The Origin and Well-Formedness of Tonal Pitch Structures

Aline Honingh

The Origin and Well-Formedness of Tonal Pitch Structures

ILLC Dissertation Series DS-2006-05

Institute for Logic, Language and Computation

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
phone: +31-20-525 6051
fax: +31-20-525 5206
e-mail: illc@science.uva.nl
homepage: http://www.illc.uva.nl/

The Origin and Well-Formedness of Tonal Pitch Structures

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.mr. P.F. van der Heijden
ten overstaan van een door het college voor
promoties ingestelde commissie, in het openbaar
te verdedigen in de Aula der Universiteit
op vrijdag 20 oktober 2006, te 12.00 uur

door

Aline Klazina Honingh

geboren te Broek in Waterland

Promotores:
Prof.dr. R. Bod
Prof.dr. H. Barendregt

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

This research was supported by the Netherlands Organization for Scientific Research (NWO) in the context of the Innovation Impulse programme "Towards a Unifying Model for Linguistic, Musical and Visual Processing".

Printed and bound by PrintPartners Ipskamp.

ISBN-10: 90-5776-156-4
ISBN-13: 978-90-5776-156-0

Contents

Acknowledgments

Een woord van dank aan de mensen die er aan bijgedragen hebben dat dit proefschrift nu is zoals het is. Ik wil graag mijn promotor en begeleider Rens Bod bedanken, allereerst voor het feit dat hij mij aangenomen heeft voor deze AIO baan met daarbij het vertrouwen dat ik iets kon bijdragen aan een tot dat moment mij nog onbekend wetenschapsgebied. Ik heb veel vrijheid gekregen zodat ik me kon richten op het onderwerp van mijn interesse, maar ook kreeg ik, op cruciale momenten wanneer ik door de bomen het bos niet meer zag, de sturing die ik nodig had. Rens heeft een enorm aanstekelijk enthousiasme dat me altijd weer kon motiveren. Rens, heel erg bedankt voor de fijne samenwerking, begeleiding en inspiratie. Mijn tweede promotor, Henk Barendregt, ben ik in de eerste plaats dankbaar voor het accepteren van het promotorschap, iets wat geenszins vanzelfsprekend was aangezien ik dat hem een jaar geleden pas gevraagd heb. De intensieve reeks afspraken die we gehad hebben, zijn heel waardevol geweest voor het uiteindelijke resultaat. Bedankt.

Besides my two supervisors, also several other people have taught me a lot and influenced my work. I am grateful to Thomas Noll, for the private course on mathematical music theory during a bus trip in Italy; for the discussions on 19-tone equal temperament in Paris; for the discussions during my visit to Berlin; and for all the feedback given on my work. I am grateful to David Meredith, who read and commented on my whole thesis, and whose work was of great inspiration over the last four years. Especially, my chapter on pitch spelling had benefited a lot from his dissertation on the subject. I want to thank Kamil Adiloglu, Elaine Chew, Nick Collins, Jan van de Craats, Peter van Emde Boas, Jörg Garbers, Dion Gijswijt, Henkjan Honing, Benedikt Löwe, Frédéric Maintenant, Michael McIntyre, Wim van der Meer, Dirk-Jan Povel, Remko Scha, Stefan Schlobach, Michiel Schuijer, Leigh Smith, Paul Tegelaar, Dan Tidhar, Leen Torenvliet, Henk Visser, Anja Volk, Frans Wiering and Menno van Zaanen for helpful suggestions and discussions about my work. My research has furthermore benefited from the kind correspondence with David Benson, Peter Cariani, Paul Erlich, Ernst Terhardt, and many people from the 'Alternate Tunings Mailing List'.

Neta, I have enjoyed working together in the project during the time that we spent together in Amsterdam and Cambridge. I am grateful to Alan Blackwell for inviting me as a visiting scholar to Cambridge, and to Ian Cross and all the people of the science and music group for contributing to the unforgettable time I had in Cambridge.

I have very much enjoyed our monthly reading group, where people with various backgrounds came to discuss topics within the common field of interest: music. I want to thank all members of this reading group for their contributions and enthusiasm.

Het ILLC is altijd een heel plezierig instituut geweest om te werken. Mijn dank gaat uit naar Frank Veltman, Ingrid van Loon, Marjan Veldhuisen, Tanja Kassenaar, Jessica Pogorzelski en René Goedman voor alle hulp en de fijne werkomgeving. Yoav en Jelle, dank jullie wel, voor de gezelligheid, maar ook voor het geduld bij het uitleggen van lastige mathematische en computationele problemen. Ik ben Neta, Merlijn, Yoav, Sieuwert en Olivia, mijn kamergenoten in chronologische volgorde, dankbaar voor de sfeer op de gezelligste kamer van het instituut. Een paar speciale woorden voor Merlijn: heel erg bedankt voor de gezellige tijd die we hier samen hebben doorgebracht, voor al het 'gekwebbel' maar ook voor de fijne gesprekken over complexiteit, muziek en het leven. De hele werkvloer droeg bij aan de goede sfeer: Brian, Clemens, Eric, Fenrong, Joost, Leigh, Nick, Olivier, Reut, Stefan, Ulle, en alle anderen: heel erg bedankt!

Ook buiten de werkkring is er een aantal mensen geweest die me de afgelopen vier jaar geïnspireerd en gemotiveerd hebben. Als grote bron van inspirate wil ik als eerste Bas Pollard noemen, van wie ik veel geleerd heb over intonatieproblematiek maar ook over muziek in het algemeen, al lang voordat ik met dit onderzoek begon. Ik bedank Nick Devons voor de fijne vioollessen die vaak een therapeutische werking hadden en ervoor zorgden dat ik muziek áltijd leuk ben blijven vinden, ook al vlotte het onderzoek op dat moment misschien niet zo.

En dan zijn er nog de mensen die me gesteund hebben en voor prettige afleiding hebben gezorgd in de avonduren en weekenden. Bedankt lieve NSO vriendinnen Elske, Margriet en Janneke; CREA vriendinnen Simone, Jeantine en Nienke; Winston kwartet Tessa, Matthijs en Maurice ("o, wat speel ik gevoelig"); Etain trio, Marjolein en Roeselien - dank jullie wel voor alle gezellige etentjes, fijne gesprekken en muzikale hoogtepuntjes. Tessa, behalve een strijkkwartet-vriendin was je ook een hele fijne hardloop- en praat-vriendin. Bedankt Ydeleine en Dille, gezellig als altijd; Jantien, met wie ik inmiddels in vier verschillende orkesten heb gespeeld; Aafke, met wie ik vriendinnen was vanaf dag één van de natuurkunde studie; en Hylke (Koers), gezellig samen ontspannen in vele moeilijke yoga houdingen. Heel erg bedankt allemaal voor jullie vriendschap!

Lieve pap en mam, Berenda, Alwin en Quinten, Nelleke en Pim, en Johan, bedankt voor alles, voor de warmte zoals alleen een familie die kan geven. Tenslotte, lieve Hylke, bedank ik jou. We hebben dit promotietraject samen van alle kanten meegemaakt. Je was er altijd, als inspiratie, steun en ook als afleiding. Samen de bergen in, samen muziek maken, samen promoveren. Dank je wel voor je energie, relativeringsvermogen, rust, vertrouwen en liefde - dank je wel voor alles.

Amsterdam Aline
Augustus, 2006.

Chapter 1

Introduction and musical background

1.1 Questions to address in this thesis

Music has a long history, being closely related to language and dance, and being part of every culture. Since tonal pitch structures such as scales and chords have evolved in music tradition, it is not always clear where their origin lies. For example, what is the reason that the Western diatonic scale consists of 7 notes, the Japanese pentatonic scale of 5 notes, etc.? Are these numbers arbitrary and have they arisen from different cultures, or are these numbers related and have they arisen from a common origin? Many people believe that the latter hypothesis is true and much research has been done in a variety of areas to come up with a possible answer. A research area of 'evolutionary musicology' exists to study the analysis of music evolution, both its biological and cultural forms, see Wallin, Merker, and Brown (2000) and Mithen (2005). Sethares (1999) proposed that consonance depends on timbre and concluded that scales in different cultures have therefore arisen from the timbre of their musical instruments. Another possible answer to the question on the origin of scales has to do with equal temperament. Equal tempered scales have been constructed for several reasons, for example, to approximate certain ratios from just intonation, and for its modulation properties. Finally, there are investigations in the well-formedness or geometrical goodness of scales (Carey and Clampitt 1989), which has led to models that cover large numbers of existing scales. This thesis discusses the latter two approaches to investigate a common origin of musical scales. Besides explaining a possible origin, these approaches may also serve as an evaluation of certain existing scales, and finally, the scales resulting from these approaches can be interpreted as suggestions for new scales that have not been explored until now. Within these two approaches, the representation for pitch structures that is used, is a geometrical tone space. A new notion of well-formedness is proposed and discussed, after which two final chapters are dedicated to computational applications thereof.

The term 'pitch structures' from the title, refers foremost to musical scales, whose origin and well-formedness are addressed (ch. 3,4). Furthermore it refers to chords whose well-formedness is also addressed (ch. 4), which is in turn used to model the preferred intonation of these items (ch. 5). Finally, the term pitch structures refers to a lesser extent to the music-theoretical notion of harmonic reductions which is also interpreted in terms of well-formedness (ch. 4).

Methodological preliminaries that are used in this thesis include mathematics, computation and empirical testing. The question that may arise is: in what field does this research fit? The field that is known for music research is musicology, and musicology does not normally use mathematics and computations, but merely studies the history and practice of music. However, since as long ago as Pythagoras there has been interest in the mathematics of music and it has played a major role in the development of understanding of the mechanics of music. Furthermore, mathematics has proved to be a useful tool for defining the physical characteristics of sound, and also abstractly underlies many recent methods of analysis. In the first half of this thesis mathematics is used to learn more about temperament systems in music. Mathematics can also be used to model a specific phenomenon in music; the model can in turn be used for making predictions. It is here where computation comes in: the conversion of these mathematical models into algorithms. Since the seminal work by Longuet-Higgins (1976), there has been an increasing interest in the computational modeling of music cognition. Therefore, this thesis fits both in the research area of mathematical musicology as well as in the area of computational musicology. Finally, we will test our mathematical models and computational algorithms, on empirical, musical data, such as scales, chords and musical scores, which means that our research fits also into the field of empirical musicology. We will use the same musical scores for our experiments as employed by other researchers, thus allowing for proper systematic comparison.

Since this thesis has been written in a research group situated at the faculty of science, I will assume that the readers have some background in mathematics, but not necessarily a background in music. Sections 1.2 to 1.5 in this chapter have been written to provide the reader with the necessary background on music. Pitch structures like scales and chords are embedded in the concept of tonality which is related to consonance. The notion of consonance is in turn related to human perception of musical tones and gives rise to tuning and temperament systems.

1.2 Perception of musical tones

“Pitch is a basic dimension of a musical tone that is defined to be that attribute of auditory sensation in terms of which sounds may be ordered from low to high” (definition of American National Standards Institute). Therefore, pitch is entirely subjective and cannot be obtained in an analytical way; perceptual experiments are required. In many cases however, the fundamental frequency of a (musical) sound represents the pitch. When referring to a tone of a certain frequency, what is meant is either a simple (sine) tone consisting of only that frequency, or a complex tone having that frequency as its fundamental frequency. Human pitch perception is logarithmic with respect to fundamental frequency. This means for example that the perceived distance between two tones of 220 Hz and 440 Hz is the same as the perceived distance between two tones of 440 Hz and 880 Hz.

In this thesis, we will treat tones and pitches in an abstract way, referring to them as frequency ratios or note names. However, before making this 'simplification', we will address the perception of some musical phenomena, since those contribute to the notion of consonance and dissonance, which is an important notion is this thesis. This section is heavily based on Rasch and Plomp (1999).

1.2.1 Beats

If two sine tones are equal in frequency and played simultaneously, they sound as one sine tone which can be louder or softer. If the two tones have the same phase (starting point) then the amplitudes of the two signals can be added and the resulting tone has an amplitude which is the sum of the two individual amplitudes. If the phases of the tones are opposite (the peaks of the one align with the troughs of the other) the resulting tone will be softer and the amplitude is the difference of the amplitudes of the two individual tones. If the amplitudes of the two tones (being in opposite phase) are the same, they can cancel each other.

What happens if two sine tones that are not equal in frequency are played simultaneously? If the frequencies differ slightly, and the two signals start with the same phase, the result is a signal with an amplitude that slowly oscillates from large (when in phase) to small (when out of phase), see figure 1.1. Even

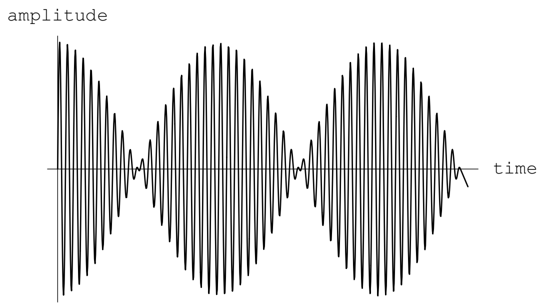


Figure 1.1: Sum of two sine waves with frequencies in the ratio 1 : 1.05.

though there are really two sine tones, it sounds like there is only one tone with a slow amplitude variation. This amplitude variation is called *beating*, and the beat frequency is $f_1 - f_2$, the difference in frequencies of the two original tones, as we will see. The beat frequency can be understood by superposition of two sine waves with slightly different frequency. If the two sine waves start in phase, they go more and more out of phase until they are at opposite phase and then turning back into phase again, and so on. The overall shape represents the beat frequency which is a variation in the loudness of the signal (fig. 1.1). A stimulus equal to the sum of two simple (sine) tones with frequencies f and g, and equal amplitudes, is represented by

$$p(t) = \sin[2\pi f t] + \sin[2\pi g t] \tag{1.1}$$

and can be written differently as:

$$p(t) = 2\cos[2\pi \frac{1}{2}(g - f)t] \cdot \sin[2\pi \frac{1}{2}(f + g)t]. \tag{1.2}$$

If $g - f$ is small, this sound is perceived as a signal with a frequency that is the average of the original primary frequencies, and an amplitude that fluctuates slowly with a beat frequency of $g - f$ Hz.

1.2.2 Critical bandwidth and just noticeable difference

If two tones are close enough in frequency such that their responses on the basilar membrane in the ear overlap, these tones are defined to be within the same critical band. The perceptual implication of the (frequency dependent) critical band is related to the finding that the ear can only make sense of one signal per critical band (Plomp 1964; Plomp and Mimpen 1968). Therefore, critical bandwidth is said to be the maximal frequency difference between two simultaneously presented notes that are not resolved by the ear (that means, not processed separately but combined). If for example two tones of 100 Hz and 110 Hz are played simultaneously, then only one beating or unresolved sound is heard (not two distinct tones) because the critical band at 100 Hz is larger than 10 Hz. The width of the critical band is roughly constant below frequencies of 500 Hz, and increases approximately proportionally with frequency at higher frequencies (see figure 1.2).

Whereas the critical bandwidth represents the ear's resolving power for simultaneous tones or partials, the Just Noticeable Difference (JND) deals with the distinction of two consecutive tones. The JND between two notes is the smallest change in frequency that a listener can detect. The JND depends on frequency and is furthermore highly dependent of the method with which it is detected. Furthermore it varies with duration, intensity of tones and training of the listener. The JND is roughly a constant percentage of the critical band for varying frequency (see figure 1.2). The JND is an important measure for a number of matters discussed in this thesis. For example in the discussion about a suitable number for the equal division of the octave (chapter 3), a limit could be set at

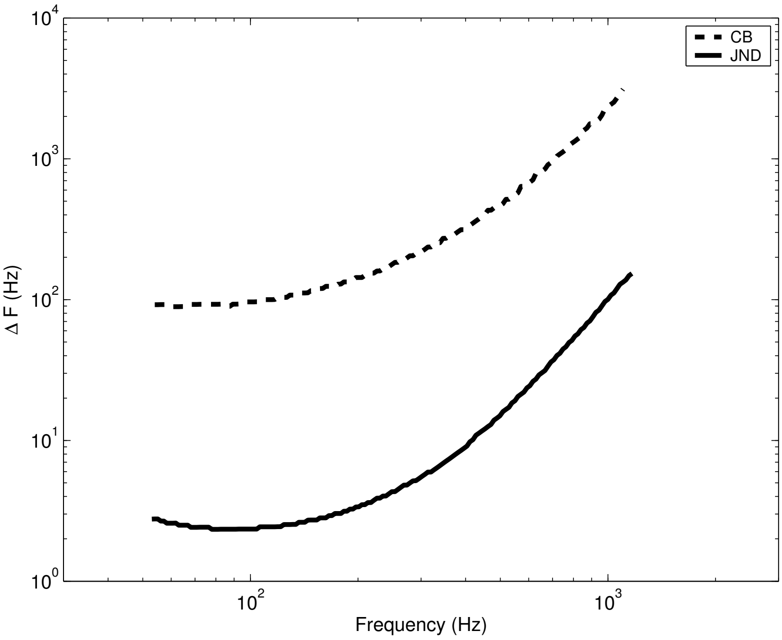


Figure 1.2: Global fit of critical band (CB) and just noticeable difference (JND) plotted as a function of their center frequency, shown over a part of the audible frequency range for humans (which is from about 20Hz to about 20kHz).

the division whereby the smallest parts are equal to the JND. However, it would not be helpful to express small frequency distances in units of JND since, as mentioned above, this measure depends on many variables.

1.2.3 Virtual pitch

The perceived pitch of a complex tone consisting of a fundamental and a number of partials, is usually the same as that of a sine wave with a frequency equal to the fundamental of the sound. However, it has turned out that, when this fundamental is removed from the sound, this doesn't change the sound, the fundamental frequency can still be observed. This is called the missing fundamental or *virtual pitch* (Terhardt 1974; Terhardt, Stoll, and Seewann 1982). One can wonder if a virtual pitch can always be observed or how many partials are needed to create a virtual pitch. Experiments have pointed to a dominance region which goes from roughly 500 Hz to 2000 Hz (Plomp 1967; Ritsma 1967). The partials that are falling in this region have a bigger influence on the pitch than other partials. Smoorenburg (1970) showed that it is possible to create a virtual pitch with only two partials in the dominance region, and Houtgast (1976) showed that virtual pitch could be obtained even with one partial with noise (see also Rasch and Plomp 1999 for the conditions under which this virtual pitch could be perceived.). The phenomenon of the missing fundamental was already observed by Seebeck (1841) and was brought under the attention of modern psychoacoustians by Schouten (1938). The observations led Schouten to the formulation of the theory of periodicity pitch, according to which the pitch is derived from the waveform periodicity. The periodicity is not changed if the fundamental is removed. However, as early as the 1950s new observations had indicated that the time-domain model in its original design was not fully adequate, since non-periodic sounds can produce virtual pitch as well. Tones with inharmonic partials can produce a virtual pitch which will be the fundamental of the harmonic series which is the closest to the inharmonic partials in the sound (Rasch and Plomp 1999). When there is ambiguity about which harmonic series the partials of a sound belong to, more than one virtual pitch is possible and can be perceived (but not at the same time) depending on the context (see Schulte, Knief, Seither-Preisler, and Pantev 2001). This ambiguous perception of pitch has often been compared to visual illusory contours (see for example Sethares 1999).

1.2.4 Combination tones

Two simple tones at a relatively high sound pressure level and with a frequency difference that is not too large can give rise to the perception of so-called combination tones (see for example Jeans 1968). These combination tones arise in the ear as a product of nonlinear transmission characteristics. The combination tones are not present in the acoustic signal, however, they are perceived as if they were present. The ear cannot distinguish between perceived components that are “real” (in the stimulus) and those that are not (combination tones). The combination tones are simple tones that may be canceled effectively by adding a real simple tone with the same frequency and amplitude but opposite phase. This cancellation tone can be used to investigate combination tones.

The possible frequencies of combination tones can be derived from a general transmission function. Assume a stimulus with two simple tones (as in 1.1):

$$p(t) = \cos 2\pi ft + \cos 2\pi gt, \tag{1.3}$$

f and g being the two frequencies. Linear transmission is described by $d = c_0 + c_1 p$ (c_0 and c_1 being constants). If transmission is non-linear, higher order components are introduced: $d = c_0 + c_1 p + c_2 p^2 + c_3 p^3 + \ldots$ The quadratic term can be developed as follows:

$$\begin{aligned} p^2 &= (\cos 2\pi ft + \cos 2\pi gt)^2 && (1.4) \\ &= 1 + \frac{1}{2}\cos 2\pi 2ft + \frac{1}{2}\cos 2\pi 2gt + \cos 2\pi(f+g)t + \cos 2\pi(f-g)t && (1.5) \end{aligned}$$

It can be seen that components with frequencies $2f$, $2g$, $f+g$, and $f-g$ are introduced in this way. The components $f+g$ and $f-g$ are the first order combination tones. Similarly, the cubic term can be developed and results in components with frequencies $3f, 3g, 2f + g, 2g + f, 2f - g, 2g - f$, where the latter four are the second order combination tones. The higher terms of the nonlinear transmission formula can be worked out analogously. The components of the form $nf + mg$ are called sum-tones, and the components of the form $nf - mg$ are the difference tones[1]. The first order combination tones are more audible than the higher order combination tones, and the difference tones are generally more audible than the sum tones (for information on the audibility region of combination tones, see Smoorenburg 1972). If both main tones belong to a common harmonic series (see section 1.3), then the combination tones can easily be calculated. In the harmonic series on C, the C and E are the fourth and fifth harmonics respectively. The first order difference tone of a simultaneous sounding C and E is the first $(5 - 4 = 1)$ harmonic C, the summation tone is the 9^{th} $(5 + 4 = 9)$ harmonic, D.

