necessary.

Truncated tunings work well for music that avoids wolf intervals. This type of music typically visits a bounded range of pitch classes that does not span a wolf fifth. Indeed, the vast majority of early baroque keyboard music obeys such bounds since it was intended for instruments using truncated meantone tunings [3, 179-87].

3.8.6 Enharmonic Expectations

Tunings whose size is larger than 12, be they regular or truncated, could cause problems in pieces that rely on enharmony between pitches separated by a d2. Two pitches are enharmonic if and only if they are tuned the same. An example of reliance on such enharmony would be a D♯ and an Eb tied together. Even if such a rarity were to occur, it could be interesting to use different tunings for these pitches in order to highlight the meaning of what is occurring.

In cases where music modulates by a d2 over an extended period of time (for instance, starting in Eb major and ending in D♯ major), it is probably not so important that pitches separated by d2 be enharmonic. Those listeners who have perfect pitch might be slightly disturbed by the failure to arrive back where the music started, but, on the other hand, this could be fun for listeners with perfect pitch: only they would know that the music seemed to return where it started but in fact was slightly altered in the process.

3.8.7 Well Temperaments

Well-temperaments, like truncated tunings, are limited in size, and, in addition, do not have unpleasant wolf intervals. As with many topics in tuning theory, there is much misunderstanding surrounding the notion of a well temperament. Most people are familiar with the term though exposure to Bach’s masterpiece, Das Wohltemperierte Klavier (The Well-Tempered Klavier). Unfortunately, most people assume that “wohltemerierte” is just an antiquated German word for “equal-tempered.” This is not the case, since even during Bach’s
time there was a specific word for “equal-tempered” (gleichschwebend) and so it seems likely that if he meant “equal-tempered” he would have entitled the piece accordingly [19], [14, 59–71,195].

A well temperament is any tuning satisfying the following two criteria. First, that it be a tuning of size 12. A tuning of size twelve tunes a range of 12 pitch classes and tunes all other pitch classes to be enharmonic to ones inside this range. Mathematically, this means that a tuning of size 12 obeys the rule

$$\phi(x_f) = \phi(x_f \mod 12) + 7(x_f \div 12).$$

The second criteria for a well temperament is that all intervals be playable. (This condition is often stated as “playable in all keys,” but we have not defined what a key is, nor is it very easy to define.) The definition of “playable” is of course subjective, but, for example, the wolf P5 that is lowered by a Pythagorean comma from just would probably be universally agreed to be unplayable.

Thus 12TET is a well temperament, but not the only well temperament. It is unique in that it is the only regular well temperament. All other well temperaments are similar to 12TET, but they impart a slightly different character to music in different keys since they are irregular.

Well temperaments can be specified by tabulating values of \( \Phi(x_f) \) over a range of 12 values of \( x_f \). Instead of doublings, it is useful to express \( \Phi(x_f) \) in Pythagorean commas. This is a convenient unit because of a special property of tunings of size 12 which is that for all \( i_0 \),

$$\sum_{i=i_0+11} \Phi(i) = -C_P$$

A proof of this is left to the interested reader.

Several well temperaments appear in Table 3.3, taken from [14, 62]. The abbreviations identifying the temperaments are as follows: W = Werckmeister 1681, Y1 = Young 1800, V = Vallotti circa 1750, L = Lambert 1774, Y2 = Young 1800, N = Neidhardt 1724, ET = 12TET. Although we have chosen to model diatonic tunings as transposable, historically, well temperaments were not thought of as transposable. This is why Y1 and V are listed as separate tunings even though they are just transpositions of each other.

<table>
<thead>
<tr>
<th>( \Phi(x_f) )</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>0</td>
<td>0</td>
<td>-4</td>
<td>-4</td>
<td>-4</td>
<td>0</td>
<td>0</td>
<td>-4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Y1</td>
<td>0</td>
<td>0</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
</tr>
<tr>
<td>V</td>
<td>0</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
<td>-6</td>
</tr>
<tr>
<td>L</td>
<td>0</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
<td>-7</td>
</tr>
<tr>
<td>Y2</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
</tr>
<tr>
<td>N</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
</tr>
<tr>
<td>ET</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
<td>-12</td>
</tr>
</tbody>
</table>

Table 3.3: Some well temperaments (in units of \( C_P \))

Figure 3.11 shows these well temperaments in terms of \( \Phi(x_f) \) and \( \Theta(x_f) \). (The Valotti temperament is excluded because of its similarity to the first Young
temperament.) The points on the graphs in Figure 3.11 are shaped differently in order to indicate what pitch class they belong to. Figure 3.10 illustrates this mapping of shape to pitch class. This shaping, reminiscent of a clock hand, is designed so that multiple points may lie on top of each other and still be distinguished.

\[
\begin{array}{cccccccccccc}
2 & -1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 0 & E & C & D & 0 & 3 & 6 & 9 & 12
\end{array}
\]

Figure 3.10: Mapping of point shape to pitch class

### 3.8.8 Review of Diatonic Tunings

For situations where large (or infinite) tuning size is not a problem, such as in the use of electronic instruments, a regular tuning is a good choice. If tuning size is a problem (such as in the use of acoustic instruments), but the piece avoids wolf intervals by visiting only small range of pitch classes (usually 12), a truncated tuning is a good choice. If tuning size is a problem and the piece visits a large range of pitch classes, a well temperament is a good choice. Historical factors may influence these decisions as well: it may be desirable to render a piece in the tuning that the composer intended it for. In addition, irregular temperaments might be employed to create subtle effects of key coloration.