1.3 Just intonation and the compromises of temperaments

When two tones have a pitch relation such that the ratio of their (fundamental) frequencies is a rational number, the interval between the tones is called a just interval. This means that the two notes are members of the same harmonic series. Just intonation is any musical tuning in which the frequencies of notes are related by rational numbers (Lindley 2005). Just intonation is generally referred to as the tuning system that is used by violinists and other musicians using non-fixed note instruments. However, for instruments like a piano, it is not possible to play in just intonation as we will see, and a temperament system has to be developed.

Since the just intonation concept is as old as Pythagoras (as we will see), this section on just intonation and temperament systems is preceding the section on the concept of consonance and dissonance (sec. 1.4). In the present section the just intonation system is treated to be the most preferred tuning system; the next section will then discuss possible explanations of the fact that intervals which are related by whole number ratios, are said to be consonant (and thus preferred).

1.3.1 Harmonic series

Harmonics are generated by all natural vibrating systems. A vibrating string produces under normal conditions not only the fundamental tone but also the other harmonics, or overtones. The harmonic series with fundamental C is shown

[1] Since a difference tone exist at frequency $f - g$, which is the same as the beat frequency between two tones with slightly different frequencies, one could think that these two phenomena are related or arise from a common origin. However, Hall (1981) explained that these two phenomena are totally different.

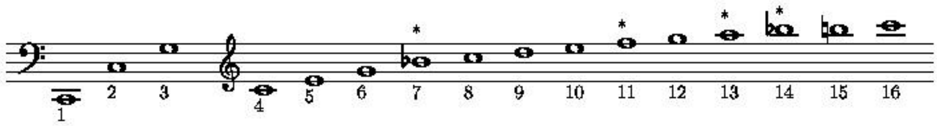


Figure 1.3: Harmonic series with fundamental C.

in figure 1.3. Each harmonic is indicated with a number, starting with the fundamental C as number one, the (first harmonic), the c (octave higher) number two (second harmonic) and so on as can be seen from figure 1.3. The number indicating each harmonic is also the denominator of the fraction representing the length of the string segment producing the tone. For example, if a string produces the fundamental tone, then the string which is half as long (length divided by two) produces the tone which is an octave higher, the second harmonic. If the initial string length is divided by three then the tone sounds an octave and a fifth higher (the third harmonic), as can be seen from the harmonic series. By definition, the harmonic series is that sequence of frequencies that represents all whole-number multiples of any particular fundamental frequency. If the fundamental C has frequency f, then the second harmonic (c) has frequency $2f$, the third harmonic has frequency $3f$, and so on. Therefore, it is possible to extract the frequency ratios for some important musical intervals from the harmonic series. For example, the interval of a perfect fifth can be found between the second and the third harmonic. Therefore, the frequency of the third harmonic equals $\frac{3}{2}$ times the frequency of the second harmonic: the interval of a perfect fifth is characterized by the ratio $\frac{3}{2}$. The frequency ratios of some other basic intervals are listed in table 1.1.

interval	**ratio**
octave	2/1
major sixth	5/3
minor sixth	8/5
fifth	3/2
fourth	4/3
major third	5/4
minor third	6/5

Table 1.1: Ratios of several intervals derived from the harmonic scale.

Tuning to whole number ratios is referred to as just intonation. The just intonation major diatonic scale is defined by tuning the tonic-, subdominant- and dominant-triad as $4 : 5 : 6 = 1 : 5/4 : 3/2$. This means that in a major diatonic scale the tones do, mi, sol have frequency ratios $4 : 5 : 6$ in relation to each other, as well as the tones fa, la, do and sol, ti, re (the last one an octave higher). Table 1.2 shows the ratios of the notes in the just major scale compared to the fundamental. The tones in the just intonation scale are defined

Note	do	re	mi	fa	sol	la	ti	do
Ratio	1:1	9:8	5:4	4:3	3:2	5:3	15:8	2:1

Table 1.2: Frequency ratios between the different notes of the major scale and the fundamental 'do'.

in relation to each other, if one plays in A major with a 'do' of 440Hz, the 're' equals $9/8 \cdot 440 = 495$Hz and so on.

In just intonation, modulations are problematic. Say, for example you play in C major and want to change key to G major. In C, the tones of the scale: $c, d, e, f, g, a, b, (c)$ are tuned as in table 1.2, so for example the a is tuned as 5/3 times the frequency ratio of c. If the key is then switched to G major, again the tones of the scale, now: $g, a, b, c, d, e, f\sharp, (g)$ are tuned to the ratios as in table 1.2. The a is now tuned as 9/8 times the frequency of g. However, the a in the scale of C was tuned as 5/3 times the frequency ratio of c, and the g in the scale of C was tuned as 3/2 times the frequency ratio of c. This means that the a (in the scale of C) is tuned as $(5/3)/(3/2) = 10/9$ times the frequency of g, which is different from the tuning of the a in the scale of G (9/8 times the frequency of g). As a result one would need two tunings for the a (and also for other notes), which is highly impractical, and even impossible when playing on an instrument with a fixed note system, like a piano.

Even when playing in one and the same key, just intonation may not be preferable. Consider the note sequence given in figure 1.4. Starting with the first

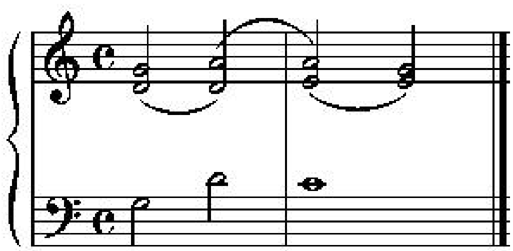


Figure 1.4: Pitch drift illustrated by $\frac{3}{4} \times \frac{3}{2} \times \frac{3}{5} \times \frac{3}{2} = \frac{81}{80}$. The pitch of the final G will be tuned as 81/80 times the frequency of the first G.

G, in just intonation the D will be tuned a perfect fourth 4/3 below the G. In the adjacent chord, the A will then be tuned as a perfect fifth 3/2 above the D. In the chord thereafter, the C is to be tuned as a major sixth 5/3 below the A. The final G is then tuned as a perfect fifth 3/2 above this C. Comparing the tuning of the first and the last G, we can calculate that the final G is tuned as $\frac{3}{4} \times \frac{3}{2} \times \frac{3}{5} \times \frac{3}{2} = \frac{81}{80}$ of the frequency of the first G. This is called a pitch drift and is a familiar problem related to just intonation. To solve the above mentioned problems, a temperament system can be developed, in which a compromise of the tunings of notes exists.

1.3.2 Temperament difficulties

Temperament deals with the division of the tones within the octave. The problem of constructing a good temperament is generally viewed as how to unite perfect fifths with major thirds. If we add four perfect fifths, like in figure 1.5, the highest note is a little bit higher than the perfect third from the harmonic series. The highest note in the sequence of fifths would be a high E whose frequency relative to the low C would be $(\frac{3}{2})^4$ which is bigger than 5, the harmonic number of the same E in the harmonic series. This difference of $\frac{81}{80}$ is called the syntonic comma.

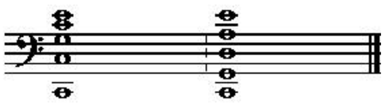


Figure 1.5: Chord from harmonic series (left) and chord of pure fifths (right) representing the syntonic comma between the two highest notes of both chords.

Another comma which plays a role is the Pythagorean comma. As we calculated, the ratio of a perfect fifth is $\frac{3}{2}$. If we start with the lowest C on the piano, and then go up in perfect fifths twelve times, we end up at a $B\sharp$ (figure 1.6) whose frequency relative to the low C would be $(\frac{3}{2})^{12}$. We will see that this note is very

Figure 1.6: Sequence of twelve fifths.

close to the highest C on the piano, seven octaves above the lowest. According to the ratio of a perfect octave the frequency of this note would be 2^7 relative to the low C. A little calculation

$$\left(\frac{3}{2}\right)^{12} = 129.746 \neq 2^7 = 128$$

shows us that the $B\sharp$ will be higher than the C. The difference in pitch is called the comma of Pythagoras and measures 531441/524288. Although the notes $B\sharp$ and C have different names, it can be understood that they are in pitch too close together to make separate keys on a piano (or another fixed tone instrument) for them[2]. Furthermore, *if* we would distinguish these notes from each other, this could also be done for $B\sharp\sharp$ and $D\flat$, and so on adding more sharps and flats, such that a piano would need an infinite number of keys. Notes such as $B\sharp$ and C are called enharmonically equivalent to each other. The infinite line of fifths $\ldots B\flat, F, C, G, D, A, E, B, F\sharp \ldots$ is considered to be a circle consisting of 12 notes by identifying the enharmonically equivalent notes $G\sharp$ and $A\flat$.

This problem concerning the commas was solved by the introduction of equal temperament in the 16^{th} century. In this 'theory' the octave is divided into twelve equal semitone intervals. The ratio between two semitones (consecutive chromatic notes) is then:

$$2^{\frac{1}{12}} : 1$$

The major second then has the ratio $2^{\frac{2}{12}} : 1$, the minor third the ratio $2^{\frac{3}{12}} : 1$ and so on. The octave has ratio $2^{\frac{12}{12}} : 1 = 2 : 1$ and is 'in tune' according to just intonation. The rest of the tones are slightly out of tune, see table 1.3. We note that in equal temperament the fifth is slightly smaller than $\frac{3}{2}$ which

	harmonic series	equal temperament
octave	2	$2^{\frac{12}{12}} = 2$
fifth	$\frac{3}{2} = 1.500$	$2^{\frac{7}{12}} = 1.498$
fourth	$\frac{4}{3} = 1.333$	$2^{\frac{5}{12}} = 1.335$
major third	$\frac{5}{4} = 1.250$	$2^{\frac{4}{12}} = 1.260$
minor third	$\frac{6}{5} = 1.200$	$2^{\frac{3}{12}} = 1.189$

Table 1.3: Some ratios from just intonation compared with those from equal temperament.

makes the Pythagorean comma disappear. A convenient unit to express intervals in, is the *cent* which is based on equal tempered tuning. One cent is defined as one hundredth part of an equal tempered semitone. That means that an interval expressed in frequency ratios $\frac{x}{y}$ has a width of $1200 \cdot \log_2(\frac{x}{y})$ cents. An equal tempered minor second measures 100 cents, a major second 200 cents, and so on, until the octave which measures 1200 cents. In equal temperament, enharmonically equivalent notes are tuned exactly the same, and the circle of fifths is automatically closed since the $G\sharp$ and the $A\flat$ are equivalent. In any

[2]However, both the syntonic comma and the Pythagorean comma lie, for a large range of the audible frequencies, above the just noticeable difference. This means that listeners can usually detect this difference in frequency.

unequal temperament there is a problem closing the circle of fifths, since the interval between the $G\sharp$ and $E\flat$ (the fifth above the $A\flat$) which is known as the wolf fifth, is considerably out of tune compared to the fifth from just intonation.

1.3.3 Tuning and temperament systems

Various tunings and temperaments exist as solutions for the comma-problem. A tuning is a system, all of whose intervals can be expressed as rational numbers, which leaves the commas to fall as they must. Examples of tuning systems are the just intonation system from table 1.2, and the Pythagorean tuning as we will explain below. A temperament involves deliberately mis-tuning some intervals to obtain a distribution of the commas which is more useful in a given context. A regular system is a tuning or a temperament in which all the fifths but one are of the same size (Barbour 1951).

It is not exactly known when temperament was first used. Vicentino (music theorist and composer, 16th century) stated that fretted instruments have always been in equal temperament; since one fret applies to more than one string this was the only useful temperament (Barbour 1951). For keyboard instruments, Zarlino (Italian music theorist and composer, 16th century) declared that temperament was as old as the chromatic keyboard (Barbour 1951).

Until about 1500 the tuning according to Pythagoras was used. Pythagoras (sixth century B.C.) found (using a monochord) the ratios for the octave 2:1 and fifth 3:2. With these ratios he calculated the rest of the intervals. For example the major second is calculated as $3/2 \cdot 3/2 \cdot 1/2 = 9/8$, the major third as $9/8 \cdot 9/8 = 81/64$. Pythagoras also calculated that the diatonic semitone is not equal to the chromatic semitone, the difference between those tones is the Pythagorean comma as we have seen above. In this tuning, a $B\sharp$ is therefore higher than a C, a $C\sharp$ higher than a $D\flat$ and so on. According to this Pythagorean tuning, the octave, fifth and fourth are pure, the thirds and sixths are dissonant. This tuning was suitable for monophonic music and for the medieval polyphonic music in which only the octave and fifth were considered to be consonant.

When harmony had evolved to the stage that a more pure third was required, a new temperament was devised with all thirds pure and the fourths and the fifths as nearly pure as possible. Quarter-comma Mean-tone temperament was constructed by tuning $C - E$ pure and then tuning each fifth within that third ($C - G$, $G - D$, $D - A$, $A - E$) a quarter (syntonic) comma flat. This (regular) temperament is called Mean-tone because the pure major third, which measures 386 cents, consists of exactly two whole tones of 193 cents. Since the temperament had many pure thirds, as a consequence the wolf fifth was very large and hence unusable. There were also some wolf thirds which were very sharp, but these were kept in keys which composers took care to avoid. There were several different mean-tone temperament schemes, each involved slight adjustments to the sizes of the major thirds and fifths. For example, in sixth-comma mean-tone, the thirds were slightly worse, the fifths and fourths equally slightly better, and the wolves were smaller. With these Mean-tone systems, about 16 of the 24 major and minor keys were usable.

From the 17th century irregular temperaments such as those devised by Vallotti, Werckmeister, Kirnberger and others were used (for an overview, see Barbour 1951). These temperaments were based on the keys that are most used. All octaves were pure, keys related to C had nearly pure major thirds and fifths and keys distant from C had much less pure intervals. In fact, every key had its own character, which could be used by the composer. It has been demonstrated that Bach's Well-tempered Clavier with 48 prelude and fugues, two in each of the 12 major and 12 minor keys, used such a temperament, with some movements showing the purity of the better keys and others, with rapid note passages, disguising the impurity of the less good keys (Montagu 2002).

One way to avoid the wolf notes and to keep as many pure thirds and fifths as possible, was to increase the number of notes in the octave. Then, on a keyboard with separate notes for $D\sharp$ and $E\flat$, and for $G\sharp$ and $A\flat$, (and sometimes for some other notes as well) it was possible to play in almost every key. A problem with this was that the more keys there were on a keyboard, the more difficult is was for the keyboard player to remember which key should be used in which chord.

Since the middle of the 19^{th} century, the most popular form of tuning has been Equal Temperament. As discussed before, this is a scale with pure octaves, equally divided into twelve parts.

1.4 Consonance and dissonance

Consonance is an important concept since it may form the basis for tunings, temperaments and pitch structures like scales and chords. Sensory consonance refers to the immediate perceptual impression of a sound as being pleasant or unpleasant. It may be judged for sounds in isolation (without a musical context) and by people without musical training (Palisca and Moore 2006). In section 1.4.1 we will go into the history of explanations of sensory consonance. In section 1.4.2 we will see that besides sensory consonance, more types of consonance can be distinguished.

1.4.1 Explanations on sensory consonance and dissonance

Playing in just intonation is related to sensory consonance as the frequencies are related by small whole number ratios. Explanations of sensory consonance are concerned with the fact that common musical intervals correspond - at least in Western, Indian, Chinese and Arab-Persian music (see Burns 1999) - to relatively simple ratios of frequencies, although some of these explanations do not require exact integer tunings, only approximations. Pythagoras already found that if two similar strings under the same strength sound together, they give a pleasant sound if the lengths of the strings are in the ratio of two small integers.

Small integer ratios

Galilei (1638) explained the preference for small integer ratios from the regularity of the resulting signal which is pleasant for the ear. Also Euler (1739) had a conscious feeling for ordered as opposed to disordered relations of tone, and he even proposed a measure of consonance based on this theory. When complex tones are considered, two tones whose frequencies are in small integer ratios have partly overlapping harmonics which makes the sound 'rich'. However, also simple (sine) tones are said to sound more consonant when tuned to simple integer ratios according to this theory. Modern expositions of this idea exist as well (Boomsliter and Creel 1961; Partch 1974), in which consonance is viewed in terms of the period of the combined sound. If two frequencies form an interval of a small integer ratio, the period of the combined sound is shorter which should be more pleasant for the ear. This is in fact a testable hypothesis and these so-called periodicity theories of consonance assume some time-based detector in the ear. Neurological evidence for such temporal models exist. Cariani (2004) provides evidence to ground pitch-based theories of tonal consonance in inter-spike interval representations. He finds that "for both pure and complex tones, maximal salience is highest for unison and the octave separations and lowest for separations near one semitone". Tramo, Cariani, Delgutte, and Braida (2001) claim that 1) pitch relationships among tones in the vertical direction influence consonance perception and 2) consonance cannot be explained solely by the absence of roughness (see further below). They provide neurophysiological, neurological and psychoacoustic evidence to support these claims.

Tonal fusion

Stumpf (1898) proposed the idea that consonance was based on tonal fusion. The fusion of two simultaneously presented tones is proportional to the degree to which the tones are heard as a single perceptual unit. The idea behind this theory is, that when enough harmonics of two tones coincide, the tones perceptually fuse together.

Virtual pitch

The existence of virtual pitch has also been regarded to explain consonance. Terhardt (1974) and Terhardt, Stoll, and Seewann (1982) emphasize the role of learning in the perception of intervals. Different learning experiences lead to different intervals and scales, and hence, to different notions of consonance and dissonance. Virtual pitch tries to locate the nearest harmonic template when confronted with a collection of partials (or overtones). This is ambiguous if the sound is not harmonic (see section 1.2.3). According to Terhardt's view, dissonance is a negative valanced sensory experience that arises when a sound evokes highly ambiguous pitch perceptions. This explanation of consonance using virtual pitch is related to the periodicity theory since one of the most important explanations of virtual pitch is the periodicity of the sound. Cariani (2004) found neurological evidence supporting the virtual pitch theory. He concludes that "the present simulation demonstrates that interval-based models of low, virtual pitch can plausibly account for the consonance of pairs of pure and complex tones that are presented in isolation".

Combination tones

Another possible explanation of consonance was the existence of combination tones. Among others, Krueger (1904) proposed that dissonance is proportional to the number of distinct difference tones. Consonance then occurs if there are only a few difference tones. Husmann (1953) looked at the fitting of the combination tones of two tones with their overtones. He calculated the percentage of over-tones that coincided with the combination tones. This gave an order of consonance. Plomp (1965) performed experiments regarding these combination tones, and argued that the nonlinear distortion of the hearing organ is so small that it cannot be regarded as a constitutive basis for consonance.

Roughness

Helmholtz (1863) was the first to propose a consonance theory based on the phenomenon of beats. Recall from section 1.2.1 that two sine tones produce beating when their frequencies are close together. Slow beating is generally perceived as being pleasant, fast beating as being rough and unpleasant with maximum roughness occurring at a beat rate of (around) 32 times per second. Since sound can be decomposed into sine wave partials, Helmholtz argued that dissonance is due the rapid beating of the partials of a sound. Consonance is then the absence of such beats. Helmholtz's theory resulted in a measure of consonance in which he made the assumptions that 32 Hz gives maximal roughness, and that roughnesses can be added. Plomp and Levelt (1965) did experiments on the perception of consonance with musically naive subjects and found that the dissonance of an interval is primarily due to rapid beats between the compound tones, which supports Helmholtz's theory. They found that the minimal and maximal roughnesses of an interval are not independent of the mean frequency of the interval. They are related to critical bandwidth with the rule of thumb that maximal tonal dissonance is produced by intervals subtending 25% of the critical bandwidth, and that maximal tonal consonance is reached for intervals greater than about 100% of the critical bandwidth (see fig. 1.7). This is a modification of Helmholtz's 32Hz criterion for maximum roughness because the critical bandwidth is not equally

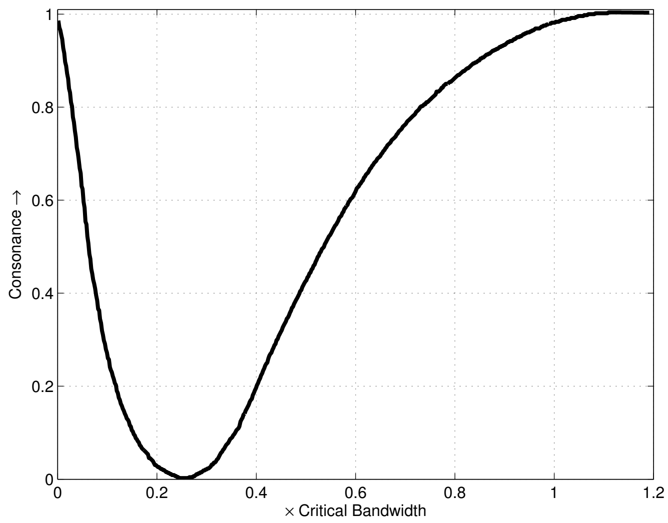


Figure 1.7: Standard curve from Plomp and Levelt (1965) based on their experiments, representing the consonance of two simple (sine) tones as a function of the critical bandwidth. The consonance scale is arbitrary.

wide at all frequencies. Due to the dependency of the critical bandwidth on frequency, intervals (like minor thirds) that are consonant at high frequencies, can be dissonant at low frequencies.[3] Note that this is different from the small frequency ratio hypothesis, not all intervals that are in small integer ratios produce consonant sounds. Similarly to Helmholtz, Plomp and Levelt (1965) claim that the dissonance can be calculated by adding up all of the dissonances between all pairs of partials. Kameoka and Kuriyagawa (1969a, 1969b) make a reproduction and an extension of Plomp and Levelt's tonal consonance ideas. For dyads, consisting of two partials, they find that the consonance gradually decreases as the frequency separation increases and is least when the frequencies are separated by approximately 10% in the middle frequency range. They establish a theory for calculating the subjective magnitudes of the dissonance of complex tones. A theoretical investigation clearly showed that the consonance of chords is greatly dependent on the harmonic structure. For example, a complex tone that includes only odd harmonics shows no consonant peak for the fifth (2:3), but does have consonant peaks for 3:5 and 5:7. Following Helmholtz (1863), Plomp and Levelt (1965) and Kameoka and Kuriyagawa (1969a), who state that the consonance of a sound depends on the absence of roughness of its partials, Sethares (1993) argues that the partials or harmonics of a sound define its timbre, and therefore consonance is dependent on timbre. Hence, the origin of consonance would be the instruments of a specific culture. String and wind instruments naturally produce

[3]This is consistent with musical practice where small intervals appear in the treble parts and the larger intervals like octaves and fifths appear in the bass part.

a sound that consists of exact multiples of a fundamental frequency. They are therefore appropriate for playing music in the 12-tone just intonation scale (or equal temperament, since that is a close approximation). However, in Indonesian Gamelan music for example, the instruments are all percussive, and do not produce exact integer multiples of a frequency. Therefore, the Western scale is not appropriate for that type of instrument, and indeed not used (Sethares 1999).

Culture

Finally, a cultural explanation of consonance exists. Cazden (1980) argued that the wide variety of scales and tunings used throughout the world serves as evidence that cultural context plays a key role in notions of consonance and dissonance. Cazden furthermore argued that an individual judgment of consonance can be modified by training, and so cannot be due entirely to natural causes.

As may be clear from the above, the notions of consonance and dissonance have changed significantly over the years and several theories exist next to each other. How much the perception of consonance and dissonance is due to basic sensory and perceptual factors and how much to learned ones remains unresolved (Palisca and Moore 2006).

1.4.2 Different types of consonance

We have addressed the explanations of sensory consonance, that apply to sounds in isolation, without a musical context. One may wonder if the notion of consonance changes when a sound is presented in a musical context. Some authors have described distinct types of consonance and dissonance.

Tenney (1988) identifies five different forms of what he calls the 'consonance/dissonance concept' (CDC). The five forms are summarized below.

- CDC-1: Melodic consonance. This type of consonance/dissonance refers exclusively to the relatedness of pitches sounded successively. This concept is related to and derived from the pre-polyphonic era, when music was primarily conceived melodically.

- CDC-2: Polyphonic consonance. During the early polyphonic period (900 − 1300) the consonance concept became a function of the interval between two simultaneously sounding tones. Stumpf's (1898) idea of tonal fusion was proposed to explain the rank ordering of intervals during this period.

- CDC-3: Contrapuntal consonance. In this type, consonance was defined by its role in counterpoint. These are the rules that describe voice-leading techniques. Thus, the context of the notes was important here, as opposed to the physical properties of the sound.

- CDC-4: Functional consonance. Influenced by Rameau (1722), notes and intervals were judged consonant or dissonant according to whether they have a simple relationship to the fundamental root or tonic. Dissonant notes have the implication of 'motion', since they set up the expectation to return to the root.

- CDC-5: Psychoacoustic consonance. This is the most recently developed concept of consonance and focuses on the perceptual mechanisms of the auditory system. One view on consonance within this class is called sensory consonance and is usually credited to Helmholtz (1863) and Plomp and Levelt (1965). Another component of psychoacoustic consonance is 'tonalness', which is based on Rameau's fundamental bass and Terhardt's (1984) notions of harmony, which was extended by Parncutt (1989). A major component of tonalness is the closeness of the partials to the harmonic series.

Terhardt (1977) developed a two-component model of musical consonance. He argues that the concept of consonance obviously implies the aspect of pleasantness, but that pleasantness is not confined to musical sounds. Therefore, he has termed this aspect of consonance 'sensory consonance'. He argues that, as sensory consonance was not conceptualized to explain the essential features of musical sounds, there must be another component to account for this. This other component was termed 'harmony'. Thus, musical consonance consists of sensory consonance and harmony. Sensory consonance consists in turn of roughness, sharpness and tonalness, and harmony consists of 'affinity of tones' and 'root-relationship' (virtual pitch). Roughness (as defined by Helmholtz) is the major component of the sensory consonance, sharpness is a kind of spectral envelope weighted loudness, and tonalness is the opposite of noisiness. Tone affinity means that tones may be perceived as similar in certain aspects. That is, that in some respect, a tone may be replaced by another one. The concept of root-relationship indicates that the root of a musical chord is not merely a theoretical concept, but that it is an attribute of auditory sensation, i.e., a virtual pitch. Terhardt's (1977) model is closely related to that of Helmholtz (1863). A difference between the two is that Helmholtz did not take the virtual pitch into account.

Many different concepts of consonance have arisen, and a clear distinction between the consonance of simultaneously presented tones and the consonance of successively presented tones seems to exist. While simple ratios may be preferable for simultaneously presented tones, it is not clear whether this is the case for tones presented successively (Palisca and Moore 2006). A number of experiments investigating intonation in performances have been performed (see for example Kopiez 2003; Rakowski 1990; Loosen 1993) and show that there is no simple answer. Boomsliter and Creel (1963) found that, within small groups of melodic notes, simple ratios are preferred, although the 'reference' point may vary as the melody proceeds. However, others have concluded in contrast that there is no evidence that performers tend to play intervals corresponding to small integer ratios for either melodic of harmonic situations (Burns 1999). There is a tendency for small intervals to be tuned smaller and for large intervals to be tuned larger than equal temperament (Burns 1999).

Intonation

A concept closely related to the consonance of both simultaneous and successive notes is intonation. Intonation has been extensively researched with the goal of maximizing harmonic consonance. However, besides its role in harmonic consonance, intonation also has roles in, e.g., key coloration and harmonic meaning. An important melodic example concerns upward steps from pitches that function as leading notes. The higher the intonation of this leading note, the more a pull is felt towards the pitch above. Fyk (1995) makes a division of intonation in diatonic music[4] into four classes, based on a number of experiments with violinists.

1. Harmonic tuning. Many of the deviations from equal temperament take place in trying to tune in just intonation (small integer ratios) with the underlying harmony.

2. Melodic tuning. It is observed that the intonation is raised in the context of an ascending and it is lowered in the context of a descending melody.

3. Corrective tuning. When a performer perceives a small deviation in playing a melody, he/she corrects it by adjusting the note itself or the note that follows immediately. The former kind of adjustment occurs frequently in string quartet playing in order to create a given chord in just intonation.

4. Colouristic tuning. When the performer plays a rising series of two notes with an octave interval between them, the interval is in precise just tuning (2 : 1). However, when the same series of notes has an overlapping period in between them, it is observed that the performer raises the higher note as soon as the lower note disappears.

The concept of harmonic tuning as defined by Fyk (1995) is in agreement with the notion of sensory consonance. Although melodic consonance has not been clearly defined as discussed above, it is important to note that melodic tuning and melodic consonance do not necessarily have the same intention. For example, in melodic tuning it can be desirable to tune a leading note a little bit high so that the direction to the next note is clear. This does not necessarily result in the most consonant melodic interval between those two notes. However, Fyk's definition of melodic tuning seems to be consistent with the observations published in Burns

[4]See also Kanno (2003) who has written about intonation and performance practice in non-diatonic, microtonal and new music.

(1999) that there is a tendency for small intervals to be tuned smaller and for large intervals to be tuned larger than equal temperament.

1.5 Tonality

Tonality is a term that refers most often to the orientation of melodies and harmonies towards a referential (or tonic) pitch class. In the broadest sense, however, it refers to systematic arrangements of pitch phenomena and relations between them (Hyer 2006). A large number of definitions for tonality have been proposed among which there is no consensus about whether it applies to both Western and non-Western music. Another area of disagreement, similar to the discussion on consonance and relating to the origin of the term tonality, is whether, and to what extent, tonality is natural or inherent in music, and whether, and to what extent, it is constructed by the composer, performer and listener.

The vocabulary of tonal analysis consists of scales and chords (in case of harmony). Tonal music is described in terms of a scale of notes. Chords are built on the notes of that scale. In the context of a tonal organization, a chord or a note is said to be "consonant" when it implies (perceptual) stability. Note that this use of the term consonance is different from 'sensory consonance' and perhaps closest to what Terhardt (1977) defined as the 'harmony' component of musical consonance. A tonal piece of music will give the listener the feeling that a particular chord or note (the tonic) is the most stable and final. Establishing a tonality in Western tradition is accomplished through a *cadence*, some chords in succession which give a feeling of a completion or rest. When the sense of a tonic chord is changed, the music is said to have changed key, or *modulated.* In a tonal context, a dissonant chord is in tension with the tonic. *Resolution* is the process in which the dissonant chord or note moves to a consonant chord or note in the tonal organization. The majority of tonal music assumes octave equivalence, which means that the notes an integer number of octaves apart are perceived to have the same function. In primitive music, the tonic has three main determinants. They are 1) great frequency (number of occurrences) and length compared to the other tones, 2) final position in individual sections and phrases, and 3) terminal position in the song (Nettl 1956).

Historically, theories of tonal music have generally been dated from Rameau (1722), who has described music written through chord progressions, cadences and structure. In 1844, Fétis defined tonality as the set of relationships, simultaneous or successive, among the tones of a scale, allowing for other types of tonalities among different cultures. Moreover, Fétis believed that tonality was entirely cultural. In contrast, Riemann believed that tonality was entirely natural in the way that all types of tonality can be derived from a single principle based on the chordal functions of the tonic, dominant and subdominant (Dahlhaus and Gjerdingen 1990).

Reti (1958) differentiates between harmonic tonality and melodic tonality. In this context he argues that in harmonic tonality the chord progression $V - I$ is the only step "which as such produces the effect of tonality", and describes melodic tonality as a type wherein "the whole line is to be understood as a musical unit mainly through its relationship to this basic note [the tonic]".

1.5.1 Scales

A scale is a sequence of notes in ascending or descending order of pitch (Drabkin 2005), that provides material for a musical work. In the rest of this thesis, we do not use the property of orderedness[5], we will treat a scale as a set of tones.

Some scales, such as the chromatic scale of the piano, are sufficiently defined by a sequence of notes within a single octave, which can be extended without limit in either direction by octave transposition. Others, such as the medieval gamut, are complete in themselves. Scales may be described according to the intervals they contain - for example diatonic, chromatic, whole tone, or by the number of different pitch classes they contain - for example pentatonic, hexatonic, octatonic, etc. Scales are often abstracted from performance or composition, though they are often used pre-compositionally to guide or limit composition. Each note in a scale is referred to as a scale degree. The simplest system is to name each degree after its numerical position in the scale, for example: the first or I, the fourth or IV. In the Western major diatonic scale, the degrees are named in order: tonic, supertonic, mediant, subdominant, dominant, submediant and leading-tone.

Burns (1999) notes that "the evidence from ethnomusicological studies indicates that the use of discrete pitch relationships is essentially universal". One possible explanation of the human propensity to discretize pitch space involves the idea of categorical perception (Burns and Campbell 1994; Burns and Ward 1978; Locke and Kellar 1973). The brain tries to simplify the world around it.

In Western traditional music theory, scales generally consist of seven notes and repeat at the octave. Most familiar scales are the major and minor (harmonic and melodic) diatonic scales. Additional types include the chromatic scale, whole tone scale, pentatonic scale and octatonic scale. Many other musical traditions employ scales that include other intervals and/or a different number of pitches. Gamelan music uses a variety of scales including Pélog (5 note scale) and Sléndro (7 note scale), neither of which lies close to the familiar 12-tone equal temperament. In Indian music, ragas - consisting of 5, 6 or 7 notes - are the melodic framework embedded in the 22-tone shrutis scale (Burns 1999). In Maquams, the melodic modes from Arab music, quarter tone intervals may be used (Zonis 1973). Microtonal scales, formed by the division of the octave into intervals smaller than a semitone, have long been used in Eastern cultures (e.g. in Hindu ragas), and

[5]The fact that a scale is defined as an *ordered* set is not acknowledged by everybody, see for example: `http://launch.groups.yahoo.com/group/tuning/` for a discussion on this topic.

they have also been adopted by a number of Western composers (Latham 2002). For example, Partch (1974) proposed a 43-tone scale (in 11-limit just intonation).

Dowling (1982) differentiated between a scale and tonal material by defining a psychophysical scale as the general system by which pitches are related to the frequencies of tones. He defined tonal material as the entire set of pitch intervals available in a given musical culture. In Western music this would constitute the set of semitone intervals of the 12-tone equal-tempered scale (Dowling 1982). Considering the 12-tone chromatic scale to be a valid scale, one can understand that a scale can sometimes be similar to the tones available from a temperament.

As in the discussion on consonance and tonality, many conflicting opinions exist about the origin of (traditional) scales. It has been proposed that the origin of scales lies in speech inflections. Nations or regions adopt a particular type of inflection which may distinguish them from others. Similarly, in music, groups of people have adopted characteristic idioms and inflections, which in course of time took the form of favoring some pitches of pitch intervals and avoiding others (Latham 2002). However, this view is not restricted to a purely cultural origin of scales. Natural principles such as mathematical, psychological and neurophysiological views can be incorporated as well.

Besides the various views on the origin of scales, also many different ways to evaluate the goodness, reasonableness, fitness, well-formedness and quality of a scale exist, each criterion leading to a different set of best scales or tunings.

1.6 What lies ahead

Now that we have a basic understanding of the established theories of consonance, tuning, temperament and scales, we will see how to apply these notions and also investigate these from a different viewpoint in the coming chapters. Our investigations may lead to different insights and new explanations on existing temperaments and scales. In this thesis, we will discuss several criteria to derive a scale. These criteria can operate as models explaining the existence of some scales, or they may serve as an evaluation of certain existing scales. Furthermore, the resulting scales can be interpreted as suggestions for new scales that have not been explored until now.

In chapter 2 we will interpret the tonal material of Western music in an algebraic way, where we focus on just intonation. A geometrical description of the just intonation framework can be made, which leads also to a geometrical interpretation of the Western notation system as well as the familiar 12-tone equal temperament. We start by describing tones in terms of frequency ratios, (Western) note names, or (equal tempered) pitch numbers. These three concepts are meant abstractly and are mathematically connected through (homomorphic) projections. The geometrical description of the tone system will be used throughout this thesis. Since several geometrical models for tone systems have been proposed, a comparison will be made.

Chapter 3 goes into the details of equal temperament, where also microtonal systems with $n > 12$ will be addressed. There are several ways to evaluate the 'goodness' of an n-tone equal tempered system. We will address two of them, the first being the requirement that an equal temperament should approximate some ratios from just intonation as well as possible. The second condition has to do with the application of equal temperament to a notational system where Western harmony is incorporated. Together, these conditions result in n-tone equal tempered systems where n equals $12, 19, 31, 41$ or 53, which have been covered by previous literature as we will see. The temperaments can be represented in the tone space developed in chapter 2.

Chapter 4 addresses the question of the quality and origin of scales from a mathematical viewpoint. We will focus on the notion of mathematical beauty or geometrical well-formedness applied to pitch structures, that can possibly serve as the principled basis for tonal music. It will turn out that there is a highly persistent principle holding for pitch structures like scales (also non Western scales), diatonic chords and harmonic reductions: if presented in the tone space described in chapter 2 they will form compact and convex or star-convex shapes. Convexity and star-convexity may be explained in terms of consonance, such that in a convex musical pitch structure the consonance is optimized. For the star-convex scales it turns out that consonance is optimized according to the tonic of that scale.

In chapters 5 and 6, applications of compactness and convexity in the tone space are addressed. In chapter 5 compactness and convexity are used as measures of consonance with which the preferred intonation of chords in isolation is modeled. Euler's Gradus function is used as the measure of consonance, and it turns out that compactness represents this measure of consonance best. Chapter 6 discusses two computational applications, the first being a model for modulation finding using the notion of convexity and the second being a model for pitch spelling using the notion of compactness on the tone space. The first model turns out to be only moderately successful. However, the second is much more successful. Pitch spelling is the process of disambiguating equal tempered pitch numbers and transcribing into Western note names. The pitch spelling model based on compactness has been tested on the first book of the Well-tempered Clavier by Bach and has resulted in a percentage of 99.21% correctly spelled notes.

We will conclude by stating that the principles of convexity and compactness may reflect universal properties of tonal pitch structures.

Chapter 2

Algebraic interpretation of tone systems

Attempts to capture musical pitch in geometrical models have a long history. Representing pitches in a geometrical structure is useful for a number of reasons. The quality of a musical pitch is dependent on its relation to other pitches. This could be represented in a geometrical model. The spatial features of such a model are useful when describing relations of distance between pitches. Furthermore, geometrical models are usually possible to visualize and therefore more easy to understand.

In this chapter, an algebraic and eventual a geometrical interpretation of tone systems is described. The algebraic interpretation is important for understanding the mathematical properties of the tone system that we will use in the rest of this thesis. These mathematical properties have implications for the pitch structures that can be described in the tone space. In section 2.2 we will give the formal derivation of this tone space. To have a full understanding of the mathematics that is used in this derivation, section 2.1 introduces the necessary terms and definitions. Finally, in section 2.3 a comparison of the derived tone space with other geometrical systems of musical pitch is made. The readers that are not interested in the mathematical background of tone systems can skip most of this chapter, but are advised to read the second half of section 2.2.2 and section 2.3.

2.1 Group theory applied to music

In this section we will provide some basic definitions from group theory that we will use especially in the next section, and in a few instances in some other chapters in this thesis. After every definition an example is given that applies to music.

We start with the definition of a mathematical group. A *group* consists of a non-empty set G and a binary operation on G (usually written as composition with the symbol $\circ$) satisfying the following conditions[1].

[1] The definition includes the fact that, if $x, y \in G$, the product $x \circ y$ is also a member of the

- The binary operation is associative: $(x \circ y) \circ z = x \circ (y \circ z)$ for any $x, y, z \in G$.
- There is a unique element $e \in G$, called the identity element of G, such that $x \circ e = x$ and $e \circ x = x$ for any $x \in G$.
- For every $x \in G$ there is a unique element $x^{-1} \in G$, called the inverse of x, with the property that $x \circ x^{-1} = x^{-1} \circ x = e$.

For example, the integers $\mathbb{Z}$ under the operation of addition forms a group, since 1) addition is associative: $(a + b) + c = a + (b + c)$, 2) the identity element is 0: $0+a = a$ for any $a \in \mathbb{Z}$, and 3) the inverse for a is $-a$ since $a+(-a) = 0 = (-a)+a$ for all $a \in \mathbb{Z}$. This group could represent a musical tone system if we say that every integer represents a pitch, and consecutive integers represent pitches that are a semi-tone away from each other. In the MIDI pitch system the value of 60 represents middle C and each integer is a step on a standard piano keyboard (for example, 61 is $C\sharp$ above middle C). Actually, MIDI only defines the pitches in the integer range 0 to 127, but in theory, one could go up and down indefinitely. Another musical interpretation of the group $\mathbb{Z}$ could be formed by the line of fifths:

$$\ldots - D\flat - A\flat - E\flat - B\flat - F - C - G - D - A - E - B - F\sharp - C\sharp - G\sharp \ldots \tag{2.1}$$

which is an infinite line in both directions representing all possible note names. If these note names are identified with integers in the following way

$$\begin{array}{cccccccccccccccc} \ldots & D\flat & A\flat & E\flat & B\flat & F & C & G & D & A & E & B & F\sharp & C\sharp & G\sharp & \ldots \\ \ldots & -5 & -4 & -3 & -2 & -1 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \ldots \end{array} \tag{2.2}$$

the elements of the group $\mathbb{Z}$ can be used to indicate the note names.

A group G is *abelian* if all pairs of elements of G commute, $x \circ y = y \circ x$, in which case the order of the elements in a composition is irrelevant. Otherwise, the group G is *non-abelian*. The group $\mathbb{Z}$ under addition is abelian, since $a+b = b+a$.

A group G is said to be *finite* if it has a finite number of elements, and *infinite* otherwise. The order of a finite group G is defined as $|G|$, the number of elements of G. If x is an element of G, then for $n \in \mathbb{N}$, x^n is used to mean the composition $x \circ \ldots \circ x$ involving n terms. x^0 is defined to be e. The element x is of *finite order* if there is some $n \in \mathbb{N}$ such that $x^n = e$. If x is of finite order, then we define the order of x to be the least positive integer n such that $x^n = e$.