Still, in general, for electronic instruments, regular tunings seem like the best we can do as far as our goal of just M3 and P5. Nonetheless, it is natural to wonder if we can somehow do better, perhaps by allowing dynamic intonation, multiple tunings of the same pitch depending on context. This might allow just M3 and P5 to be simultaneously possible. This idea will be the subject of the next section.

### 3.9 Triadic Tunings

#### 3.9.1 The Representation of Triadic Pitch

Previously we have seen that it is impossible to achieve just M3 and P5 in any diatonic tuning. It was suggested that multiple tunings of the same pitch, i.e. dynamic intonation, be allowed in order to admit this possibility. But we cannot allow this since our theory requires a tuning to be a function, which requires the production of a unique output from any given input. The answer to this dilemma is to switch to a richer description of pitch class. This allows us to discuss dynamic intonation while still using most of the theory we have developed for static intonation (tuning).

Diatonic pitch classes (traditional or FRV) are identified by a distance in P5. We would like to be able to identify triadic pitch classes in terms of a two-dimensional distance in P5 and M3\(_{-1}\), where M3\(_{-1}\) is a new kind of M3
not defined in terms of $P_5$. We will first enrich traditional representation with this extra information and then enrich FRV similarly.

Let’s look at an example of what this representation will need to do and then present a way to accomplish it. The diatonic pitch class $E$ identifies itself as $4P_5$ away from $C$. We need to differentiate this $E$ from the one that is $M3_{-1}$ away from $C$. To accomplish this differentiation, let us call $C + 4P_5$ “$E_0$” and $C + M3_{-1}$ “$E_{-1}$.”

The general rule of this new representation is that the familiar diatonic pitch classes are subscripted by zero, and all the classes that are $M3_{-1}$ above them retain the letter name one would expect but are subscripted by $-1$. For example, the $F\sharp$ that is $M3_{-1}$ above $D_0$ is $F\sharp_{-1}$, and a major triad based on $G3$ would consist of $G3_0$, $B3_{-1}$, and $D4_0$. This notation extends to subscripts other than $-1$. For example the pitch class $M3_{-1}$ below $C$ is $A\flat_1$.

Other theoreticians use a system of subscripting that looks similar to this but means something quite different. This other system has its roots at least as far back as Helmholtz and Ellis [12, 276-8]. New to this work is the idea that a subscripted pitch is a triadic pitch, not a frequency ratio. Conventionally, $X^i_n$ means $\tau_P(X) + nC_S$ where $\tau_P(X)$ is the Pythagorean tuning of diatonic pitch $X$ and $C_S$ is the syntonic comma. The notion of triadic pitch is far more theoretically powerful because it does not imply a specific tuning. As we shall see below, the conventional meaning of subscripting implies just triadic tuning. Our use of subscripting is more closely related to Lindley and Turner-Smith’s concept of the pitch class relation for $M3$ (“III” in their notation) [14, 24].

Traditional notation will be enriched to describe triadic intervals in the following manner. The triadic interval from triadic pitch $X_i$ to $Y_j$ is the diatonic interval from $X$ to $Y$ subscripted by $j - i$. For example, the interval from $D3_0$ to $F3_1$ is $m3_{-1}$.

We will enrich the FRV representation by adding an initial “thirds” component to the vector, forming a third-fifth-register vector (TFRV) $[x_t \ x_f \ x_r]^T$ that represents the interval

$$x_tM3_{-1} + x_fP5 + x_rP8$$

or the pitch that is that distance away from $C0_0$. For example, A major triad based on $C0_0$ would consist of $[0 \ 0 \ 0]^T$, $[1 \ 0 \ 0]^T$, and $[0 \ 1 \ 0]^T$. The three-dimensional vector representation of TFRV represents a significant and original extension to Regener’s work in two-dimensional pitch representations.

A traditionally represented triadic pitch or interval $X_i$ can be converted to a TFRV $y$ in the following manner. Consider $X_i$ as

$$X_i = X_0 + iP1_1.$$ 

Note how the subscript $i$ functions very similarly to the way sharps do in diatonic notation: it defines a multiple of the smallest positive prime interval. Diatonically, this is $A1$, and triadically it is $P1_1$. By the way, what exactly is this strange interval $P1_1$? Well, for example, it is the interval from $E_{-1}$ to $E_0$. 

39
To convert \( X_0 + iP_1 \) to a TFRV, we need to express \( X_0 \) and \( iP_1 \) as TFRV. \( X_0 \) is easy to express as a TFRV; if \( X \) has FRV \([x_f \ x_r]^T\), it is just \([0 \ x_f \ x_r]^T\). An easy way to calculate \( P_1 \) in TFRV form is to reckon it as the difference between \( M_{3_0} \) (our old, P5-based notion of M3) and \( M_{3_{-1}} \) (our new notion of M3). By this method,

\[
P_1 = M_{3_0} - M_{3_{-1}} = [0 \ 4 \ -2]^T - [1 \ 0 \ 0]^T = [-1 \ 4 \ -2]^T.
\]

Therefore, the TFRV \( y \) for the traditional triadic pitch or interval \( X_i \) can be expressed as follows:

\[
y = [0 \ x_f \ x_r]^T + iP([-1 \ 4 \ -2]^T).
\]

### 3.9.2 Just Triadic Tuning

A just triadic tuning in which just M3 and P5 are possible is easy to construct.

\[
\tau(x) = \frac{5}{4} x_t + \frac{3}{2} x_f + x_r.
\]

This maneuver of adding a “thirds” dimension seems to magically solve all our problems. But as is always the case with tuning (and life), you can’t get something for nothing; it is all about compromises.

One potential problem with the just triadic tuning is that, like regular diatonic tunings with irrational \( y \), its size is infinite. Mathematically, this means that just triadic tuning is a one-to-one mapping, i.e. any change to \( x \) results in a change of value for \( \tau(x) \). To show this, let us look at \( \tau(x) \) (not its log),

\[
\tau(x) = \left(\frac{5}{4}\right)^{x_f} \left(\frac{3}{2}\right)^{x_t} 2^{x_r} = 5^{x_t} 3^{x_f} 2^{x_r}.
\]

Since 2, 3, and 5 are prime, any change to their exponents will result in a different value of \( \tau(x) \). Any change to \( x \) results in a change to these exponents, so \( \tau(x) \) is one-to-one and the just triadic tuning has infinite size. Of course size is not a problem for implementation on electronic instruments.

A big problem with just triadic tuning is that much of the repertoire relies on enharmony between pitches separated by \( P_1 \), like \( A_0 \) and \( A_1 \). This enharmony does not exist in just triadic tuning since

\[
\tau(P_1) = \tau([-1 \ 4 \ -2]^T) = 81/80 = C_S \approx 17.9 \text{ mil.}
\]

In other words, \( P_1 \) is tuned to a syntonic comma.

Let us examine how this enharmony is required in a simple musical example. Consider the chord progression I-IV-ii-V-I, whose diatonic pitch classes are indicated in Table 3.4 for C major.

How shall we assign triadic pitch classes to this progression?

Harmonically, we would like to think of major chords as having a M3\(_{-1}\) between root and third and minor chords as having a M3\(_{-1}\) between third and fifth. We will assume that diminished triads are tuned as two stacked m3\(_1\), and
augmented triads are tuned as two stacked M3_1. Examples of subscripting using these harmonic criteria appear in Table 3.5.

Melodically, it would be nice to keep the triadic pitch class the same if two consecutive notes have the same diatonic pitch class.

As listed in Table 3.6, there is a solution to the harmonic and melodic constraints we have discussed above, but unfortunately it results in a P1₁ slip, a phenomenon whereby all parts become transposed by a P1₁ (usually downward). This is often called a comma slip, but it seems more appropriate to discuss these phenomena without reference to a specific tuning.

Table 3.6: I-IV-ii-V-I with P1₁ slip

It is possible to imagine music in which P1₁ enharmony is not relied upon and therefore slips cannot occur. For example, music restricted to the five interlocking triads listed below cannot slip.

Unfortunately, this would only be useful for music in C major that stayed strictly within the key and did not use triads based on B or D. Little if any such music exists. Richer sets of triads can be accommodated, for instance the set of seven interlocking triads below.

Still, this triad set would only be useful for music in A minor that stayed strictly within the key and did not use triads based on G, G♯, or B. Again, little if any such music exists.

Arbitrarily large “slip-free” triad sets can be formed. This is an important point, because many people assume that in order to avoid slippage, music must be limited to small triad sets such as the examples above. The problem with slip-free triad sets is not that they are inherently small (they are not),
it is that they do not suit the needs of traditional Western music. It turns out that \( P_1 \) enharmony is a critical aspect of Western music, because a single pitch must be able to have multiple harmonic functions.

For example, if a piece is to have a \( C_0 \), \( F_0 \), and \( G_0 \), it cannot have a \( D \) that is \( m3_1 \) below the \( F_0 \) (\( D_{-1} \)) and \( P_5_0 \) above the \( P_5_0 \) (\( D_0 \)). This \textit{supertonic problem} is one of the well-known consequences of slip-free triad sets [18, 61] [27].

Having decided that slip-free triad sets are not practical, how else can we avoid slippage? One option is to use tunings that enharmonically identify \( P_1 \). Unfortunately, this is the same thing as returning to a diatonic representation of pitch and using a diatonic tuning.

Another option is to relax our harmonic and/or melodic constraints. Harmonically, we could allow \( M3_0 \) (and its co-conspirator, \( m3_0 \)) back into the picture for occasional appearances. We could even allow \( P5_{-1} \) to appear on rare occasions, even though it is best not to tamper with this all-important interval. Melodically, we could allow a repeated note to have two different subscripts. We could even allow \textit{melismatic intonation}, the adjustment of the intonation of a note while it is sounding. This way, a note held between two chords could have its subscript change as its function changes [30], [17, 143]. A change in subscripting to a held or repeated note could be disconcerting, especially if it were in an outer voice.

\( P_1 \) slippage aside, another problem is that it is not always easy to assign triadic pitch classes (subscripts) to the notes of a piece of music. The primary reason for this is that not all notes have a triadic (or even harmonic) function. For example, what is the harmonic function of an “A” passing tone between a C major and a G major chord, since it belongs to neither? Another example is the seventh of a dominant seventh chord. It certainly has harmonic function, but how should it be subscripted triadically? In a dominant seventh on \( G_0 \), is it \( F_0 \), functioning as a subdominant to \( C_0 \), or is it \( F_1 \), functioning as an \( m3_1 \) stacked on top of \( D_0 \)?

So not only is it difficult to avoid \( P_1 \) slips, it is also difficult to know how we would like to assign subscripts even ignoring them. In other words, we can’t always get what we want, and we don’t always even know what we want! In some sense this is good, though, because if it is unclear how to subscript a pitch, we can let the avoidance of slippage dictate the subscripting.

At some point, judgments about subscripting cannot be made without knowledge of the triadic tuning that will be applied to them. For instance, without knowing just how bad an \( M3_0 \) sounds, it is difficult to decide how much it should be avoided in favor of \( M3_{-1} \).