2.1.1 Cyclic groups

A group G is called a *cyclic* group if there exists an element $g \in G$, such that every element in G can be represented as a composition of g's. The group can

set G, since binary operations are required to satisfy closure. This is sometimes stated as a separate axiom (closure).

then be represented as

$$\langle g \rangle = \{x \in G | x = g^n \text{ for some } n \in Z\} \tag{2.3}$$

and is called the cyclic group generated by g. If g is of order n, then $\langle g \rangle = \{e, g, \ldots, g^{n-1}\}$. The set $\mathbb{Z}$ under addition is an infinite cyclic group since every element in $\mathbb{Z}$ can be represented as a composition of 1's ($1+1$, $1+1+1$, $1+(-1)$, etc.). In this case 1 is the generator. The sets $\mathbb{Z}_n = \{0, 1, 2, \ldots, n\}$ with addition modulo n are finite cyclic groups. For example, the group $\mathbb{Z}_3$ is the set $\{0, 1, 2\}$ under addition modulo 3. The cyclic group can also be denoted by C_n.

When the octave is divided into 12 equal semitones, as in equal temperament, we can count from C, 12 semitones up to finish at C again. Octave equivalence means that we identify notes with each other which differ by a whole number of octaves. If we associate a number with each pitch, starting from $C = 0$ then we can count up to 11, and then start back from 0. In terms of group theory, we can say that the set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ under 'addition modulo 12' is a group. The identity element is 0 and the inverse of n is $12 - n$. This group is written as $\mathbb{Z}_{12}$ or C_{12} and is a cyclic group. The generators for $\mathbb{Z}_n$ are precisely the numbers i in the range $0 < i < n$ with the property that n and i have no common factor, they are said to be relatively prime. In the case for $n = 12$, the possibilities for i are $1, 5, 7, 11$. In terms of musical intervals, this means that all notes can be obtained from a given note by repeatedly going up by a semitone (1), fourth (5), fifth (7) or major seventh (11). In fact, the sequence of semitones is generated by 1 (forwards) and 11 (backwards). The elements 5 and 7 generate the group in consecutive fifths or fourths representing the circle of fifths.

2.1.2 Properties of groups and mappings

Let A and B be groups, $\circ$ the binary operation in A, and $*$ the binary operation in B. The map $f : A \mapsto B$ is a *homomorphism* if, for all $a_1, a_2 \in A$:

$$f(a_1 \circ a_2) = f(a_1) * f(a_2), \tag{2.4}$$

For example, the map from the integers $\mathbb{Z}$ under addition, to the cyclic group $\mathbb{Z}_{12}$, formed by projecting each element from the former group onto an element of the latter group by taking the value modulo 12, is a homomorphism.

The *kernel* of f consists of the elements of A that are mapped onto the identity element of B:

$$Ker f := \{a \in A | f(a) = e_B \in B\}. \tag{2.5}$$

The *image* of f is given by:

$$Im f := \{b \in B | b = f(a) \,\text{for}\, a \in A\}, \tag{2.6}$$

which consists of those elements of B which are maps of elements of A. If the homomorphism f is a bijection (one-to-one correspondence), then one can show that its inverse is also a homomorphism, and f is called a *isomorphism*. In this case, the groups G and H are called isomorphic: they differ only in the notation of their elements. In an isomorphism f, $Kerf = e$. An isomorphic projection of a group onto itself is called an automorphism. For example, the mapping of $\mathbb{Z}$ whereby every integer is multiplied by -1, projects $\mathbb{Z}$ onto itself and is an automorphism.

A subset H of a group G is said to be a *subgroup* of G if it forms a group under the binary operation of G. Equivalently, $H \subseteq G$ is a subgroup if the following conditions hold:

- The identity element e of G lies in H.
- If $x \in H$, then its inverse x^{-1} in G lies in H.
- If $x, y \in H$, then their composition $x \circ y$ in G lies in H.

For example, $2\mathbb{Z}$, the group of all even integers under addition is a subgroup of $\mathbb{Z}$, the group of integers under addition. The set e as well as the group G itself, are also subgroups of G. Going back to the example of C_{12} representing the 12-tone scale, we have seen that the elements 1, 5, 7 and 11 can generate the group. The rest of the elements of the set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ generate subgroups of C_{12}. To illustrate: the elements 2 and 10 are both of period 6 (this means $(2)^6 \bmod 12 = (10)^6 \bmod 12 = e$), they both generate the set $\{0, 2, 4, 6, 8, 10\}$. This set can be interpreted as a whole tone scale and constitutes the group C_6, a subgroup of C_{12}. Similarly, 3 and 9 are of period 4 and generate the subgroup $C_4 : \{0, 3, 6, 9\}$ which could represent a diminished seventh chord. The other subgroups of C_{12} are $C_3 : \{0, 4, 8\}$, generated by 4 and 8 and corresponding to an augmented triad, and $C_2 : \{0, 6\}$, the tritone that generates only itself and the identity.

Given a subgroup $H = \{h_1, h_2, \ldots\}$ of a group G, a *(left) coset* of H written gH, with $g \in G$, is defined as the set of elements obtained by multiplying all the elements of H on the left by g:

$$gH := \{gh_1, gh_2, \ldots\}. \tag{2.7}$$

A *normal subgroup* H of G is one which satisfies

$$gHg^{-1} = H, \qquad \text{for all } g \in G. \tag{2.8}$$

This means that the sets of left and right cosets of H in G coincide. Necessarily, all subgroups N of an abelian group G are normal since $gHg^{-1} = gg^{-1}H = H$.

The *quotient group* or *factor group*, of G over its normal subgroup H is defined by the set of all cosets of a subgroup H, and is intuitively understood as a group that "collapses" the normal subgroup H to the identity element. The quotient group is written G/H. For example, the quotient group $\mathbb{Z}/2\mathbb{Z} = \mathbb{Z}_2$ is the cyclic group with two elements. This quotient group is isomorphic with the group C_2.

The direct product of the groups A and B is the group $A \times B$ whose elements are ordered pairs (a, b) with $a \in A$ and $b \in B$. The group operation is defined by

$$(a_1, b_1)(a_2, b_2) = (a_1 a_2, b_1 b_2). \tag{2.9}$$

The inverse of (a, b) is (a^{-1}, b^{-1}) and the identity element is formed by the identity elements of A and B. For example, the direct product $\mathbb{Z} \times \mathbb{Z}$, which is the lattice $\mathbb{Z}^2$, has identity element (0,0). The group operation is a simple vector addition[2]: $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$, and the inverse of an element (a, b) is $(-a, -b)$.

For a formal description of group theory, one can look at for example Jones (1990) or Alperin and Bell (1995). Furthermore, both Balzano (1980) and Carey and Clampitt (1989) have given useful introductions to group theory which has been directly applied to music, and Benson (2006) has written a whole chapter dedicated to group theory and music.

2.2 Group theoretic and geometric description of just intonation

In this section, which is based on Honingh (2003a) we will use group theory to describe just intonation and our main focus will be on just intonation to the 5-limit. We will make a representation of all intervals in this tuning system within one octave. We will see that this representation is isomorphic with $\mathbb{Z}^2$ so that it can be shown as a two-dimensional lattice. This two-dimensional representation of a tone space will form the basis of some considerations in the field of tuning in section 2.2.2 and will be used as a reference tone system throughout this thesis.

2.2.1 Just intonation in group theoretic terms

Musical intervals can be expressed in terms of frequency ratios. Any positive integer a can be written as a unique product $a = p_1^{e_1} \cdot p_2^{e_2} \dots p_n^{e_n}$ of positive integer powers e_i of primes $p_1 < p_2 < \dots < p_n$. Hence, any rational number can be expressed as

$$2^p 3^q 5^r \dots \tag{2.10}$$

with $p, q, r \in \mathbb{Z}$. For example $2^{-1}3^1 (= \frac{3}{2})$ represents a perfect fifth and $2^{-2}5^1 (= \frac{5}{4})$ a major third. Tuning according to rational numbers is referred to as just intonation (Lindley 2005). If the highest prime that is taken into account in describing a set of intervals is n, then we speak about n-limit just intonation.

[2]We use the term 'vector addition' here, since $\mathbb{Z}^2$ is considered as a subgroup of the vector space $\mathbb{R}^2$. However, note that the lattice $\mathbb{Z}^2$ itself is not a vector space. See Mazzola (1990) for information on linear algebra of modules. We will continue to use the terms 'vector' and 'vector addition' in this chapter at occurrences that are similar to this.

Considering powers of the first two primes, we can create the following set of numbers and ratios

$$\{2^p 3^q | p, q \in \mathbb{Z}\}, \tag{2.11}$$

in which integers like $1, 2, 3, 4, 6, \ldots$ can be found, but also fractions like $\frac{3}{2}, \frac{4}{3}, \frac{9}{8}$ etc. It is possible to rewrite $2^p 3^q$ as:

$$2^p 3^q = 2^{p+q}(\frac{3}{2})^q = 2^u(\frac{3}{2})^v \tag{2.12}$$

with $u = p + q$ and $v = q$. If we consider all numbers resulting from expression 2.11 to be frequency ratios, we see now that those intervals can all be built from a certain number of octaves($\frac{2}{1}$) and fifths ($\frac{3}{2}$). This is called Pythagorean tuning, and is a special form of just intonation. The set of numbers in 2.11 form a subgroup of the positive rational numbers with respect to multiplication. Using group theory we can make an abstraction of several tuning systems. We will call (2.11) the group P_2 (since all elements are powers of the first two primes) and write:

$$P_2 = \{2^p(\frac{3}{2})^q | p, q \in \mathbb{Z}\}. \tag{2.13}$$

The group has identity element 1 and the inverse of an element a is a^{-1}. It is an abelian group (i.e. the elements commute) with an infinite number of elements.

The group P_3 (taking into account the first three primes) can be defined by $\{2^p 3^q 5^r | p, q, r \in \mathbb{Z}\}$ or, equivalently by

$$P_3 = \{2^p(\frac{3}{2})^q(\frac{5}{4})^r | p, q, r \in \mathbb{Z}\}. \tag{2.14}$$

so that the elements can be seen as all intervals built from octaves ($\frac{2}{1}$), perfect fifths ($\frac{3}{2}$) and major thirds ($\frac{5}{4}$). More groups like this can be defined, taking into account multiples of 7 and higher primes. All groups represent a form of just intonation: P_2 represents 3-limit just intonation, P_3 represents 5-limit just intonation, P_4 represents 7-limit just intonation and so on. Every defined group is a subgroup of the group which takes into account higher primes.

$$\begin{aligned}\{2^p | p \in \mathbb{Z}\} \quad \subset \quad & \{2^p 3^q | p, q \in \mathbb{Z}\} \quad \subset \quad \{2^p 3^q 5^r | p, q, r \in \mathbb{Z}\} \quad \subset \\ & \{2^p 3^q 5^r 7^s | p, q, r, s \in \mathbb{Z}\} \quad \subset \quad \ldots\end{aligned} \tag{2.15}$$

Our focus is on P_3 which represents the intervals in 5-limit just-intonation. It can be shown that the group P_3 is isomorphic to the group $\mathbb{Z}^3$, under the projection: $2^p(\frac{3}{2})^q(\frac{5}{4})^r \mapsto (p, q, r)$. For the proof, see appendix A.1. Hence we can represent the elements of P_3 in a 3-dimensional lattice labeled by the elements of $\mathbb{Z}^3$. For simplification however, we want to consider only the intervals lying within one octave. This means, considering the elements of P_3 lying within the interval $[1, 2)$. To accomplish this, we make a projection $\varphi : P_3 \mapsto \mathbb{Z}^2$ that divides P_3 in equivalence classes of (the same) intervals over all octaves. Of each equivalence class a representative is chosen that lies within the interval $[1, 2)$. For example, the intervals $3/2$, $6/2$, $3/4$, $12/2$ and so on, are in the same equivalence class. The representative of this equivalence class is $3/2$ since this is the interval lying within $[1, 2)$. The projection is given by

$$\phi : 2^p(\frac{3}{2})^q(\frac{5}{4})^r \mapsto (q, r). \tag{2.16}$$

This map is a group homomorphism since

$$\phi(xy) = \phi(x) + \phi(y) \qquad \text{for all } x, y \in P_3 \tag{2.17}$$

The kernel of the map is $\{2^p|p \in \mathbb{Z}\}$ which are the elements that are projected on the unit element $(0, 0)$ of $\mathbb{Z}^2$. The quotient-group

$$\hat{P}_3 = P_3/\{2^p|p \in \mathbb{Z}\} \tag{2.18}$$

is the group of cosets of the subgroup $P_1 = \{2^p|p \in \mathbb{Z}\}$. We write

$$\hat{P}_3 = P_3/P_1 = \{xP_1|x \in P_3\} \tag{2.19}$$

which means that every element of P_3/P_1 is an equivalence class of the elements $\{2^p(\frac{3}{2})^\alpha(\frac{5}{4})^\beta|\alpha, \beta \,\text{fixed}, p \in \mathbb{Z}\}$. From every coset we can choose one representative that lies within the interval $[1, 2)$ (one octave). The group P_3/P_1 is isomorphic to $\mathbb{Z}^2$

$$P_3/\{2^p|p \in \mathbb{Z}\} \cong \mathbb{Z}^2 \tag{2.20}$$

since the map φ is a homomorphism and $\{2^p|p \in \mathbb{Z}\}$ is its kernel. In figure 2.1 the representatives of the elements of P_3 are shown ordered according to the 2-D lattice of $\mathbb{Z}^2$. The lattice is unbounded but only part of it is shown. Notice that (for $r = 0$) on the q-axis the (representatives of the) group $\hat{P}_2 = P_2/P_1$ can be found. The representation of intervals like in figure 2.1 has been referred to as the 'Euler-lattice' or 'harmonic network' which we will discuss in section 2.3. For a representation and discussion of the 3-dimensional Euler-lattice (including the octave coordinate), see for example Noll (1995).

We see from figure 2.1 that all frequency ratios (in 5-limit just-intonation) can be built from perfect fifths (3/2) and major thirds (5/4) (and transposing octaves back)[3]. The q-axis represents the number of perfect fifths, the r-axis the major thirds. Every frequency ratio represents a musical interval, like 3/2 representing the perfect fifth and 5/4 representing the major third. Other note intervals can be obtained from these intervals as well. From figure 2.1 we can for example see that two perfect fifths and one major third up (and one octave down) gives an augmented fourth of ratio 45/32. Interval addition can be seen here as vector addition (all vectors with their origin in $(0, 0) \in \mathbb{Z}^2$). In figure 2.2 is shown that

	25/18	25/24	25/16	75/64	225/128		
40/27	10/9	5/3	5/4	15/8	45/32	135/128	
32/27	16/9	4/3	1	3/2	9/8	27/16	81/64
256/135	64/45	16/15	8/5	6/5	9/5	27/20	81/80
	256/225	128/75	32/25	48/25	36/25	27/25	

Figure 2.1: Lattice constructed according to projection $\phi : 2^p(\frac{3}{2})^q(\frac{5}{4})^r \mapsto (q, r)$

	A4	A1	A5	A2	A6	
p5	M2	M6	M3	M7	A4	A1
m3	m7	p4	p1	p5	M2	M6
D1	D5	m2	m6	m3	m7	p4
	D3	D7	D4	D1	D5	

Figure 2.2: Tone space of note intervals, the letters p, M, m, A and D mean perfect, major, minor, augmented and diminished respectively. It is illustrated that adding a major third to a major seventh results in an augmented second.

a major seventh added to a major third results in an augmented second. These intervals are note intervals, resulting from the Western note name system, as we will see in the next section. Putting figure 2.2 on top of figure 2.1 with the (perfect) unison aligned at frequency ratio 1, it becomes clear which frequency ratios belong to which intervals. Note that the inverses of the intervals lie exactly at the other side of the center 1 (point symmetry).

2.2.2 Different realizations of the tone space

There are several ways to make a homomorphic projection from P_3 to $\mathbb{Z}^2$ (as in expression 2.16). The group $\mathbb{Z}^2$ has generating subset $\{(1,0),(0,1)\}$. This means that all elements from $\mathbb{Z}^2$ can be represented as linear combinations of basis

[3] see also Regener (1973, pp. 49,50).

vectors $(1,0)$ and $(0,1)$. In figure 2.1 these elements (or vectors) are associated with the intervals $\frac{3}{2}$ and $\frac{5}{4}$ meaning that all intervals in this figure can be built from perfect fifths and major thirds (and octaves, to get back to the representatives from the coset that lie in the interval $[1,2)$) (see also Regener 1973, ch. 8). These generating elements are not unique. We can for example write:

$$2^u3^v5^w = 2^{u+v+2w}(\frac{5}{4})^{v+w}(\frac{6}{5})^v = 2^k(\frac{5}{4})^l(\frac{6}{5})^m, \qquad u,v,w,k,l,m \in \mathbb{Z} \tag{2.21}$$

with $k = u+v+2w$, $l = v+w$, $m = v$, such that $(u,v,w) \Rightarrow (k,l,m)$ is a bijective map (i.e. a one to one correspondence), proving that the intervals from P_3 can also be built from octaves (2/1), major thirds (5/4) and minor thirds (6/5). One could ask what all possibilities for generating elements for P_3 would be. If the homomorphic projection from P_3 to $\mathbb{Z}^2$ is given by

$$2^pa^qb^r \mapsto (q,r), \tag{2.22}$$

the elements a and b correspond to the generating subset $\{(1,0),(0,1)\}$ of $\mathbb{Z}^2$. The question of which possibilities exist for values of a and b has two directions. The first direction applies to the possible transformations of generating elements of P_3 (e.g. eq. 2.21) or $\mathbb{Z}^3$ since these groups are isomorphic. The second direction applies to the possible transformation of generating subsets of $\mathbb{Z}^2$. The transformations of $\mathbb{Z}^3$ induce transformations of $\mathbb{Z}^2$, and vice versa, the transformations of $\mathbb{Z}^2$ can be lifted to $\mathbb{Z}^3$. To address the point of transformations of $\mathbb{Z}^2$, that is, what are the possible basis-vectors of this space, it turns out that the area of the parallelogram spanned by the two (alternative) basis vectors (or generating elements) should equal 1, in order to be able to represent every element of $\mathbb{Z}^2$ as a linear combination of those vectors. This is equivalent to saying that the determinant of the matrix with the basis vectors as columns should equal 1 or -1. For the proof, see appendix A.2. Concerning the transformations of $\mathbb{Z}^3$, the question is whether the transformations of $\mathbb{Z}^2$ can be lifted to $\mathbb{Z}^3$. Fleischer (Volk) (1998) has shown that lifted maps exist, with restrictions only in connection with the torsion parts of the underlying space. In the case of the projection from $\mathbb{Z}^3$ to $\mathbb{Z}^2$ there is no torsion involved and thus each map on $\mathbb{Z}^2$ may be lifted[4] to $\mathbb{Z}^3$.

There exists an infinite number of possibilities to choose a basis of $\mathbb{Z}^2$. By choosing the map given by (2.16) we accomplished that the basis vectors were corresponding to the perfect fifth and major third and obtained the lattice shown in figure 2.1. Here, we will follow Balzano (1980) and choose major third and minor third, from which the tone space can be build. Hence, we choose the projection:

$$\varphi : 2^p(\frac{5}{4})^q(\frac{6}{5})^r \mapsto (q,r). \tag{2.23}$$

In figure 2.3a, the space constructed according to this projection is shown.

[4]For other examples of lifting a musical meaningful map to a richer space, see Noll (1995).

			216/125	27/25	27/20	27/16
		144/125	36/25	9/5	9/8	45/32
	192/125	48/25	6/5	3/2	15/8	75/64
128/125	32/25	8/5	**1**	5/4	25/16	125/64
128/75	16/15	4/3	5/3	25/24	125/96	
64/45	16/9	10/9	25/18	125/72	625/576	
32/27	40/27	50/27	125/108			

(a) frequency ratios

			Bbb	Db	F	A
		Ebb	Gb	Bb	D	F#
	Abb	Cb	Eb	G	B	D#
Dbb	Fb	Ab	C	E	G#	B#
Bbb	Db	F	A	C#	E#	
Gb	Bb	D	F#	A#	C##	
Eb	G	B	D#			

(b) note names

			9	1	5	9
		2	6	10	2	6
	7	11	3	7	11	3
0	4	8	0	4	8	0
9	1	5	9	1	5	
6	10	2	6	10	2	
3	7	11	3			

(c) pitch numbers

Figure 2.3: Three representations of tone space: frequency ratio space, note-name space, and space of pitch numbers. In figure b, C is chosen to be the key.