When such painstaking judgments are being made, it becomes unclear whether the abstraction of triadic pitch class is really offering anything, and in fact one should not just take recourse to “raw” intonation rather than tuning. For instance, instead of assigning a subscript to each note, an intonation instruction in the form of a deviation from Pythagorean tuning could be given.

Indeed, for the most demanding applications this raw intonation approach is probably the best. In many circumstances, it makes sense to use a diatonic tuning for most notes, and carefully choose the intonation only in parts of a piece
where it is most important, for instance in slow sections or final, held chords.

### 3.9.3 Other Triadic Tunings

The notion of a triadic tuning can be useful to resolve certain theoretical dilemmas that arise in diatonic tunings in which a d4 is tuned such that it forms a better M3 than M3 itself. For example, in Pythagorean tuning, M3 are $C_S \approx 17.9$ mil away from just, but d4 ($[-8\ 5]^T$) are only $C_S - C_P \approx -1.6$ mil away from a just M3. Thus it is desirable to render a notated M3 like C3−E3 as C3−F♯, although at other times the “real” E3 would be desirable, for instance in the P5 A3−E3. The solution to this dilemma is to construct a Pythagorean triadic tuning

$$\tau(x) = \left(-8 + \frac{3}{2} \right) x_t + \frac{3}{2} x_f + x_r.$$ 

Reformulating the situation in this way shows that this is simply an instance of triadic tuning, not some aberrant use of Pythagorean diatonic tuning. This reformulation is also useful in understanding the approximation to just intonation offered by large equal temperaments such as 53TET, which have attracted theoretical attention since the early 17th century [14, 148]. The triadic formulation of 53TET is

$$\tau(x) = \left(\frac{17}{53}\right) x_t + \left(\frac{31}{53}\right) x_f + x_r.$$ 

As one might imagine, the problems of P11 slips still apply to these other triadic tunings.
Figure 3.11: Comparison of well temperaments
Chapter 4

Intonation Software

This section describes Helm, the software suite that was developed to realize pieces using dynamic intonation. (As the astute reader may be able to guess, it is named after the great acoustician and tuning theorist, Hermann Helmholtz.) Helm operates on score files written in a simple Score Description Language (SDL) that gives pitch, duration, and intonation information. It can create Standard MIDI Files (SMF) [5] from these scores, and then the SMF can be played back to realize the piece. Helm can also follow along in the score as it "listens" to the MIDI data from a performer's keyboard, transmitting a retuned version of this MIDI data to a MIDI synthesizer in real-time.

Below, we will first discuss two topics fundamental to both the SMF and score following modes of operation. These topics are the score representation used by Helm (SDL) and the technique used to tune MIDI synthesizers. Then we will proceed to discuss the SMF and score following modes of operation.

4.1 The Score Description Language

Before diving into the details of the Score Description Language, we will give an example of a score fragment in traditional and SDL notation. Figure 4.1 is the traditional notation for the first phrase of Bach's realization of the chorale O Welt, ich muss dich lassen [1, 205]. Figure 4.2 is the SDL notation for the same phrase.

Table 4.1 shows the syntax (in Backus-Naur Form) for the score description language. The definitions of the elements Integer and Real is not repeated here since they are defined conventionally.

We will define the semantics of this language by explaining the rather straightforward manner in which it corresponds to traditional notation. A Score is a score for NumParts monophonic parts, at Tempo quarter notes per minute. A System has no musical meaning; it is merely a syntactic convenience to break up the input into short lines. Each Part consists of sequential events, each of which has a Duration. A Part with PartNum not between 0 and NumParts – 1
A *Duration* can have the following meanings: thirty-second, sixteenth, eighth, quarter, half, whole. These are indicated by the first letter of their name. A *Duration* can be “dotted” by using the capitalized version of the base duration. For example a dotted eighth is indicated by ‘E’ and a normal eighth by ‘e.’ A *Rest* consists solely of a *Duration*, but a *Note* has a *Pitch* and an *Intonation* annotation.

A *Pitch* is indicated in the same manner used throughout this work, except that, due to the constraints of the ASCII character set, the flat sign (b) is replaced by a lowercase bee (‘b’). *Intonation* can be an integer, indicating a subscript to be used in just triadic tuning, or it can be an underscore (‘_’), indicating 12TET. Thus the example in Figure 4.2 starts in 12TET and switches to just triadic tuning.

This language is bare-bones; for example, there are no provisions for bar lines, slurs, articulations, or dynamics. Nonetheless, it does provide enough information for simple MIDI file realizations and score following of a piece.
4.2 MIDI Tuning Technique

A MIDI Note On message does not specify a frequency. Rather it specifies a pitch in a 128 element pitch set whose elements are identified by integers in the range $0 \ldots 127$. All MIDI synthesizers provide the relative tuning $\tau(n) = n/12$ with a range of reference frequencies such that $n = 69$ (the pitch normally associated with A4) is tuned to around 440 Hz.

In addition to this default tuning, many synthesizers can map the MIDI pitch set to frequency ratios defined by the user. These can usually be defined via front panel controls or System Exclusive messages. Unfortunately, although there is a proposed MIDI Tuning Standard, all manufacturers implement these tuning capabilities using different System Exclusive message formats.

A design goal of Helm was to be synthesizer independent. As such, the System Exclusive tuning technique was not a viable option. In addition, the System Exclusive technique can only be used to realize pieces with less than 128 distinctly tuned pitches. In practice, this would probably be sufficient, but it is best not to have such limitations.

Instead of the System Exclusive technique, a technique was used whereby the default tuning ($\tau(n) = n/12$) is dynamically adjusted via Pitch Bend messages. A Pitch Bend message sends a new pitch bend state to a single channel. A pitch bend state is an integer in the range $-8192 \ldots 8191$. At time $t$, the tuning of a pitch $n$ on channel $c$ is

$$\tau(n) = n/12 + a \left( \frac{P(c,t)}{8192} \right)$$

where $P(c,t)$ is the pitch bend state of channel $c$ at time $t$ and $[-a,a)$ is the pitch bend range.
The pitch bend range, as defined by \( a \), is adjustable via a standard System Exclusive message. We use \( a = 1/6 \), which is the default on most synthesizers anyway. (For reference, \( \tau(M2) = 1/6 \) in 12TET.)

Thus if we wish to play a pitch with frequency ratio \( r \), we find the pitch \( n \) whose unadjusted tuning (\( n/12 \)) is closest to \( r \) and then we find the pitch bend value \( p \) that most closely adjusts this tuning towards \( r \). More formally,

\[
n = \text{rnd}(12r)
\]

\[
p = \text{rnd} \left( \frac{8192}{a} \left( r - \frac{n}{12} \right) \right) = \text{rnd}(49152(r - n/12))
\]

where \( \text{rnd}(x) \) is any rounding operation, such as

\[
\text{rnd}(x) = \lfloor x + 1/2 \rfloor.
\]

Once we find \( n \) and \( p \), we send out a pitch bend of \( p \) followed by a note on of \( n \) on the same channel.

This uses only about a quarter of possible pitch bend range since \( r - n/12 \) is limited to \([-1/24, 1/24)\). This “headroom” was left so that future versions of the program could incorporate melismatic intonation, the adjustment of the intonation of a note while it is playing. Melismatic adjustments could exceed the pitch bend range if headroom is not provided.

With \( a = 1/6 \), pitch bends can adjust tuning in increments of about 0.02 mil. For all practical purposes, this provides continuous resolution, because this is a very small frequency ratio. In practice, tuning resolution is limited not by the finite precision of pitch bend but by synthesizer implementation, which often throws away some of the lower bits of pitch bend.

In this sense, Helm is somewhat synthesizer dependent since it relies on precise pitch bend response. Although MIDI puts an upper limit on this precision by limiting pitch bend messages to 14 bits, it does not put a lower limit on it. For instance, some synthesizers are known to discard the entire lower 7 bits of pitch bend. This leads to unsatisfactory resolution for use with Helm. This work used the Creative AWE32 and E-mu Morpheus synthesizers, both of which had adequate pitch bend resolution. Although the Morpheus was not tested, the AWE32 was found to have a resolution of about 0.3 mil.

In order to use the pitch bend tuning technique, simultaneous notes that require distinct pitch bends must be played on different channels. This limits the polyphony to 16 notes in the worst case. Larger polyphony can be accommodated if some of the simultaneously sounding notes are separated by a doubling since these notes can be played on the same channel.

Certain multitimbral situations can be tricky to realize using this technique. These are situations in which the sum of the polyphony of each timbre is greater than 16. For example, consider a piece that has a 9-note chord in one timbre followed by a 9-note chord in another. The instantaneous polyphony of this piece is only 9, but the sum of the polyphonies of each timbre is 18. In situations like this, program (timbre) changes as well as pitch bend changes must
be sent between notes of the piece. This can clog the MIDI stream, and some devices may not switch programs quickly enough to make such adjustments feasible. If the sum of the polyphonies of each timbre is 16 or less, timbres can be statically assigned to channels and there is no problem.

4.3 SMF Mode

One major function of Helm is to allow the realization of score files as MIDI files. There are four major software components that make this possible. All of these were written in the Java programming language [10]. The first is an SDL (Score Description Language) parser. This reads a score file and turns it into a corresponding data structure in memory. The second major component is the score serializer. This converts the representation of the score from a set of parallel parts with events that have duration to a single series of point-like events in time. The third major software component takes this serial version of the score and converts it a series of MIDI events. The major work of this component is implementing the MIDI tuning technique described above. The fourth major component is the MIDI file writer, which takes a series of MIDI events in memory and turns them into a Standard MIDI File.

The SDL parser does its lexical analysis using the StreamTokenizer utility class of the Java Language. The data structure it uses to represent scores is a nested hierarchy of objects that reflects the nested nature of musical structure. This nested structure can be seen from the Backus-Naur syntax; for example, a letter is part of a pitch which is part of a note which is part of an event which is part of a part which is part of a score. Although SDL has no provisions for it, this data structure can handle melismatic intonation information.

The major challenge of the score serializer is to convert events with duration to a chronological series of events without duration. The most important example of this is converting a single note event to a note on event and a note off event. What is slightly tricky about this task is using a data structure that can serialize events but still maintain associations between them. In particular, when the note on and off are separated due to chronological ordering, it is still important to be able to find the “on” that corresponds to a given “off.” This becomes important during the implementation of the MIDI tuning technique because the note off event must have the same channel as its corresponding note on.

The conversion from a serial score to a series of tuned MIDI events carefully seeks to minimize the number of pitch bend messages that are created. The idea behind this is that some synthesizers may not respond well to the dense MIDI stream that would result from sending pitch bend changes before every note. The most naive implementation of the MIDI tuning technique is to send a pitch bend before every note, rotating through all 16 channels sequentially. This ignores the fact that a pitch bend needs to be sent out before a note of class $c$ only if more than 15 unique pitch classes have been used since the last time a note of class $c$ was used. Pitch class is the relevant datum here because no retuning is necessary to play an octave multiple of a pitch that has had a channel
tuned for it. The algorithm that converts a serial score into MIDI events tries to reuse channels that are already tuned correctly, and when it must retune a channel, it picks one that has not been heavily used, under the assumption that it probably won't come in handy later.