The frequency ratios in the tone space can be identified with note names if a reference note is chosen and identified with the prime interval 1. In turn, all note names can be identified with the keys on a piano, or numbers 0 to 11 representing all 12 equal tempered semitones within an octave. The actual projections can be made after a reference point is chosen. We choose to map the prime interval 1 onto the note name C and onto the pitch number 0. The full projections are then obtained by mapping the generating elements 5/4 and 6/5 onto the E and $E\flat$ for the note names, and onto the elements 4 and 3 for the pitch numbers. With these unit elements the rest of the elements are obtained by using vector addition. In this way, two homomorphic projections from the group of frequency ratios to the

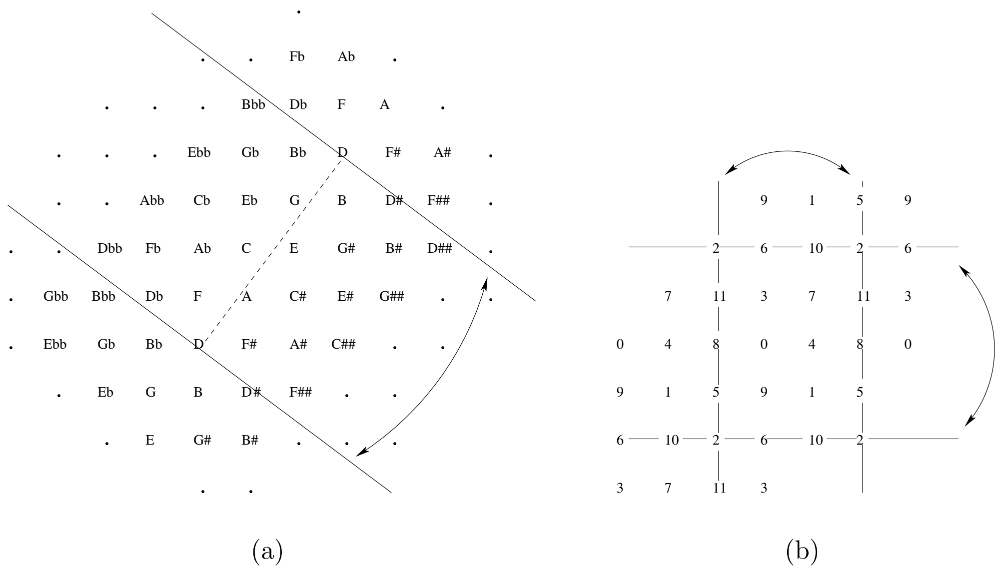


Figure 2.4: Construction of cylinder (a) and torus (b). The points in the tone space that have the same note name or pitch numbers attached to it can be identified. The space can then be rolled up by identifying the lines like indicated in the figure, to become a cylinder or a torus respectively.

group of note names and pitch numbers arise. In figure 2.3, the tone space of note names[5] and the tone space of $\mathbb{Z}_{12}$ pitch numbers are represented according to these projections.[6] In figure 2.3, we left out some complex frequency ratios and note names with many accidentals; however all three spaces can be displayed in infinite horizontal and vertical directions. We already explained that we use octave equivalence in the interval tone space. We do the same in the note name space and in the pitch number space.

All three figures 2.3a, 2.3b and 2.3c show different realizations of the same structure: one step to the right means a major third up, one step up means a minor third up. The frequency ratios space is an infinite space in all directions. If the discrete lattice of note names and pitch numbers are embedded in the continuous space $\mathbb{R}$, the note-name space can be rolled up in one direction by identifying corresponding note names with each other (fig. 2.4)a; the pitch number space can be rolled up in two directions (to become a torus) since the 12-tone system treats enharmonically equivalent notes as the same element (fig. 2.4b). For the note

[5]The terms note-name space and pitch number space are used throughout this thesis to indicate figures b and c. However, in a strict mathematical way, these figures would not be regarded as spaces but perhaps referred to as tables or charts.

[6]see also Regener (1973) for an in-depth analysis of the relation between frequency ratios and note names.

name space, the dimension that is rolled up is generated by the syntonic comma. All multiples of the syntonic comma belong to the kernel of this projection and thus there is only one point on the rolled up circle (indicated by the two D's at both ends of the dotted line in fig. 2.4a). There is no torsion, i.e. there is no element whose multiple is zero. In the case of the pitch number space there are two cases of torsion. Three major thirds as well as four minor thirds add up to zero (or unison in terms of musical intervals). The lines in fig. 2.4b indicate the two cases of torsion. The pitch numbers within the rectangle represent all pitch numbers on the surface of the resulting torus. Thus, we can say that the note name space can be obtained from the frequency ratio space making an identification in one direction; the pitch number space can be obtained from the frequency ratio space by making identifications in two directions. We will further explain this idea in section 3.1.2.

In terms of group theory, the following observations about these three spaces can be made. The set $\mathbb{R}$, representing the real numbers, is a group under addition. Musically, this group can be interpreted as the pitch continuum that consists of the logarithms of all possible frequency ratios (keep in mind that pitch perception is logarithmic with respect to fundamental frequency, see section 1.2). Octave classes of pitches correspond the group $\mathbb{R}/\mathbb{Z}$, i.e. the real numbers mod 1. Subgroups of this group can now be recognized in the spaces of frequency ratios, note names, and pitch numbers in the following way. The group consisting of the (base 2) logarithms of the elements of P_3 (representing the frequency ratios from 5-limit just intonation) under addition, will be denoted by $\log_2(P_3)$. The set of octave classes of elements of this group, denoted by $\log_2(P_3)/\mathbb{Z}$ or equivalently by $\{\log_2(2^p 3^q 5^r) \bmod 1 | p, q, r \in \mathbb{Z}\}$, forms a subgroup of $\mathbb{R}/\mathbb{Z}$. The group $\mathbb{Z}$ under addition, can represent the note names by ordering them in a line of fifths, as we have seen (eq. 2.2). Since this projection from frequency ratios to note names is generated by the fifth (with frequency ratio 3/2), the note names can also be indicated by the group $\{(\frac{3}{2})^p | p \in \mathbb{Z}\}$. Therefore, the embedding of the note names in $\mathbb{R}/\mathbb{Z}$ yields $\{p \log_2(\frac{3}{2}) \bmod 1 | p \in \mathbb{Z}\}$. The pitch numbers can be homomorphically embedded into octave classes of pitches ($\mathbb{R}/\mathbb{Z}$) by sending $p \in \mathbb{Z}_{12}$ to $\frac{p}{12} \in \mathbb{R}/\mathbb{Z}$. The group of pitch numbers is then given by $\{\frac{p}{12} \bmod 1 | p \in \mathbb{Z}\}$, which is a subgroup of $\mathbb{R}/\mathbb{Z}$. But while the embedding of note names is a subgroup of the octave classes of $\log_2(P_3)$ this is not the case for the pitch number embedding representing the 12-tone equal temperament. For further discussion on the pitch continuum $\mathbb{R}$ and its subgroups as described above, see Carey (1998, appendix B).

The projections from the frequency ratio space onto the note name space and pitch number space as shown in figure 2.3, are homomorphic projections. Notice that the group of note names ($\mathbb{Z}$) can be represented on a one-dimensional infinite line. The group of pitch numbers ($\mathbb{Z}_{12}$) can be represented on a one-dimensional finite line. However, here we have chosen to represent all three spaces (figure 2.3) in two-dimensions, to make clear how to project the lattices onto each other (see section 4.2.2 for more explanation of the note name space, and section 4.1.2 and Balzano (1980) for more explanation of the pitch number space).[7]

We can now understand (and see from figure 2.3) that our note name system is not sufficient to distinguish between all frequency ratios. From figure 2.3a we see that, going three major thirds up and four minor thirds up (and two octaves down), we are not back on the note we started from - which is suggested by figure 2.3b - but one syntonic comma (81/80) higher. Therefore, in figure 2.3b every note differs one or more syntonic commas from another note with the same name. For example, comparing the ratios 9/8 and 10/9, both corresponding to a D in figure 2.3b gives us a difference in pitch of $\frac{9}{8}/\frac{10}{9} = \frac{81}{80}$ which is equal to 21.51 cents.

The four 2's present in figure 2.3c represent the same note in the pitch number space, but four different frequency ratios in the frequency ratio space (figure 2.3a). The four frequency ratios corresponding to the four 2's are, from left to right, from top to bottom: $144/125, 9/8, 10/9, 625/576$. The note names of these frequency ratios (compared to C) are $E\flat\flat, D, D, C\sharp\sharp$ respectively. The two D's differ one syntonic comma (= 21.51 cents) from each other. The difference in cents between $E\flat\flat$ and $D(10/9)$, and between $D(9/8)$ and $C\sharp\sharp$ is $62,57$ cents, which is called the greater diesis. The difference between $E\flat\flat$ and $D(9/8)$, and between $D(10/9)$ and $C\sharp\sharp$ is 41.06 cents, which is called the lesser diesis. Finally the difference between $E\flat\flat$ and $C\sharp\sharp$ is 103.62 cents[8], which is more than one (equal tempered) semitone!

2.3 Other geometrical representations of musical pitch

There have been various other proposals to geometrically represent musical pitch. In this section we will give a short overview of some prominent geometrical models of musical pitch that have been used. For a more complete overview, see Krumhansl (1990) or Shepard (1982). Musical pitch is a logarithmic function of the frequency of a note. This can be translated to the fact that the ratio of the frequencies describing an interval is more important for the perception of music than the physical frequencies. For example, the interval between a 100 Hz and 150 Hz tone sounds musically similar to the interval between a 200 Hz and 300 Hz, since their ratio is 3/2 in both cases. Therefore, the simplest geometrical model of pitch can be visualized as a (one dimensional) line, representing the logarithm of ascending or descending frequency. The note names in equal temperament can

[7]For further algebraic investigations of tone spaces, see Cafagna and Noll (2003), Balzano (1980), and Benson (2006).

[8]The Pythagorean comma is not visible here, since this comma is constructed from 12 fifths and hence needs more space.

be represented on a line such that they correspond to ascending frequencies

$$\ldots - G - A\flat(G\sharp) - A - B\flat(A\sharp) - B - c - c\sharp(d\flat) - d - d\sharp(e\flat) - \ldots \qquad (2.24)$$

where c is an octave higher than C and so on. To represent the line of pitches from equation 2.24 such that height corresponds to pitch height, a number of spiral configurations of this line have been proposed (among others Drobisch 1855 Shepard 1982, Bachem 1950). In this representation, the line is presented in a spiral, such that tones at octave intervals are located one above the other (see figure 2.5). The circle projection on the plane perpendicular to the axis of

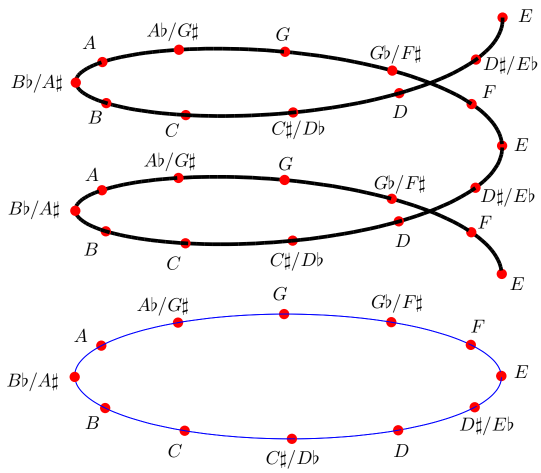


Figure 2.5: Pitch representation of Shepard (1982).

the spiral is sometimes called the chroma circle.

A disadvantage of the above pitch representations is however, that it does not reflect the consonance of the intervals. If proximity should be indicative of consonance, an option would be to represent pitches on the line of fifths, since the fifth is considered to be a consonant interval.

$$\ldots - B\flat - F - C - G - D - A - E - B - F\sharp - C\sharp - G\sharp - \ldots \qquad (2.25)$$

Note that, on the line of fifths, octave equivalence is assumed, in which case a C is equivalent to a c and so on. Many other attempts have been made to model consonance as proximity in geometrical models. The model which has been best well-known is shown in figure 2.6 and emphasizes the major thirds and fifths. A tone in this model has on its left and right its neighbors from the line of fifths, and above and below neighbors from the tones in the sequence of major thirds. This lattice representation and minor variants of it have been introduced in numerous articles (Helmholtz 1863, Riemann 1914, Fokker 1949

D#	A#	E#	B#	F##	C##	G##
B	F#	C#	G#	D#	A#	E#
G	D	A	E	B	F#	C#
Eb	Bb	F	C	G	D	A
Cb	Gb	Db	Ab	Eb	Bb	F
Abb	Ebb	Bbb	Fb	Cb	Gb	Db
Fbb	Cbb	Gbb	Dbb	Abb	Ebb	Bbb

Figure 2.6: Geometrical model of musical pitch referred to as the harmonic network or Euler-lattice.

(who attributes it originally to Leonhard Euler), Longuet-Higgins 1962a 1962b 1987b) and are known under the names 'harmonic network', 'Euler-lattice' and 'Tonnetz'. In this thesis, we will use the term Euler-lattice. As may be clear, the space we obtained in figure 2.3b is a minor variant of the Euler-lattice. It is the same lattice under a basis-transformation[9]. We have seen that this Euler-lattice of note names can be derived from the ratios in 5-limit just intonation (figures 2.3a,b). If the enharmonic equivalent notes are identified with each other, spaces similar to figure 2.3c appear, which have been proposed by Balzano (1980) and Shepard (1982).

Another geometrical model of musical pitch has been developed by Chew (2000) and is called 'the spiral array'. By identifying the points in figure 2.6 indicating the same note names, a vertical cylinder having note names situated on its surface can be obtained[10]. Alternatively, the spiral array can be seen as a spiral configuration of the line of fifths, so that pitches a major third apart line up above each other four steps later. The spiral array is shown in figure 2.7. Chew (2000) has calibrated the parameters in the model such that pitch proximities represent the perceived interval relations in Western tonal music. She furthermore argues that the depth added by going to three dimensions (for example, the key can be represented by a point interior to the spiral) "allows the modeling of more complex hierarchical relations in the spiral array" (Chew and Chen 2005).

Throughout this thesis, we will use the three representations of the tone space or Euler-lattice, as presented in figure 2.3: the tone space of frequency ratios, the tone space of note names which is similar to Chew's (2000) spiral array (although

[9]In the previous section, if we had used the space 2.1 to project the note names onto, we would have obtained figure 2.6

[10]It may be clear that the implied frequency ratios that differ one or more syntonic commas are then identified as well.

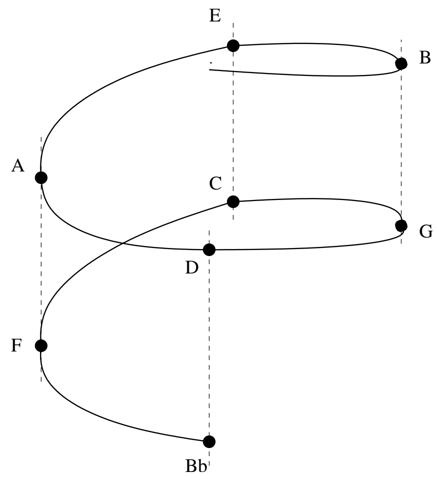


Figure 2.7: Spiral array model developed by Chew (2000).

we will only use the surface of the cylinder representing the note names), and the tone space of pitch numbers which is similar to the torus as suggested by Balzano (1980). As already mentioned in the previous section, the frequency ratios can be projected to the note names, which in turn can be projected to the pitch numbers. The backwards projections however, are ambiguous. The correct projection from pitch numbers to note names and from note names to frequency ratios represent two known problems that we will address in chapters 5 and 6.

Chapter 3

Equal temperament to approximate just intonation

Equal temperament (ET) is the temperament system in which an interval (usually the octave) is divided into a certain number of equal units. The most common equal tempered system consists of 12 tones to the octave. However, in theory, any equal division of the octave is possible. One might wonder, and many music theorists have, if there exist ETs other than the 12-tone ET that can be used for musical practice. Tuning to just intonation can be preferable for its consonance properties. However, ET can be preferred for its practical modulation properties. Combining these two useful properties, an equal tempered system is usually judged on its quality of how well it approximates certain ratios from just intonation. However, other criteria for judging an equal tempered system have been proposed as well (see section 3.1).

Concerning the questions raised in the introduction, an n-tone equal tempered scale that results from imposing certain criteria, gives a possible explanation of the existence of some scales. Furthermore, these criteria may serve as an evaluation of certain existing scales, and finally, the resulting n-tone equal tempered scales can be interpreted as suggestions for new scales to explore. In this chapter, we investigate equal tempered scales that consist of all notes within the temperament system (like the chromatic 12-tone equal tempered scale consists of all notes from the 12-tone ET).[1]

In the first half of this chapter which is based on Honingh (2003b), we will measure the goodness of an n-tone equal tempered system with a designed goodness-of-fit function, based on how well this temperament approximates certain ratios from just intonation. In the second half of this chapter which is based on Honingh (2004) we will concentrate on the Western music notation system combined with equal temperament. We will see that this music notation system restricts the possibilities for n in an n-tone ET system as well. We focus on the problem of

[1] Investigations of scales embedded within an n-tone equal tempered system have been made by for example Balzano (1980) and Krantz and Douthett (2000).

"which pitch number belongs to which note names", which has been illustrated by figures 2.3b,c for 12-tone ET.

3.1 Short review of techniques of deriving equal-tempered systems

Over the past centuries, much has been written about equal tempered systems by music theorists, composers and many others. Many conditions of how to construct the best ET have been proposed. One condition most authors agree on, is to have an equal division of the *octave*, as opposed to an equal division of another interval. Therefore, in this thesis we restrict ourselves to octave based ETs. Besides this, various other conditions have been proposed ranging from good approximation of the perfect fifth (Schechter 1980; Douthett, Entringer, and Mullhaupt 1992), to good scale and chord structures (Erlich 1998) and preservation of the difference between major and minor semitones (Fokker 1955). In this chapter, we do not intend to give an overview of all literature on ET or proposed temperament systems. Instead we focus on two issues 1) the approximation of just intonation ratio by ET, and 2) some limitations on ET caused by Western harmony and the music notation system. An overview and explanation on (equal) temperament systems can be found in Bosanquet (1874a, 1874b).

A mathematical technique known as the *continued fraction compromise*, was used by Schechter (1980) and Douthett, Entringer, and Mullhaupt (1992) to derive divisions of the octave that approximate one specific ratio from just intonation. The perfect fifth was usually chosen to represent this interval since this was considered to be the most important interval after the octave. In section 3.1.1 a short explanation on continued fractions is given. Investigations of how to approximate more than one interval from just intonation were done by De Klerk (1979) and Fuller (1991). Some mathematical functions having the ratios that are to be approximated as input and the desired n-tone ETs as output, have been designed by Hartmann (1987), Hall (1988) and Krantz and Douthett (1994). Hall (1973) judged (not necessarily equal) temperaments with a goodness-of-fit function on the basis of the approximation of just intonation ratios. Yunik and Swift (1980) have written an algorithm to derive the most successful equal tempered scales based on the approximations of 50 consonant ratios from just intonation.

Knowing the values of suitable divisions for the octave (that approximate certain ratios from just intonation), does not explain yet the projection from just intonation to an equal tempered system. Scholtz (1998) explains the relations between just intonation, Pythagorean tuning and equal temperament by giving algorithms to project one onto the other. Regener (1973) provides the conditions for correspondence between pitch notation and regular systems. Constructing an equal tempered system by approximating ratios from just intonation can be understood as tempering out certain commas. Recall from sections

1.3.2 and 1.3.3 that 12-tone ET closes the circle of fifths and tempers out the Pythagorean comma. This means that the frequency ratio 531441/524288 known as the Pythagorean comma, is approximated by unison ($531441/524288 \rightarrow 1$), and the width of the comma is equally divided over tones (in a fifth sequence) lying between the two tones $B\sharp\sharp$ and C. Rapoport (1993) calculates the values of n for n-tone equal tempered divisions by approximating several commas by unison. Erlich (2005) explains that temperament consists in altering the tuning of the frequency ratios so that some of the (small) intervals in the tuning system become perfect unisons. He explains that when tempering out one interval (a comma) from just intonation, the dimensionality of the system is reduced by one dimension. The *dimensionality* of a system is determined by the number of generating elements of the system. Equal temperament is a one-dimensional system: a system where all intervals can be represented as compositions of only one interval (cf. section 2.1.1). Pythagorean tuning is a two dimensional system (its intervals built from powers of 2 and 3, see section 2.2.1), and can be reduced to a one dimensional equal tempered system by tempering out one interval: the Pythagorean comma. To reduce 5-limit just intonation, which is a 3-dimensional system[2] to an ET system, one needs to temper two intervals to unison. Fokker (1955) has described a procedure to project the 5-limit just intonation tone space (fig. 2.3a) to an n-tone equal tempered system, by using *periodicity blocks.* In section 3.1.2 a short explanation on periodicity blocks is given.

3.1.1 Continued fractions

The continued fraction compromise has been used to find temperaments that approximate one interval from just intonation (Schechter 1980; Douthett, Entringer, and Mullhaupt 1992). Especially the fifth, which was considered to be the most important interval after the octave, has been used for approximation. Continued fraction notation is a representation for the real numbers that is finite if the number is rational. Truncating the continued fraction representation early, gives a very good rational approximation. This is an important property of continued fractions and an advantage over decimal representation, since the latter usually does not give a good approximation when truncated.

The continued fraction of a number x is given by

$$x = a_0 + \cfrac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}} = [a_0; a_1, a_2, a_3, \ldots] \tag{3.1}$$

where a_0 is an integer and the other numbers a_i are positive integers. For example, the number 415/93 can be written as $[4; 2, 6, 7]$ in continued fraction form, since $4 + 1/(2 + 1/(6 + 1/7)) = 415/93$. Another example, to show that

[2]Note that in figure 2.3 we have shown this 3-dimensional just intonation system in only 2 dimensions by dividing the octave out.

truncations of continued fractions yield good approximations, is the continued fraction for π. It is an infinite representation, the beginning of which given by $[3; 7, 15, 1, 292, \ldots]$. When this representation is truncated, the approximations yield $3, 22/7, 333/106, 355/113$, which are better approximations to π than truncations of the decimal form $3, 31/10, 314/100, 3141/1000$. To calculate the continued fraction representation of a number r, write down the integer part of r, subtract this integer part from r, find the inverse of the difference, and follow the same procedure again. For example, for the number 2.45, $2.45 - 2 = 0.45$ (write down 2) $\rightarrow 1/0.45 = 2.222.. \rightarrow 2.222.. - 2 = 0.222..$ (write down 2) $\rightarrow 1/0.222.. = 4.5 \rightarrow 4.5 - 4 = 0.5$ (write down 4) $\rightarrow 1/0.5 = 2 \rightarrow 2 - 2 = 0$ (write down 2). The continued fraction representation for 2.45 is $[2; 2, 4, 2]$.