Channel assignment is also affected by the presence of unisons. To understand this, we need to introduce a little history. Initially, the primary application of MIDI was as a way for keyboards to control separate synthesizers. Due to the nature of the keyboard mechanism, a unison can never be played. This means that two Note Ons for the same key are always separated by a Note Off for that key. Because of its heritage as a keyboard control protocol, MIDI does not define the behavior a synthesizer should have in the presence of a unison.

Even if MIDI did define a behavior in the presence of unisons, it could not be defined in a musically satisfying manner. This inherent limitation comes from the fact that it is impossible for a synthesizer to figure out which Note Ons match with which Note Offs. Imagine a MIDI stream of the following three messages, all on the same channel and for the same key, with 500ms between them: Note On with velocity 30, Note On with velocity 120, Note Off, Note Off. How should the Note Ons and Offs be paired? It is impossible to tell, yet the musical result is quite different due to the fact that the two pending Note Ons have widely different velocities.

As a result of the fact that MIDI’s behavior in the presence of unisons is undefined, the two (or more) notes of a unison must be assigned to separate channels. Thus a channel can be reused to play note n with pitch bend p only if its pitch bend state is already p and it is not currently in use playing note n. This seems like a minor point, but unisons are very common in polyphonic vocal music.

The conversion from a series of MIDI events in memory to a MIDI file is accomplished via the Rogus McJava MIDI Library. This is a library that allows the reading and writing of Standard MIDI Files from the Java language. It was written for use in Helm but is completely general, i.e. it is in no way specific to tuning.

In addition to creating a Standard MIDI File from an SDL score, Helm can also create a Csound [26] Score File (.sco file). The conversion from an SDL to a Csound score does not even require serialization of the score, since Csound can accept scores with parallel events. The Csound score can be used to create audio file realizations with stunning audio quality and intonation accuracy, all without the use of any external synthesizer hardware.

### 4.4 Score Following Mode

Besides the creation of MIDI files, the other important function of Helm is to follow along in the score as it “listens” to the MIDI data from a performer’s keyboard, transmitting a retuned version of this MIDI data to a MIDI synthesizer in real-time. There are five major software components that are used to accomplish score following. The first two are the SDL parser and score serializer
already discussed. The third creates an input/output (I/O) map file that describes what the follower is looking for and what it should translate its findings into. Another way of describing this file is a map from the untuned MIDI data of the score (what it will be receiving from the performer) to the tuned MIDI data of the score (what it will be sending to the synthesizer). The fourth and fifth components are the only part of Helm that are not written in Java; they are written in C++ [25] because they need to access operating system calls that deal with MIDI input and output. The fourth major component of the software is the program that reads in the I/O map file and actually performs the score following and remapping. This program is built upon the fifth and final major component, the Rogus McBobus C++ MIDI Library [6].

The serial score to I/O map converter works as follows. There is no timing information in the I/O map file; the score is just segmented into clumps of inputs (Note Ons and Offs) that appear simultaneous according to the score, along with the output they are supposed to map to. This timelessness is due to the fact that the score follower that uses this information is naive and cannot take advantage of timing information. The MIDI output information does not include channel assignment: that is done later by the score follower itself.

Figure 4.3 contains a fragment of an I/O map file for the phrase of Bach introduced in Figures 4.1 and 4.2. A clump is a pair of NOF/NON lines. A NOF line lists the MIDI note numbers of any Note Offs expected from the keyboard around the time of the clump. A NON line lists comma-separated I/O pairs for any Note Ons expected from the keyboard around the time of the clump. The I/O pairs consist of three numbers: the input Note On expected from the keyboard, and the corresponding Note On and Pitch Bend values to be sent out as a result.

So, for example, the second clump contains the Note Offs expected from the first chord (74, 70, 65, and 46) and the mappings for the Note Ons of the second chord (70, 67, 63, and 51). Because the second chord is tuned to 12TET as per the underscores in the input file (see Figure 4.2), the Note Ons of the second chord map to the same Note On value with a Pitch Bend of zero, i.e. the I/O pairs are all of the form $X (X \ 0)$.

As was already mentioned, the score follower is naive. In fact, it is not so much a score follower as an interactive score player. It will never get lost because it basically uses the keyboard input to step through the score. In fact, the major problem with the so-called score follower is that it is easy for the performer to get lost, whereas in real score followers, the problem is just the opposite: it is easy for the computer to get lost! It was decided that since this project was about tuning, not score following, a naive score follower that worked well enough to serve as a proof of principle would suffice.

The score follower has three modes of operation: performance mode, note mode, and chord mode. In performance and note mode, the program uses only the note on information in the I/O map file. When a key is released, it looks up the most recent Note On it sent as a result of a press of that key and sends an appropriate Note Off. Thus the duration of notes is entirely controlled by the performer. In performance mode, the only way to cause a note to be played on the
synth is to press a key that corresponds to a Note On in the current clump. Recall that a clump is a bunch of inputs that appear simultaneous in the score and the outputs that they should produce. When all the keys in the current clump have been pressed, the program moves its “clump pointer” to the next clump.

Note mode is the same as performance mode except that input from any key will be accepted, and the clump will map it to whatever output appears first in the clump. The ordering of outputs in the clump is by part number. So, for example, in performance of a score with part numbering as in Figure 4.2, repeatedly pressing a single key will step through all the notes of the piece with chords being stepped through from top down. To play the piece polyphonically, all that is necessary is to perform the rhythm correctly since which key is depressed does not matter. Chord mode plays out an entire clump (Note On and Note Off outputs) every time it receives a key press. All of the modes map the input velocity through to the output velocity, giving the performer the normal expressive control of a MIDI keyboard. In the case of chord mode, all the notes of a chord have the same velocity as the input keypress that triggered it.