The continued fraction of an irrational number is unique. For finite continued fractions $[a_0; a_1, \ldots, a_n, 1] = [a_0; a_1, \ldots, a_n + 1]$, so there are always two continued fractions that represent the same number. For example $[2; 2, 4, 2] = [2; 2, 4, 1, 1] = 2.45$. As already mentioned, for infinite continued fractions, good approximations are given by its initial segments. These rational numbers are called the *convergents* of the continued fraction. The even numbered convergents are smaller than the original number, the odd numbered convergents are bigger. The first three convergents of $[a_0; a_1, a_2, \ldots]$ are

$$\frac{a_0}{1}, \quad a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1}, \quad a_0 + \frac{1}{a_1 + \frac{1}{a_2}} = \frac{a_2(a_0 a_1 + 1) + a_0}{a_2 a_1 + 1}. \tag{3.2}$$

When some successive convergents are found with numerators $h_1, h_2, \ldots$ and denominators $k_1, k_2, \ldots$, the next convergents can be found by calculating:

$$\frac{h_n}{k_n} = \frac{a_n h_{n-1} + h_{n-2}}{a_n k_{n-1} + k_{n-2}}. \tag{3.3}$$

Semi-convergents include all rational approximations which are better than any approximation with a smaller denominator. If $\frac{h_{n-1}}{k_{n-1}}$ and $\frac{h_n}{k_n}$ are successive convergents, then the fraction

$$\frac{h_{n-1} + a h_n}{k_{n-1} + a k_n} \tag{3.4}$$

is a semi-convergent, where a is a nonnegative integer.

Applied to music, continued fractions are very useful to find the best equal tempered approximation to a certain interval. In an octave based system, the fifth 3/2 measures

$$\log_2 \frac{3}{2} = 0.5849625... \tag{3.5}$$

It may be clear that this irrational number cannot be exactly indicated by an interval in a certain division of the octave. However, it can be approximated, and to find the best approximations, we write the continued fraction as

$$0.5849625... = [0; 1, 1, 2, 3, 4, 1, 5, 2, 23, \ldots] \tag{3.6}$$

The convergents of this representation are given by

$$0, \frac{1}{1}, \frac{1}{2}, \frac{3}{5}, \frac{7}{12}, \frac{24}{41}, \frac{31}{53}, \ldots \tag{3.7}$$

With the semi-convergents included, the sequence becomes

$$0, \frac{1}{1}, \frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \frac{4}{7}, \frac{7}{12}, \frac{10}{17}, \frac{17}{29}, \ldots \tag{3.8}$$

Musically, this means that the fifth is better and better approximated by ETs with $1, 2, 3, 5, 7, 12, 17, 29, \ldots$ notes to the octave. In an ET of for example 12 notes to the octave, the fifth is approximated by 7 steps. Using the continued fraction approach, possible good sounding ET systems can be developed. The other way around, the continued fraction compromise may explain historical choices for temperament systems.

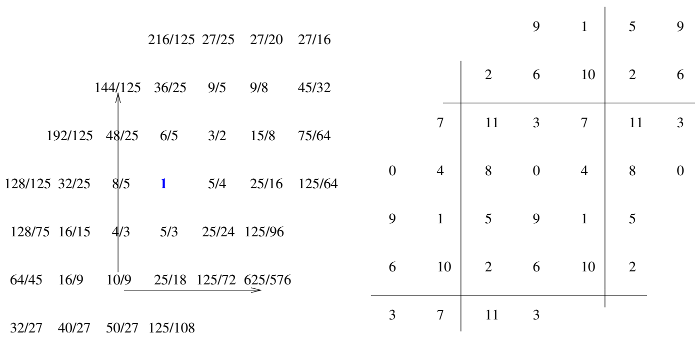
3.1.2 Fokker's periodicity blocks

Periodicity blocks are devices used in a technique for constructing musical scales (Fokker 1969). Although periodicity blocks apply to lattices of just intonation ratios (such as figure 2.3a), in this chapter, we will only study the implications of periodicity blocks for ET.

It has been observed that ETs temper certain commas to unison (see section 1.3.2). When this is turned around, ETs can be obtained by choosing certain commas that are to be tempered to unison. As mentioned before, the dimensionality of a system is reduced by one dimension, when a comma from just intonation is tempered to unison.

In an equal tempered system a finite number of pitches (within one octave) exists. Therefore, a finite number of just intonation ratios within one octave is approximated in such a system. Suppose that we want to construct an equal tempered system by choosing ratios from just intonation that we want to approximate. We add more and more ratios from just intonation until we find a ratio that is very close to a ratio we already have and decide that this ratio is therefore not worth adding. The difference between these two ratios then defines the *unison vector*. The unison vector represents the chosen interval or comma that will be tempered out. In Pythagorean tuning (presented by the diagonal axis of the tone space, fig. 3.1), where all ratios are constructed from powers of 2 and 3, one unison vector is enough to limit the sequence of ratios. However, in 5-limit just intonation two unison vectors are required to limit the group of intervals.

Recall from section 2.2.2 that the 12 tone equal tempered system results from identifying the D and the $C\sharp\sharp$, their difference representing the smaller diesis 128/125, and identifying the D and $E\flat\flat$ whose difference represents the greater



(a) Unison vectors (b) Periodicity block

Figure 3.1: Construction of periodicity block from unison vectors.

diesis 648/625.[3] These two commas represent the unison vectors for constructing the 12-tone ET, and establish that three major thirds equal one octave (or unison when using octave equivalence) and four minor thirds equal one octave (or unison). The two vectors together span a parallelogram (in this case, a rectangle) in which exactly 12 ratios lie. The parallelogram can be moved a certain number of steps downwards or to the left/right without changing the intervallic distances inside the parallelogram. In figure 3.1 the parallelogram with the above chosen unison vectors is drawn in the tone space. This is an example of a periodicity block. The whole tone space can be divided into periodicity blocks when lines are drawn parallel to the unison vectors that are defined. Then the space is divided into blocks that all represent the same intervallic distances. Every block represents the same intervals in ET. Since the commas are tempered out, the space can be identified along the lines of the unison vectors. The space can therefore be rolled up to become a torus. Periodicity blocks can be created in many ways by choosing different unison vectors.

The link between periodicity blocks and ET can now be formulated as follows. By choosing unison vectors, the values of the remaining notes can be filled in by linear interpolation to give the ET version of the scale; the equal tempered values are then regarded as approximations to the just intonation values.

3.2 Approximating consonant intervals from just intonation

Many different temperament systems have been developed in the past. Nowadays, for western music, all keyboard instruments are tuned in equal temperament

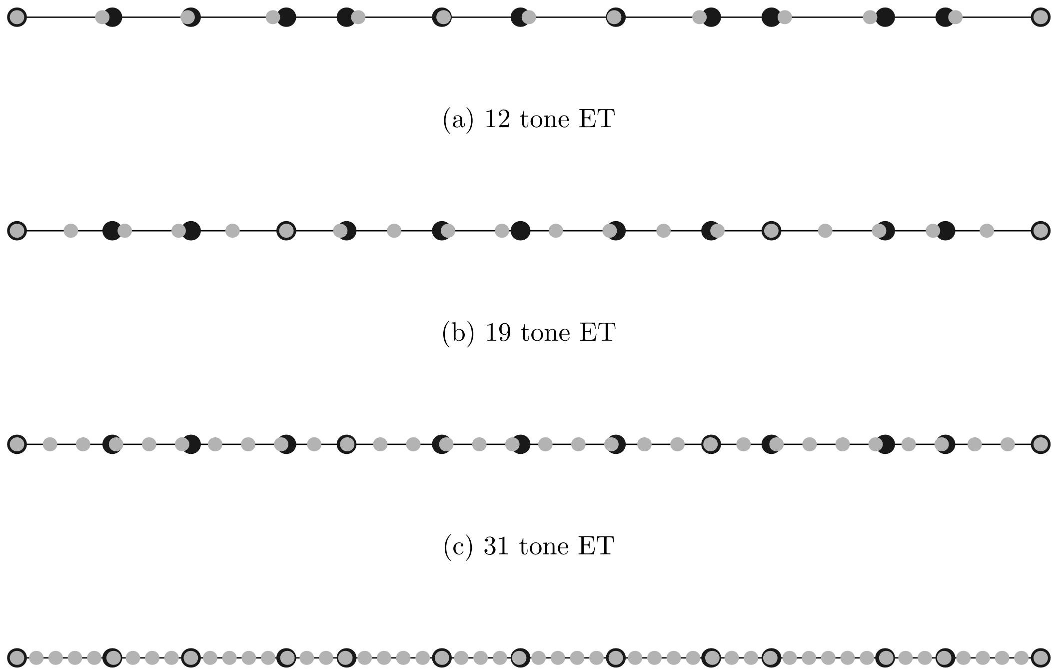
[3]The Western 12-tone ET can equally well be obtained by using the Pythagorean comma 531441/524288 and the syntonic comma 81/80 as vanishing commas, and even other combinations of commas are possible as well, see for example Erlich (2005).

where the octave is divided into 12 equal parts. Other musicians such as singers or string players, tend to play in what is called 'just intonation' (see chapter 1). This relies on the widely established idea that two tones sound best for the ear if they have a simple frequency ratio. It is well known that for keyboard instruments, it is not possible to tune to just intonation. This is inherent to the fact that just intonation cannot be described as a closed finite linear system where intervals can be added to each other so as to obtain the requested sum interval within that finite system (see section 1.3). For keyboard instruments a tuning system has to be developed which approaches just intonation as well as possible. In our 12-tone ET system the ratio of the octave is exactly $2:1$ and the fifth is approximated very closely. Since these two intervals are generally judged as the most important ones, and since all other intervals can be matched to the basic intervals from just intonation by an approximation acceptable to the ear, this temperament system is considered to be a good tuning system.

However, several researchers have investigated whether this temperament system could be improved such that more ratios from just intonation can be approximated more closely. One way to do this is to create an ET system with a division different from 12 tones per octave. To investigate what would be a suitable number of parts to divide the octave into, one has to find out what equal division of the octave approximates certain intervals from just intonation. To illustrate how well the 12-tone equal tempered system approximates certain ratios from just intonation figure 3.2a represents the octave along a line together with black points indicating the locations of the ratios of the chromatic scale defined by Vogel (1975), and gray points indicating the locations of the equal tempered tones according to 12-tone ET. Other ETs like 19-tone, 31-tone and 53-tone have also been proposed. These temperaments are visualized in figures 3.2b,c,d. We want to emphasize that, on the basis on figure 3.2, it is very difficult to decide on the best temperament. For example, it is not clear whether the ratios from the chromatic just intonation scale represent the desired intervals that have to be approximated (more, less or other just intonation intervals could have been chosen as well). Furthermore, the extent to which an n-tone equal temperament approximates these just intonation ratios, is not necessarily based on the average approximation of each individual interval. For example, it may be desirable to approximate the consonant intervals closer than the dissonant intervals. To sum up, to be able to select the best temperament from 3.2, certain assumptions have to be made or personal taste has to be expressed.

In section 3.1, we have given a short overview of authors who have investigated the possibilities of deriving n-tone ETs by approximating intervals from just intonation. Remarkably, most authors do not give a formal derivation of the intervals they choose to approximate in their ETs. Furthermore, at simultaneously fitting intervals, weightings are chosen without any motivation (Hall 1988; Krantz and Douthett 1994).

In this section, an attempt is made to formalize the choice of intervals to



(a) 12 tone ET

(b) 19 tone ET

(c) 31 tone ET

(d) 53 tone ET

Figure 3.2: Several equal tempered approximations to the 12-tone chromatic scale in just intonation. Black points indicate the ratios of this scale: $1, 16/15, 9/8, 6/5, 5/4, 4/3, 45/32, 3/2, 8/5, 5/3, 9/5, 15/8, 2$ (represented on the line by their base-2 logarithmic value), gray points indicate their approximations in n-tone ET.

approximate in equal tempered scales. Furthermore, two different weight models for simultaneously fitting intervals are presented and a number of experiments are performed to find the best equal division of the octave. This procedure is in accordance with the following musical intuition: The intervals from just intonation that should be approximated in equal tempered scales are the ones that are important in (Western) music. The more important/preferred the interval, the more weight should be applied to the fitting procedure.

3.2.1 Measures of consonance

As already mentioned, we are interested in the following question. Which ratios from just intonation should be approximated in equal temperament? One could claim that the consonances (fifth, fourth, major third, minor third, major sixth, minor sixth) according to music theory are to be approximated since these are most important. A more complete answer to the question perhaps is: all intervals from just intonation that appear in (Western) music. But which intervals are those? A possible answer is: the intervals that appear in the major scale in just intonation. Table 1.2 shows the ratios of the notes in the just major scale compared to the fundamental. Counting also the intervals internal to the scale (which means, for example taking the interval 10/9 into account which appears between the 'sol' and the 'la'), these intervals could represent the set of intervals most common in Western music. Let us call this set S_1 (see Table 3.1). Another possibility is to take the frequency ratios of the intervals appearing in the harmonic- or overtone series within a certain number of harmonics. Here the first nine harmonics are considered. Let us call this set S_2 (see Table 3.1). These

S_1	S_2
octave (2/1)	octave (2/1)
fifth(3/2)	fifth (3/2)
fourth (4/3)	fourth (4/3)
major third (5/4)	major third (5/4)
minor third (6/5)	minor third (6/5)
major sixth (5/3)	major sixth (5/3)
minor sixth (8/5)	minor sixth (8/5)
major whole tone (9/8)	major whole tone (9/8)
minor whole tone (10/9)	minor seventh (9/5)
major seventh (15/8)	sub minor seventh (7/4)
minor seventh (9/5)	sub minor third (7/6)
diatonic semitone (16/15)	super second (8/7)
augmented fourth (45/32)	sub minor fifth (7/5)
diminished fifth (64/45)	super major third (9/7)
subdominant minor seventh (16/9)	
Pythagorean minor third (32/27)	
Pythagorean major sixth (27/16)	
grave fifth (40/27)	
acute fourth (27/20)	

Table 3.1: Set S_1: intervals coming from the just major scale, and set S_2: intervals appearing in the harmonic series up to the ninth harmonic. The names of the intervals are taken from Helmholtz (1863).

tables represent the ratios from just intonation to approximate in ET, but what is the order of importance of these intervals? For example, is it more important for a certain ratio to be closely approximated than for another, or are they all equally important? From the definition of just intonation, it can be understood that there is a preference for simple frequency ratios. Accordingly, the fifth 3/2 is more preferred than the major second 9/8. So there is an order of importance. But how to compare ratios like 6/5 and 7/4? For several ratios there is no consensus according to the definition of just intonation. No unique function exists that describes the order of preferred ratios. A number of hypotheses and theories have been proposed regarding the origin of consonance and dissonance. Two prominent theories are the ones by Helmholtz and Euler which are constructed from a physiological and psychological perspective respectively. In order to let the tension-resolution process in music have the greatest effect, one could say that the most important intervals are the most consonant intervals and therefore those should be approximated most closely. Here, Euler's Gradus function and Helmholtz' roughness function are used to put the two sets of intervals in an order of preference or, in the way it is called throughout the thesis, an order of consonance. See Krumhansl (1990, pp. 55-62) for an overview of several measures of consonance, each obtaining an order of consonance for the intervals from the chromatic 12-tone scale.

Euler's Gradus function

Euler developed a Gradus function Γ that applies to whole frequency ratios x/y from just intonation (Euler 1739). The function is defined as a measure of the simplicity of a ratio.

Any positive integer a can be written as a unique product $a = p_1^{e_1} \cdot p_2^{e_2} \dots p_n^{e_n}$ of positive integer powers e_i of primes $p_1 < p_2 < \dots < p_n$. Euler's formula is defined as:

$$\Gamma(a) = 1 + \sum_{k=1}^{n} e_k(p_k - 1) \tag{3.9}$$

and for the ratio x/y (which should be given in lowest terms) the value is $\Gamma(x \cdot y)$. Applied to the problem of consonance this means that the lower the value Γ the simpler the ratio x/y and the more consonant the interval. In accordance with eq. 3.9, the order of consonance for set S_1 and set S_2 is given in Table 3.2.

S_1	S_2
2/1	2/1
3/2	3/2
4/3	4/3
5/4 , 5/3	5/4, 5/3
6/5, 9/8, 8/5	6/5, 9/8, 8/5
16/9	7/4
10/9, 9/5, 15/8	7/6, 8/7, 9/5
16/15, 27/16	7/5, 9/7
32/27	
27/20	
45/32, 40/27	
64/45	

Table 3.2: Order of consonance for ratios of set S_1 and S_2 according to Euler's Gradus function, from most to least consonant.

Note that there is not a perfect match with musical intuition. For example the major whole tone (in music considered to be a dissonant tone) is placed on the same level as the minor third (a consonance according to music theory). This is because Euler's function judges only the simplicity of the ratio.

Helmholtz's roughness function

Helmholtz defined the roughness of an interval between tones p and q on the basis of the sum of beat intensities $I_n + I_m$ associated with the n^{th} harmonic of p and the m^{th} harmonic of q (Helmholtz 1863). This roughness depends on the ratio n/m, but also on the intensity of the harmonics (and therefore on the type of sound) and on the register of the tones (in lower positions, intervals tend to sound more rough). Helmholtz calculated the roughness of intervals in the c'-c" octave, and based the intensity of the harmonics on violin sound. The roughness of two tones is expressed by:

$$r_p = 16B'B'' \cdot \frac{\beta^2\theta^2\delta^2p^2}{(\beta^2 + 4\pi^2\delta^2)(\theta^2 + p^2\delta^2)}, \tag{3.10}$$

where p refers to the p^{th} partial tone, β is the coefficient of damping, and B' and B'' are the greatest velocities of the vibrations which the tones superinduce in Corti's organs in the ear, which have the same pitch. Furthermore, $\theta = 15/264$ and $\delta = \frac{n_1 - n_2}{2}$, where n_1 and n_2 are the number of vibrations in 2π seconds for the two tones heard. Applying this formula to all the partials of the motion of the bowed violin string, which Helmholtz had measured by means of the vibration microscope, he obtained a diagram for the predicted dissonance of pairs of notes over a continuous range of two octaves. The order of consonance of set S_1 and S_2 according to this roughness function is given in Table 3.3[4].

Again, one would expect this ordering to coincide with a musical intuition, but an interval known as dissonant (augmented fourth 45/32) is placed on the same level as two consonances (the minor third 6/5 and the minor sixth 8/5). Note also that the fifth and the octave are judged to be equally consonant. The ordering according to Helmholtz differs from the ordering according to Euler (Table 3.2). Remarkably, for both measures, the preference for lowest numbers suggested by the definition of just intonation is not entirely followed. For example, for both measures the ratio 16/9 is placed on a more consonant level than the more simple, and therefore expected to be more consonant ratio 9/5.

3.2.2 Goodness-of-fit model

A goodness-of-fit model can be developed to measure which ET best approximates ratios from just intonation. Several functions have been defined to measure the

[4]Since Helmholtz (1863) did not consider the interval ratio 9/7, it is not included in Table 3.3.

S_1	S_2
2/1, 3/2	2/1, 3/2
4/3	4/3
5/3	5/3
5/4	5/4
6/5, 8/5, 45/32	7/4
16/9	6/5, 8/5
27/16	9/5, 7/6, 7/5
9/5	8/7
32/27	9/8
27/20	
64/45	
9/8	
10/9	
15/8	
40/27	
16/15	

Table 3.3: Order of consonance for ratios of set S_1 and S_2 according to Helmholtz's roughness function, from most to least consonant.

goodness of a given n-tone equal-tempered scale (Hartmann 1987; Hall 1988; Krantz and Douthett 1994). Given an n-tone ET (the octave divided into n equal parts), the ratio R is best approximated by $2^{m/n}$ when the error

$$E = |\log_2 R - \frac{m}{n}| \tag{3.11}$$

is as small as possible. The number of steps m in an n-tone scale that minimizes eq. (3.11) is

$$m = \lfloor n \cdot \log_2 R + 0.5 \rfloor, \tag{3.12}$$

where $\lfloor x \rfloor$ is the largest integer less than or equal to x. With eq. (3.12) substituted into eq. (3.11), the function E is defined as:

$$E(R, n) = |\log_2 R - \frac{1}{n}(\lfloor n \cdot \log_2 R + 0.5 \rfloor)|. \tag{3.13}$$

This is, in fact, already a measure of the goodness of an n-tone scale for a ratio R; the lower the error E the better the fit. Since the function (eq. 3.13) applies to ETs, it necessarily yields the same values for an interval and its musical inverse (for example a fifth 3/2 and a fourth 4/3). For a perfect fit, $E = 0$, the maximum error for a fit is $E = \frac{1}{2n}$. To obtain the error in cents, one just multiplies eq. (3.13) by 1200.

Previous models

Measures based on (variation of) the error function (eq. 3.13) have previously been made by Hartmann (1987), De Klerk (1979), and Yunik and Swift (1980).