The score following software is implemented on top of the Rogus McBo
gus C++ MIDI Library [6], written by the author and Patrick Pelletier. It was written primarily to support the MIDI programming required in the Brain Opera [15] project, an interactive multimedia performance piece.
Chapter 5

Conclusions

There are three main areas of this research that we will draw to some kind of conclusion here. The first area is the theory of tuning presented. The second area is the dynamic intonation software, Helm. The third area is the application of Helm to experiments in just triadic tuning.

We have presented a theory of tuning that is precise, concise, and provides insight into the structure of tunings of Western music. Included in the theory are various tuning models which may be applied to a wide variety of musics and instruments. The model best suited to Western tunings was the transposable register-zero tuning. This theory also introduced the fifth-register vector (FRV), a powerful mathematical formalization of diatonic pitch and interval. The goals of just P5 and M3 were explained and used as a basis for judging the efficacy of tunings. Transposable register-zero tunings operating on FRV were studied, illuminating some of the problems and compromises associated with diatonic tunings. The third-fifth-register (TFRV) and triadic tunings were presented along with a discussion of P11 slips and other factors limiting their feasibility.

Although this theory is similar to the work of others, including Helmholtz, Regener, Lindley and Turner-Smith, and Blackwood, it offers certain advantages over this previous work. One advantage it offers compared to existing theories is its thorough treatment of register. Other theorists have correctly identified pitch class as the fundamental element on which a tuning operates, but they have failed to present a mechanism by which a tuning of pitch classes can be extended to a real tuning of all pitches. This may seem like a simple process, and indeed this is probably what these theorists assumed, leading them to ignore it, but in fact it involves some subtlety. Regener is the only other theorist who has treated register in a satisfactory manner, although his work is restricted to regular tunings. One of the main theoretical contributions of this thesis is the synthesis of Regener’s “register-smart” representation of pitch with Lindley and Turner-Smith’s treatment of irregular tunings.

Another advantage of the tuning theory presented in this work is the notion of triadic tuning. Other theorists have presented just intonation as a series of direct modifications to Pythagorean diatonic tuning. Thus it seems more like
“raw” (ad hoc) intonation than a form of tuning. In contrast, this work arrives at just intonation as a specific tuning of triadic pitch classes. The assignment of triadic pitch class (subscripting) becomes an endeavor in the theoretical domain of harmonic analysis, not a tweaking of raw frequency ratios. This is methodologically important for a number of reasons. One, it allows for triadic tunings other than just. Two, it maintains the theoretical framework of tuning as a map from pitches to frequency ratios. Three, it allows the problems of just intonation to be framed in terms of enharmony, a concept already present in diatonic theory.

Helm, the dynamic intonation software developed as part of this research provides a flexible platform on which to investigate tuning and intonation. It is the only such platform that exists. It is based upon tuning-annotated score files written in a simple score description language. These score files can be converted into MIDI files. They can also be used by a score follower program to achieve dynamic intonation for keyboard performance, allowing one key to produce many frequencies at different points in the score.

Helm was used to conduct various experiments in just triadic tuning. The tuning theory presented in this work suggests that just triadic tuning is of limited feasibility, but it was investigated nonetheless, for a variety of reasons. One, it is important to actually hear the phenomena that the theory predicts will create problems. Two, that it is the most extreme form of tuning since it actually creates beatless, not just slowly beating chords, and therefore provides an opportunity to assess the desirability of beatlessness in musical context. Three, much debate, but little experimentation, has surrounded the feasibility of just intonation for the realization of traditional repertoire.

The musical materials used for this experimentation were as follows:

Guillaume de Machaut,
*Kyrie* from *Messe de Nostre Dame*, measures 1-8 [3, 130]

Orlando di Lasso,
motet *Ave Regina Coelorum*, measures 1-5 [3, 133]

J.S. Bach,
chorale *O Welt, ich muss dich lassen* [1, 205]

J.S. Bach,
chorale *Wie selig seid ihr doch, ihr Frommen* [1, 205]

These add up to only about 40 measures total, but some interesting conclusions can be drawn even from the experience of working with such limited musical materials.

Not surprisingly, it was difficult to assign triadic functions (subscripts) to notes. In order to avoid slippage, occasional non-just chords were created. The job was made somewhat easier for the first two excerpts because the music was taken from Blackwood [3], who had already given his interpretation of the subscripting.

The results were in no way scientifically studied; they are merely the observations of the author and the impressions he was given by the colleagues for
whom he played the examples. 12TET versions were created for comparison. Many of these experiments can be listened to on the web page http://theremin.media.mit.edu/bdenckla/thesis/main.html.

The beatlessness of the chords was noticeable, but only with certain timbres, and only for fairly long chords. It came as no surprise that beatlessness wasn’t too important for chords that even in 12TET would only beat a few times during their duration. The importance of timbre was quite surprising, though.

For many timbres, beatless and 12TET chords sounded almost identical. One reason for this is that many timbres have beating (detuning) already built in. This is usually accomplished by playing two or more slightly detuned versions of the same note. This beating within tones makes it hard to hear beating (or the lack of it) between tones. Also, many timbres have vibrato built in, which makes beating hard to perceive. Still, many timbres just didn’t seem to make beating show up, but there was no easy explanation for this such as the presence of detuning or vibrato. The timbre that worked best was a simple organ sound with voice doubling at the octave. This voice doubling probably helped because it emphasized higher harmonics, albeit only even ones.