Hall (1988) criticizes these measurements for the fact that for all sufficiently large n, the fit is equally good. But, Hall (1988) argues, for reasons of pitch discrimination, high values for n are usually rejected. Therefore Hall, and Krantz and Douthett stated that a correction should be made by multiplying the error E by n (Hall 1988; Krantz and Douthett 1994). They both developed a function to simultaneously fit multiple intervals. Hall (1988) presented his Remarkability function:

$$R = \prod(-\log_2 p_{ij}), \qquad p_{ij} = \frac{2e_{ij}}{S_N}, \tag{3.14}$$

where $S_N = 1200/N$ is the smallest step size available in N tone ET, and $e_{ij} = 1200 \cdot E$ (E as in eq. 3.11 with $R = i/j$). Krantz and Douthett (1994) proposed their generalized Desirability function:

$$D(R, n) = 10 - 20 \sum_{i=1}^{N} p_i |n \cdot \log_2 R_i - q_i|, \tag{3.15}$$

where p_i is the weight of each target interval R_i, and $\sum p_i = 1$, and $q_i = \lfloor n \log_2 R_i + 0.5 \rfloor$ (cf. eq. 3.12). In section 3.2.3 some results of the models of Hall and Krantz and Douthett are discussed.

Goodness-of-fit function

In this thesis, a number of experiments are carried out with our goodness-of-fit function that builds on these previous models. Differently from the models (3.14) and (3.15), it is believed that the rejecting rate of an n-tone scale is not per se linear in n but can vary depending on the purpose of the scale. Thus, the correction of multiplying the error E by n is not made here and the decision as to what is the best scale can be made at a later stage.

Analogous to the previous models a weight p_i for each ratio R_i is introduced, such that

$$\sum p_i = 1. \tag{3.16}$$

Different ratios can then be weighted according to consonance. For each n-tone equal tempered scale a total error $\sum_i p_i E(R_i, n)$ can now be constructed.

Since one rather wants to obtain a high value from our function when the fit is good and a low value when the fit is bad, and to make the difference between the fits more visible, it makes sense to use $(\sum_i p_i E(R_i, n))^{-1}$ as a goodness measure. However, the values of this goodness are not bounded. Therefore the final expression of the goodness-of-fit function used for the evaluation of n-tone temperaments is presented as:

$$f(n) = \frac{1}{0.01 + \sum_i p_i E(R_i, n)}. \tag{3.17}$$

The value of $\sum_i p_i E(R_i, n)$ is bounded between 0 and $\frac{1}{2n}$, therefore $f(n)$ can vary between (approximately) 2 (worst fit for $n = 1$) and 100 (perfect fit). The measure (eq. 3.17) can be compared to other measures since the scaling $\frac{1}{0.01+x}$ is a monotonically decreasing function for $x = \sum_i p_i E(R_i, n) \in [0, 1/2]$.

3.2.3 Resulting temperaments

The goodness-of-fit model presented in equation 3.17 can now be used to judge n-tone ET systems, given a set of intervals and a weight function. The function in equation 3.17 can be plotted as a function of the number of notes n to be able to compare the goodness values. Emphasizing that a variety of choices is possible for the set of intervals and the weight function, we will give some examples here.

Since the fifth (3/2), fourth (4/3), major third (5/4), minor third (6/5), major sixth (5/3) and minor sixth (8/5) are the consonant intervals according to music theory, this has been a general motivation to construct temperament systems that approximate (subsets of) these intervals best (Krantz and Douthett 1994; Hall 1988). Taking these intervals as input for equation 3.17 with equal weighting, figure 3.3 shows the result. Considering the values for n which yield better and

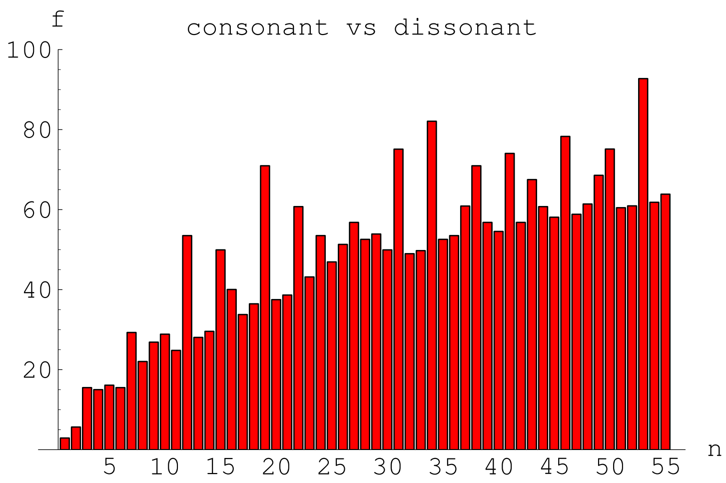


Figure 3.3: Values of f calculated from the consonances according to music theory: fifth, fourth, major third, minor third, major sixth, minor sixth, with equal weighting.

better approximations to the chosen intervals, peaks can be seen for the values $n = 12, 19, 31$ and 53, which are all covered by previous literature. For example Krantz and Douthett (1994) get the exact same results with their desirability function except that they do not notice the goodness of $n = 31$ due to the correction they made for high values of n. The good fit for $n = 12$ could explain the fact that this temperament is generally used nowadays. Furthermore, the 19, 31 and 53 tone temperament have also been recognized and used by theorists and composers (Bosanquet 1876; Yasser 1975; Huygens 1691).

Then, the sets of intervals obtained in section 3.2.1 together with a weighting according to the measures of Helmholtz and Euler are taken as input values for the model (eq. 3.17). The goodness $f(n)$ is calculated for $n \in [1, 55]$. The weights p_i for the ratios R_i are chosen inversely proportional to the values of Euler's Gradus function or Helmholtz's roughness function. The results are shown in Figures 3.4a to d.

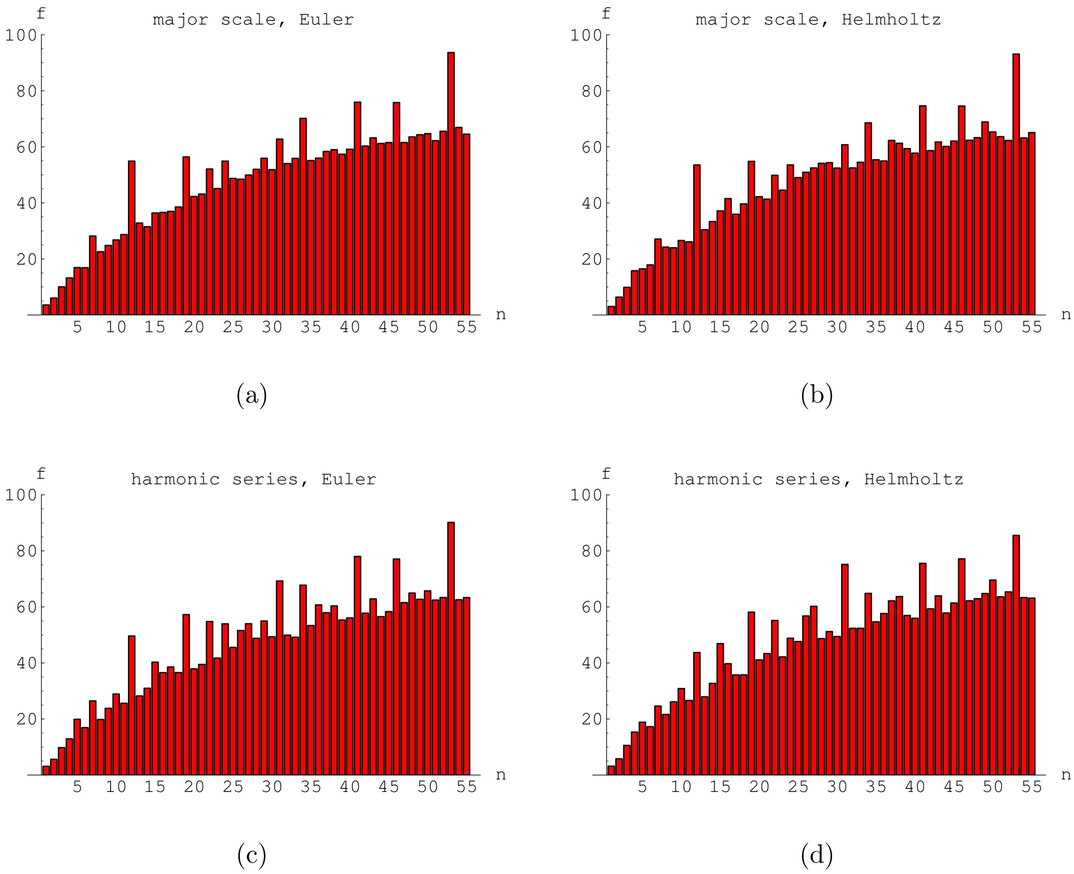


Figure 3.4: Values of f calculated from the set intervals of either the major scale or the harmonic series with weights according to either Euler's or Helmholtz measure of consonance.

If the values of n are considered which yield better and better approximations to the chosen intervals, the values $12, 19, 31, 41$, and 53 appear for all four figures. These peaks refer to good fits for these n-tone scales. The values $12, 41$, and 53 are just the continued fraction solutions to the perfect fifth, and also the other values are covered by previously published results (Hall 1988; Krantz and Douthett 1994). In Figures 3.4a and 3.4b an extra peak appears for $n = 34$, and

Figure 3.4d has also three more peaks for $n = 15$, $n = 27$, and $n = 46$. The peak for $n = 34$ is consonant with what Krantz and Douthett (1994) found. The peak at $n = 15$ is also reported by Krantz and Douthett (1994) who was fitting to the minor third (6/5) and minor seventh (7/4). The good fits for the octave divided into 27 or 46 equal pieces, are at odds with previously published results. However, compared to other peaks they don't stand out in their quality.

Comparing the figures, it can be seen that figures 3.4a and 3.4b are rather similar and differ from Figures 3.4c and 3.4d, the latter two being similar as well. This means that varying the set of intervals has a greater impact than varying the measure of consonance. This is mainly due to the model's property that judges inverse ratios in the same way, although none of the intervals are placed at the same level of consonance as their inverse according to the measures of Euler and Helmholtz. Comparing figures 3.4a to figure 3.3, note that $n = 41$ (the continued fraction compromise for the perfect fifth) occurs in figures 3.4a to d but not in figure 3.3.

As already mentioned, for reasons of pitch discriminations, high values for n are not always favored over lower values. The just noticeable difference (JND) expresses the difference of two consecutive tones (see section 1.2.2). It has been shown that the JND can be as small as two or three cents (Zwicker and Fastl 1990), which would motivate us to cut off n at around 400 (1 cent represents 1/1200 part of an octave, therefore 3 cents represent 1/400 part of an octave). However, studying figure 1.2 we see that for example around 6000 Hz the JND equals 50 Hz which roughly corresponds to 14 cents and leads to a maximum division of 85 notes per octave. Here, we have only shown the results up to $n = 55$ to be able to clearly visualize these results.

This work suggests that testing an n-tone equal-tempered system involves formally choosing a set of intervals and a measure of consonance to apply weights. These choices can in turn depend on the type of music or a special purpose.

3.3 Limitations on fixed equal-tempered divisions

In this section we address another approach to finding a suitable division of the octave for n-tone ET. Historically, it has not always been clear for what reasons certain choices for n-tone equal tempered systems have been made. Yunik and Swift (1980) write "Through often convoluted, difficult-to-follow logic, various other values for n have been proposed". In this section limitations on the choice of an n-tone ET are considered that are related to the music notational system. To be able to use an n-tone equal tempered system for a keyboard application using the Western notational system, there are some restrictions to be taken into consideration.

In chapter 2, it was shown that the projection from the frequency ratios to the 12-tone equal tempered pitch numbers constitutes a homomorphic mapping. It may be clear that also in general, the projection from the group of frequency ratios to the finite group of n-tone equal tempered pitch numbers, should be a homomorphism (see eq. 2.4). Ideally, this projection should also be monotonic, such that the ordering of the elements is maintained[5]. However, this is not possible, which we can understand as follows. Consider a homomorphic mapping from the frequency ratios from just intonation to an n-tone ET. In this ET, a specific comma (interval from just intonation) is tempered out (see section 3.1.2). When a frequency ratio is multiplied by such a comma, the ratio is changed. However, when the corresponding pitch number from the n-tone ET is raised (or lowered) by this comma, it stays the same, since the comma was approximated by unison in this n-tone ET. The pitch number stays constant even after adding a high number of this comma. However, if more and more commas are multiplied with the frequency ratio, it changes more and more. Therefore, for some frequency ratios a and b, where $a > b$, it happens that they are projected onto the pitch numbers 1 and 2 respectively $(1 < 2)$, such that the projection is not monotonic. Still, monotonicity of the projection may be possible for a part of the intervals, as we will see.

In this section, our goal is to make a homomorphic projection from the frequency ratio space to the pitch number space such that a surjective mapping from the note names to the pitch numbers result. This surjective mapping is required in order to have a suitable mapping from a score to e.g. the keys of a piano. An ET system of this kind was called a 'negative system' by Bosanquet (1874a, 1874b). Positive systems require a separate notation (Bosanquet 1874a, 1874b). However, even if a notation system is chosen that represents every frequency ratio as a separate note name, limitations related to interval addition still exist, as we will see. This section is based on Honingh (2004).

3.3.1 Attaching note-names to an octave division

The line of fifths represents all possible note-names for musical tones

$$\ldots A\flat - E\flat - B\flat - F - C - G - D - A - E - B - F\sharp \ldots \tag{3.18}$$

[5]An ordered group has a binary relation '$\leq$' satisfying the following:

$$\begin{aligned} & a \leq a \\ & a \leq b \text{ and } b \leq c \Rightarrow a \leq c \\ & a \leq b \text{ and } b \leq a \Rightarrow a = b \\ & a \leq b \text{ or } b \leq a \\ & a \leq b \Rightarrow (a+c) \leq (b+c) \\ & a \leq b \Rightarrow (c+a) \leq (c+b). \end{aligned}$$

A morphism for ordered groups is a homomorphism f that is monotonic: $a \leq b \Rightarrow f(a) \leq f(b)$

which is an infinite series in both directions. In the well-known 12-tone ET, the line of fifths is transformed into a circle of fifths by saying that the $B\sharp$ should be equivalent to the C. After 12 fifths, the circle is finished, as is illustrated in figure 3.5a. Placing the note names in scale order, figure 3.5b illustrates that the note

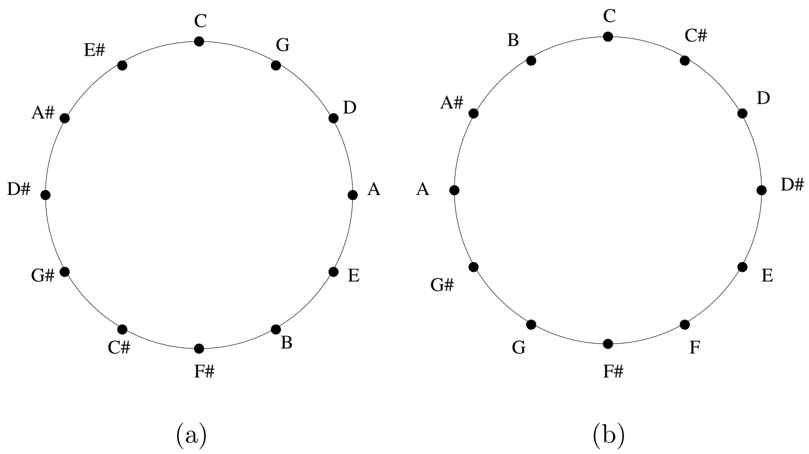


Figure 3.5: Transformation of the line of fifths into (a) the circle of fifths, and (b) the chroma circle.

names are attached to the pitch numbers in the following way:

$$\begin{pmatrix} C & C\sharp & D & D\sharp & E & F & F\sharp & G & G\sharp & A & A\sharp & B \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \end{pmatrix} \tag{3.19}$$

Here we have written all note names with sharps ($\sharp$), but in this 12 tone ET these notes can be interchanged with their enharmonic equivalents (see section 1.3). When generalizing the 12 tone equal tempered system to an n-tone equal tempered system, another mapping from note names to pitch numbers is necessary. We will now investigate how projection should be made.

When an n-tone ET is constructed, it may be not directly clear which note name(s) belong to each unit. Using this line of fifths, a possible way to attach note-names to an equal tempered division is to calculate the number of steps m in the n-tone temperament that approximates the fifth $R = 3/2$ (see eq. 3.12) and attach adjacent note-names from (3.18) to every pitch number

$$k \cdot m \bmod n, \qquad k \in \mathbb{Z} \tag{3.20}$$

of the equal tempered division. This process is illustrated in figure 3.6 for 12-tone ET. In this case, $m = 7$ since $3/2 \approx 2^{7/12}$ (see eq. 3.12). Starting with C, which is attached to pitch number 0, the G is attached to pitch number $0 + 7 = 7$, the D is attached to pitch number $(7 + 7) \bmod 12 = 2$, and so on. In this way, all

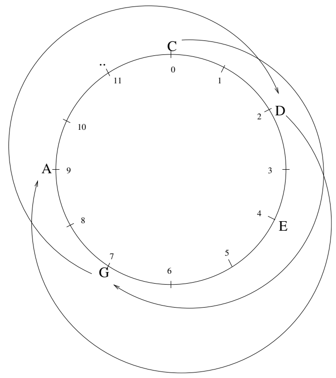


Figure 3.6: Beginning of the process of attaching note-names to the 12-tone equal tempered division. The fifth is approximated by 7 steps in the 12-tone division.

note-names are attached to a certain position on the circle. We can say that we have used the interval of the fifth as a generator of the note names.

Another way to attach note-names to an equal tempered division is to use another interval (than the fifth) as 'generator of note-names'. Figure 3.7 gives an example of the major third $5/4 \approx 2^{4/12}$ distributing note names over the 12-tone division. Using this method of distribution, only one fourth of all possible

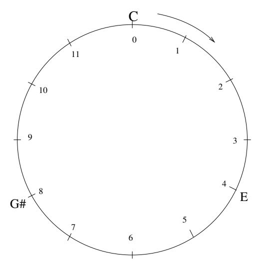


Figure 3.7: Beginning of the process of attaching note-names to the 12-tone equal tempered division using the major third that is approximated by 4 steps in this division.

note-names are used, since the line of thirds

$$\ldots - F\flat - A\flat - C - E - G\sharp - B\sharp - D\sharp\sharp - \ldots \tag{3.21}$$

represents only one fourth of the note names present in the line of fifths (3.18).

In the example of the 12-tone temperament, the note name distributions are in agreement with each other. This means that the note names selected by the major third are in the same position as those note names selected by the perfect fifth. For example, the E is attached to unit number 4 in fig. 3.6 as well as in fig. 3.7. This is not the case for all n-tone divisions. For example for the 15-tone temperament, the third $C - E$ measures 6 (out of 15) units if calculated from

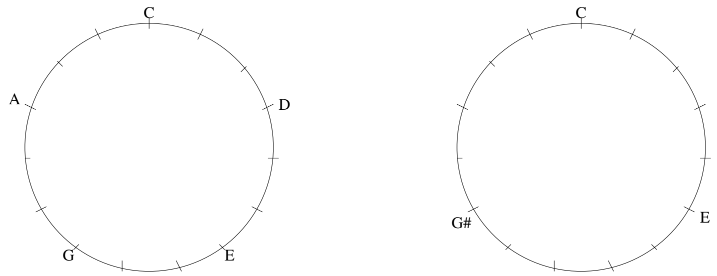


Figure 3.8: Example of 15-tone ET. Left: attaching note-names using fifth, right: attaching note-names using major thirds. Only the beginnings of the distributions (line of fifths and thirds) are shown.

the fifth, and 5 units if calculated from the major third (see figure 3.8). The E is attached to unit number 6 in the left most figure, and attached to unit number 5 in the right most figure. If an equal temperament is to be used for a keyboard application, this is not desirable since there is no consensus about which keys to press when a score is read.

Enharmonicity condition

Ideally, a surjective mapping is made from the note-names to the units of the n-tone equal division. This means, two different note-names can refer to the same unit (and are therefore enharmonically equivalent), but one note name cannot refer to two different units. Furthermore, all units should have a note-name (we come back later to this last condition). In this way, reading a score, it is clear which keys on a piano to press. To gain this result, the fifths have to match with the major thirds as illustrated in the previous section. Also, other intervals (besides the fifth and the major third) can be chosen to generate the note-names, and must therefore match with each other.

To determine whether an n-tone temperament has a good match of thirds and fifths, one has to check if the unit number that approximates the major third (5/4) is equivalent to the unit number corresponding to the Pythagorean third (81/64), which is the third constructed from four perfect fifths (modulo the octave). The difference in cents between the two intervals measures 21.51 cents. Therefore, the better these intervals are approximated, the less chance they both map on the same pitch number of the equal tempered system. And because there is always a better fit if n is chosen big enough, there will be a certain maximum to n for which the major thirds and the fifths still match. The unit-number m that approximates a frequency ratio R in an n-tone ET is given by (cf eq. 3.12):

$$m_R(n) = \lfloor n \log_2 R + 1/2 \rfloor, \tag{3.22}$$

where $\lfloor x \rfloor$ is the largest integer less than or equal to x. Since four fifths up and using octave equivalence gives a major third (C-G-D-A-E), the following equivalence relation can be demanded such that fifths and major thirds match[6]:

$$4 \cdot m_{\frac{3}{2}} \bmod n = m_{\frac{5}{4}} \tag{3.23}$$

Furthermore, three fifths give a major sixth (C-G-D-A), and demanding that fifths should match with major sixths gives the following condition

$$3 \cdot m_{\frac{3}{2}} \bmod n = m_{\frac{5}{3}} \tag{3.24}$$

It is possible to go on writing conditions like this, but first we take a look at what is covered by conditions (3.23) and (3.24).