“Mistakes” in subscripting, leading to chords very far from just and/or different versions of the same note within a short time span, were very noticeable. In the examples in question, these situations were avoidable by redoing the subscripting, but there could easily be situations for which no good solution exists.

The theory of this work and the results of tests using its software indicate that just triadic tuning has many difficulties and few benefits. This should not in any way be interpreted as a vindication of 12TET. Many other diatonic and triadic tunings remain to be investigated by listening to a variety of pieces rendered in them. In particular, the meantone tunings seem quite promising because of their small $\Phi$ and $\Omega$, as well as the historical prominence of their truncated versions. The software developed in this work currently only supports just triadic tuning and 12TET, but adding new tunings is a trivial programming task since the underlying structure of the software is completely general.

In conclusion, this work has resulted in a new theory of intonation and powerful new software for intonation research.
Chapter 6

Future Directions

This chapter presents some possible future directions for this work. These include extensions to the software, an improved model of dissonance, a theory of melodic intonation, and novel synthesis techniques related to tuning.

6.1 Extensions to the Software

Some possible extensions to the software are listed below.

- Polyphonic parts should be allowed. This will facilitate the encoding of keyboard music. Currently, scores are composed of monophonic parts.

- Different tunings should be allowed. Currently only just triadic and 12TET are supported.

- Melismatic intonation should be allowed. This will allow the frequency of a note to change while it sounds.

- The score follower should be more sophisticated, i.e. more forgiving in its handling of mistakes. This probably involves giving it some notion of rhythm.

- Bar lines, slurs, articulations, and dynamics should be allowed. This will allow more sophisticated MIDI file realizations to be produced. Some of this information could even be used by a very sophisticated score follower.

- The Pitch Bend minimization algorithm should be changed so that when it needs a new channel, it picks one that hasn’t been used in a large number of notes. Currently, it picks one that hasn’t been used heavily since it was assigned. This causes channels that have recently been assigned to be quickly reassigned, causing excess Pitch Bend messages to be sent in what might be called “channel thrashing.”
Facilities for automating the intonation annotation of scores should be added. Adding this feature might well require the development of some new theory. It might also shed some light on the issue of causal versus non-causal intonation decisions, i.e. how well can intonation decisions be made in real-time. If good intonation decisions can be made causally, facilities for real-time intonation should be added.

6.2 Extensions to the Theory

The tuning theory presented in this thesis is based on the hypothesis that beat rate determines dissonance for chords that are considered consonant. This hypothesis leads to the establishment of beatless P5 and M3 as the goals of intonation. But a more sophisticated view of dissonance is probably necessary. For example, the prevalent use of detuning as a synthesis technique suggests that at least when it is at a slow rate, beating has some desirable affects.

In addition, it would be desirable to have a model of dissonance that covers grossly dissonant cases, such as diminished seventh chords, or even wolf fifths. As it is, beating models dissonance only for chords that are considered consonant. A model of gross dissonance could help establish limits of how far away from beatless one can go while still having acceptable P5 and M3. In addition, it might shed light onto what particular tunings are desirable for dissonant chords, since surely not only the fact that they are dissonant, but how dissonant they are, and what the “color” of this dissonance is, is important to their musical meaning.

Another feature of an advanced model of dissonance would be its ability to explain timbre’s role in dissonance. In particular, it would be nice to have a model of the somewhat puzzling masking of beating by certain timbres. Along the lines of timbre, other considerations such as the interaction of vibrato with tuning and the hairy subject of intonation in large ensembles, where clouds of frequencies per pitch rather than individual frequencies occur, should be addressed.

Further research into current work in psycho-acoustics might yield some help in developing a more powerful theory of dissonance, although much work in this field studies the perception of sinusoidal dyads only, a case much simpler than real music.

Even if dissonance and consonance were completely understood, they are fundamentally harmonic phenomenon, so an important extension to the tuning theory of this work would be the consideration of melodic criteria. The most obvious and oft-cited tension between harmonic and melodic goals in intonation surrounds the leading tone. In its harmonic function, it should be close to 5/4 the frequency of the dominant. In its melodic function, many musicians like to make it lead into the tonic by playing it quite high, even higher than it is in 12TET. Melodic considerations like this one should be part of any complete tuning theory.
6.3 Tuning and Synthesis

It is the author’s suspicion that the reason why some timbres did not benefit much from just intonation in his experiments is that they were produced from electronic, rather than physical sources. Tones from polyphonic electronic instruments do not interact: they superpose in a perfectly linear fashion. In contrast, it is possible that tones played on a real polyphonic instrument such as the violin interact with each other due to physical nonlinearities in the instrument body. These interactions could serve to change the character of the sound itself depending on whether the tones were harmonically or slightly inharmonically related, making just intonation “matter” more. An investigation into the possibility of existence of such nonlinearities, and their simulation in synthesis could be an interesting extension of this work.

Another future direction related to synthesis is the idea of sounds with tempered spectra. Much of tuning theory is concerned with bringing the frequencies of tones in alignment with the harmonics of others. Why not turn this problem on its head and concern ourselves with aligning the harmonics of tones with the set frequencies of a given tuning? Early experiments have been conducted by the author in which the partials of a tone are stretched so as to be slightly inharmonic and better aligned with the fundamentals of tones in 12TET. Unfortunately such stretching results in beating within the tone itself, due to the phenomenon of second-order beats. Although this beating is quite strong, it can be made much slower than the beating normally associated with an M3 in 12TET.
Bibliography