When we look at the line of fifths:

$$\ldots A\flat - E\flat - B\flat - F - C - G - D - A - E - B - F\sharp \ldots \tag{3.25}$$

and we realize that an equal tempered approximation to an frequency ratio R is just as good as the approximation to the musical inverse of the ratio $2/R$, we understand (and it can be shown) that the following relations hold for the same values of n:

$$\begin{aligned} x \cdot m_{\frac{3}{2}} \bmod n &= m_R \\ x \cdot m_{\frac{4}{3}} \bmod n &= m_{\frac{2}{R}} \end{aligned} \tag{3.26}$$

where $x \in \mathbb{Z}^+$ and $R \in [1, 2)$ ($R \in \mathbb{Q}$). When m steps approximate the fifth, $n - m$ steps approximate the fourth:

$$m_{\frac{4}{3}} = n - m_{\frac{3}{2}} \tag{3.27}$$

Substituting (3.27) in (3.26) gives

$$-x \cdot m_{\frac{4}{3}} \bmod n = m_R, \tag{3.28}$$

and comparing (3.26) and (3.28) shows us that a number of x fifths up gives the same interval as x fourths down (or $-x$ fourths up). Thus equations (3.23) and (3.24) are not only concerned with the matching of fifths, major thirds and major sixths but also with there inverses. Therefore, equations (3.23) and (3.24) make sure that all intervals $3/2, 4/3, 5/4, 6/5, 8/5, 5/3$ match with each other. We will refer to conditions (3.23) and (3.24) together as the enharmonicity conditions. The possible values for n so that (3.23) and (3.24) are true can be obtained by running a simple Perl script:

$$\begin{aligned} n = \quad & 5, 7, 12, 19, 24, 26, 31, 36, 38, 43, 45, 50, 55, 57, 62, \\ & 69, 74, 76, 81, 88, 93, 100 \end{aligned} \tag{3.29}$$

[6]Recall the following intervals and their frequency ratios: perfect fifth 3/2, perfect fourth 4/3, major third 5/4, minor third 6/5, major sixth 5/3, minor sixth 8/5.

As we explained, there is a limit on n for which the intervals match. Apparently, this limit is reached at $n = 100$. This set (eq. 3.29) is a subset of the values for n that Rapoport (1993) found when he searched for ETs that use the Pythagorean comma and diesis, among temperaments that arose from requiring the syntonic comma to vanish. All consonant intervals from music theory (fifth, fourth, major and minor thirds, major and minor sixths) are considered when constructing the enharmonicity conditions. We think that it is not realistic to demand that more intervals should match in the n-tone system, because there is less consensus about the frequency ratios. For example, a major second is denoted by 9/8 but also by 10/9.

The fact that we have searched for ETs which are consistent with music notation, satisfying conditions 3.23 and 3.24, does not mean that other ETs can not be used. Bosanquet (1874a, 1874b) developed a notational system for so-called positive systems: systems that are not consistent with the familiar Western musical notation[7].

Adding intervals

When playing a chord in an n-tone equal tempered system, it is desirable to have all the containing intervals represented by the number of units that approximate the intervals best. For a chord consisting of three notes, the number of units that approximates an interval added to the number of units that approximates another interval should equal the number of units that approximates the sum-interval. For example, a major triad contains the intervals major third, minor third and perfect fifth. Ideally, the number of units in the n-tone ET approximating the major third adds up to the number of units approximating the minor third and equals the number of units that approximates the perfect fifth. In just intonation ratios, adding a major third to a minor third gives a perfect fifth.

$$5/4 \cdot 6/5 = 3/2. \tag{3.30}$$

Translated to the approximations in 12-tone ET this reads:

$$2^{4/12} \cdot 2^{3/12} = 2^{7/12}. \tag{3.31}$$

So here, indeed, the number of units in the 12-tone ET approximating the major third (4 units) adds up to the number of units approximating the minor third (3 units) and equals the number of units that approximates the perfect fifth (7 units). This is not the case in every n-tone ET. For example, in 14-tone temperament (3.30) would be translated to

$$2^{5/14} \cdot 2^{4/14} \neq 2^{8/14}. \tag{3.32}$$

[7]Positive systems form their major thirds by going 8 fifths down (and imposing octave equivalence) such that the ratio 5/4 is approximated by the interval $C - F\flat$.

which means that in 14-tone temperament the major and minor third do not add up to become a perfect fifth. In general, the constraint on n for the correct way of adding intervals R_1 and R_2 is:

$$m_{R_1} + m_{R_2} + k \cdot n = m_{R_3}, \qquad \text{where } R_3 = R_1 \cdot R_2 / 2^k. \tag{3.33}$$

The addition $+k \cdot n$ has the same meaning as 'mod n'. The division by 2^k in the latter equation means rescaling the interval so that it stays within one octave. Although this condition cannot be satisfied for all possible ratios R_1, R_2 for a given n, it will now be verified that for the n-tone systems from eq. 3.29, the condition 3.33 is satisfied for all R_1, R_2 chosen from the (consonant) ratios $3/2, 4/3, 5/4, 8/5, 5/3$ and $6/5$. We can understand this in the following way. Using eq. 3.26, equations 3.23, 3.24 translate into the following six equations:

$$\begin{aligned} x \cdot m_{\frac{3}{2}} \text{mod } n &= m_R, \\ \text{for} \qquad & R = 3/2\,(x=1), \\ & R = 4/3\,(x=-1), \\ & R = 5/4\,(x=4), \\ & R = 8/5\,(x=-4), \\ & R = 5/3\,(x=3), \\ & R = 6/5\,(x=-3). \end{aligned} \tag{3.34}$$

Combining these equations with 3.33, results in

$$x_1 \cdot m_{\frac{3}{2}} + x_2 \cdot m_{\frac{3}{2}} = x_3 \cdot m_{\frac{3}{2}} \tag{3.35}$$

for combinations of x_1, x_2 and x_3 from 3.34 satisfying

$$x_1 + x_2 = x_3. \tag{3.36}$$

This means that the example of adding major and minor thirds (3.30,3.31) is always correct for all n from (3.29). The complete list of equations where the addition of intervals is correctly represented is

$$\begin{aligned} 5/4 &\cdot 6/5 = 3/2 \\ 6/5 &\cdot 4/3 = 8/5 \\ 5/4 &\cdot 4/3 = 5/3 \\ (5/3 &\cdot 3/2)\,/\,2^k = 5/4 \\ (5/3 &\cdot 8/5)\,/\,2^k = 4/3 \\ (3/2 &\cdot 8/5)\,/\,2^k = 6/5. \end{aligned} \tag{3.37}$$

These equations represent the addition of all consonant intervals according to music theory.

We are now also able to say something about the monotonicity of the projection from the frequency ratios to the pitch numbers when the enharmonicity conditions are satisfied. In the hypothetical case that all frequency ratios are projected onto the pitch numbers that approximate them best, the projection is monotonic (i.e. the ordering of the intervals is maintained). Since the enharmonicity conditions establish that the frequency ratios 3/2, 4/3, 5/4, 6/5, 8/5, 5/3 are projected onto those pitch numbers that approximate them best, this implies that the projection is monotonic for these intervals.

Generating fifth

The enharmonicity conditions (3.23,3.24) allow for enharmonic equivalence of notes, but prohibit note names to refer to two different units in the equal tempered system. Furthermore, we have seen that the condition for correctly adding intervals is automatically satisfied for all intervals given in equation 3.34.

The conditions 3.23,3.24 do not yet make sure that all units in the n-tone ET get a note-name. If the note names are distributed by m, and n is divisible by m such that $n/m = t$, $t \in \mathbb{Z}$, only t units of the n-tone division get a note-name. See for example figure 3.7 where all notes are mapped onto the units 0, 4 and 8. In this particular case (where the tones are only generated by the major third) the 12-tone system can be reduced to a 3-tone system, without changing the deviation of the equal tempered notes to the frequency ratios from just intonation (the notes are equally well approximated as in the 12-tone system).

In our method we have been using the approximation to three intervals: $3/2, 5/4, 6/5$ (or their inverses), to distribute the note-names over the n units of the equal tempered system. If one of the numbers $m_{\frac{3}{2}}, m_{\frac{5}{4}}$ or $m_{\frac{6}{5}}$ establishes that all units get a note-name, this m is a so-called generator of n. The number m is a generator of n if

$$\mathrm{GCD}[m, n] = 1, \tag{3.38}$$

that is, the greatest common divisor (GCD) of m and n is 1. In this case, m is said to be relatively relatively prime to n.

It turns out that, if the n-tone equal tempered system does not have generators among $m_{\frac{3}{2}}, m_{\frac{5}{4}}$ or $m_{\frac{6}{5}}$, the n-tone temperament can be simplified to an n'-tone temperament, such that

$$n = k \cdot n' \qquad k \in \mathbb{N}. \tag{3.39}$$

The explanation hereof is given in appendix A.3. If $m_{\frac{3}{2}}$ is not a generator of n, but the enharmonicity conditions hold, then the notes generated by $m_{\frac{5}{4}}$ and $m_{\frac{6}{5}}$ will be located in the same positions as the same notes generated by $m_{\frac{3}{2}}$ (since the notes generated by the major and minor third are subsets of the notes generated by the fifth). So, if the enharmonicity conditions are satisfied, and $m_{\frac{3}{2}}$ is not a generator, then $m_{\frac{5}{4}}$ and $m_{\frac{6}{5}}$ (and their inverses) can also not be generators of the system. Thus, when $m_{\frac{3}{2}}$ is not a generator of the system, it can always be reduced to a simplified system following (3.39).

Since n-tone systems that can be reduced to n'-tone systems are not interesting here (because they do not approximate the consonant intervals better than the reduced systems), we want to find n-tone equal tempered systems of which $m_{\frac{3}{2}}$, the number of steps approximating the fifth, is a generator of n:

$$\mathrm{GCD}[m_{\frac{3}{2}}, n] = 1. \tag{3.40}$$

Now we are able to get to the main result. Combining condition 3.40 with the enharmonicity conditions, the following values for n result:

$$n = 5, 7, 12, 19, 26, 31, 43, 45, 50, 55, 69, 74, 81, 88 \tag{3.41}$$

Let us recall that the goodness-of-fit approach from the previous section has led to systems of size: $12, 15, 19, 27, 31, 34, 41, 46$, and 53. Combining these results with eq. 3.41, we see that divisions of the octave in $12, 19$ or 31 parts would be a good choice. Indeed, keyboard applications for these temperaments have been constructed, like for example the 19-tone[8] harmonium in 1854 (Yasser 1975) and the 31-tone organ by Fokker (1955).

Finally, let us mention another approach leading to a closely related result. Similar to us, Regener (1973) built an ET system by using the fifth as a generator. He stated that "the determining constant [of an ET] is the base-2 logarithm of the frequency ratio for the perfect fifth" (Regener 1973, p.97). Then, he imposed the following criterion to preserve the relationships of frequency ratios as far as possible: "the frequency ratios of the augmented prime shall not be less than that of the diminished second and not greater than that of the minor second" (Regener 1973, p.139). The n-tone systems with orders up to 100, satisfying this criterion, yield:

$$n = 12, 19, 31, 43, 50, 55, 67, 69, 74, 79, 81, 88, 91, 98. \tag{3.42}$$

One can see, that the possible values for n from eq. 3.41 and eq. 3.42 have a great overlap. This is understandable since our conditions 3.23,3.24 as well as the condition imposed by Regener, imply that the relationships of the frequency ratios belonging to the perfect fifth, major and minor third, and major and minor sixth, are preserved.

3.3.2 Equal tempered divisions represented in the tone space

So far, we have seen that the choice of using the Western note name system has resulted in two constraints (equations 3.23, 3.24 and 3.40) on the possible divisions

[8]Chalmers (1989) explained that standard musical notation can be easily extended to the 19-ET setting.

of the octave. To better understand what these restrictions mean for the resulting n-tone ETs, it is useful to present them visually in a tone space. In chapter 2 we have seen that tone-spaces can be represented as 2-dimensional lattices $\mathbb{Z}^2$. A tone space can be made from frequency ratios, note-names or equal tempered pitch numbers. The first two are shown in figure 3.9. Different versions of the

			216/125	27/25	27/20	27/16
		144/125	36/25	9/5	9/8	45/32
	192/125	48/25	6/5	3/2	15/8	75/64
128/125	32/25	8/5	**1**	5/4	25/16	125/64
128/75	16/15	4/3	5/3	25/24	125/96	
64/45	16/9	10/9	25/18	125/72	625/576	
32/27	40/27	50/27	125/108			

			Bbb	Db	F	A
		Ebb	Gb	Bb	D	F#
	Abb	Cb	Eb	G	B	D#
Dbb	Fb	Ab	C	E	G#	B#
Bbb	Db	F	A	C#	E#	
Gb	Bb	D	F#	A#	C##	
Eb	G	B	D#			

Figure 3.9: Projection of just intonation intervals generated by major and minor third on the note-names.

tone space for equal tempered pitch numbers exist, depending on the number of parts the octave is divided into. How is such a tone space constructed? Ideally, in a mapping from a pitch number tone space to the frequency ratio tone space, *every* ratio corresponds to that pitch number that approximate the ratio best. However, if that is achieved, a specific interval can have different sizes at different places in the pitch number tone space. Also, the pitch number tone space is then not a homomorphic projection from the tone space of frequency ratios and note names anymore.

As we have seen, the approximation of the fifth $m_{\frac{3}{2}}$ should be a generator of the n-tone system, and therefore it generates all the note names. Thus, the value of $m_{\frac{3}{2}}$ defines the how the pitch numbers are projected onto the note name space of figure 3.9. We can understand this as follows. The pitch number projected onto the G on lattice point $(1,1)$ is value of $m_{\frac{3}{2}}$ itself. Then the line of fifths (the diagonal in the tone space) can be represented in pitch numbers by adding the value of $m_{\frac{3}{2}}$ modulo n. Then the value of E (the major third on C) is automatically determined by adding 4 fifths on C, and thus the pitch numbers of two basis vectors $(0,1)$ and $(1,1)$ are known. The remaining lattice points can be reached by vector addition. For example, in 12-tone ET, the vectors or lattice points $(0,1)$ and $(1,1)$ correspond to the pitch numbers 4 and 7. The lattice point $(2,1)$ reached by adding the vectors $(0,1)$ and $(1,1)$ corresponds thus to the pitch-number $4+7 (\mathrm{mod}\ 12) = 11$, and so on. A consequence of obtaining the pitch numbers in this way, for a general tone space of pitch numbers is, for pitch-numbers other than $m_{\frac{3}{2}}$, one does not know whether they form the right approximations for the frequency ratios. However, in an n-tone ET where the conditions 3.23, 3.24 are satisfied, all intervals from eq. 3.34 are approximated correctly.

Here, we will study some tone spaces of pitch numbers for an n-tone equal tempered system with n from eq. 3.41. Figures 3.10 and 3.11 display the tone spaces of pitch numbers for $n = 12$, $n = 19$, $n = 26$ and $n = 31$. The tone-space

12–tone ET

1	5	9	1	5	9	1
10	2	6	10	2	6	10
7	11	**3**	**7**	**11**	3	7
4	8	**0**	**4**	**8**	0	4
1	5	**9**	**1**	**5**	9	1
10	2	**6**	**10**	**2**	6	10
7	11	3	7	11	3	7

19–tone ET

16	3	9	15	2	8	14
11	17	4	10	16	3	9
6	12	18	5	11	17	4
1	7	13	0	6	12	18
15	2	8	14	1	7	13
10	16	3	9	15	2	8
5	11	17	4	10	16	3

Figure 3.10: Two examples of n-tone equal tempered tone spaces.

displayed on the left of figure 3.10 can be rolled up along the sides of the square (indicated by boldface numbers) to become a torus. For convenience we have placed circles around a chosen number to mark the corners of the space that represent all n pitch numbers in the chosen n-tone temperament. We see from figures 3.10 and 3.11 that the n-tone temperaments we obtained can be represented in parallelograms. These parallelograms are in fact just the periodicity blocks dis-

26–tone ET

11	19	1	9	17	25	7	15
4	12	20	2	10	18	0	8
23	5	13	21	3	11	19	1
16	24	6	14	22	4	12	20
9	17	25	7	15	23	5	13
2	10	18	0	8	16	24	6
21	3	11	19	1	9	17	25
14	22	4	12	20	2	10	18
7	15	23	5	13	21	3	11

31–tone ET

13	23	2	12	22	1	11	21	0	10	20	30
5	15	25	4	14	24	3	13	23	2	12	22
28	7	17	27	6	16	26	5	15	25	4	14
20	30	9	19	29	8	18	28	7	17	27	6
12	22	1	11	21	0	10	20	30	9	19	29
4	14	24	3	13	23	2	12	22	1	11	21
27	6	16	26	5	15	25	4	14	24	3	13
19	29	8	18	28	7	17	27	6	16	26	5
11	21	0	10	20	30	9	19	29	8	18	28

Figure 3.11: Tone space represented in a 26 and a 31 tone equal tempered system.

cussed in section 3.1.2. A geometric property of these periodicity blocks is that the number of elements n is exactly the area spanned by the parallelogram.

We notice that in each n-tone system the length of one diagonal is kept constant. Comparing this space to the note name space in figure 3.9 we see that this makes sure that all notes that have the same name are identified with each other, which is a consequence of the enharmonicity conditions[9]. The other diagonal of the parallelogram indicates which other notes are identified with each other; in other words, the notes that are enharmonically equivalent.

3.3.3 Extended note systems

Now that we have given a visual representation for the n-tone systems resulting from (3.41), we wonder how we can extend this in such a way that every frequency ratio is represented by a separate note-name. This would be very useful since then we can distinguish between two different ratios having the same note-name in the Western notational system. If this can be done, the restrictions on the equal division of the octave can probably be changed such that more divisions can be used. Eitz (1891) created a note-name system that distinguishes between different ratios from just intonation. Eitz departs from Pythagorean ratios and gives them 'normal' note-names with the superscript 0. He adds a $\pm t$ superscript to note names corresponding to ratios differing by t syntonic commas (or factor $81t/80$). See figure 3.12 for a representation. With this new note-name system,

			Bbb^{+3}	Db^{+2}	F^{+1}	A^{0}
		Ebb^{+3}	Gb^{+2}	Bb^{+1}	D^{0}	F#$^{-1}$
	Abb^{+3}	Cb^{+2}	Eb^{+1}	G^{0}	B^{-1}	D#$^{-2}$
Dbb^{+3}	Fb^{+2}	Ab^{+1}	C^{0}	E^{-1}	G#$^{-2}$	B#$^{-3}$
Bb^{+2}	Db^{+1}	F^{0}	A^{-1}	C#$^{-2}$	E#$^{-3}$	
Gb^{+1}	Bb^{0}	D^{-1}	F#$^{-2}$	A#$^{-3}$	C##$^{-4}$	
Eb^{0}	G^{-1}	B^{-2}	D#$^{-3}$			

Figure 3.12: Eitz' notation for the names of the intervals.

what conditions are preferable for constructing an n-tone ET?

In the new tone-space using Eitz notation, every frequency ratio corresponds to a unique note name, such that note name identification is not necessary anymore. However, the addition of consonant intervals may still be required. From the origin of the space (indicated by 1 in the frequency ratio space and by 0

[9]The ET systems satisfying the enharmonicity conditions can therefore be grouped under the mean-tone temperaments which are characterized by tempering the syntonic comma to unison.

in the pitch number space) the intervals major and minor third (or major and minor sixth) can be seen as the unit vectors building the space. Then, the tone space of equal tempered numbers can be constructed with the values of $m_{\frac{5}{4}}$ and $m_{\frac{6}{5}}$. Ideally, the fifth should be connected to the number that approximates this interval best. Since a major third plus a minor third should equal a perfect fifth, this is established if

$$m_{\frac{5}{4}} + m_{\frac{6}{5}} = m_{\frac{3}{2}} \tag{3.43}$$

And if the major and minor third are approximated correctly, their inverses, the minor and major sixth are too, since $m_R = n - m_{1/R}$. Furthermore, eq. 3.43 makes sure that all equations from eq. 3.37 are correctly translated in the chosen n-tone temperament.

When using the Western note name system, the generating fifth condition was introduced to establish that the resulting ETs could not be reduced to a lower number system. Using Eitz's note name system, this condition is changed slightly, which we will explain. Since there is no relation between the note names generated by the fifth and the (major or minor) third anymore, two generating elements are (minimally) required in order to attach pitch numbers to all points in the lattice. The vectors representing these generating elements should form a basis of the lattice in order to attach a pitch number to every lattice point. The generating elements can therefore be chosen from the set: perfect fifth and fourth, major and minor third, major and minor sixth. Furthermore, the choice of generating elements from this set establishes that these intervals are attached to the pitch numbers that approximate them best. If, for a certain n-tone ET that is distributed by the fifth, major third and minor third (or their inverses), not all pitch numbers are used, the system can be reduced to an n'-tone system ($n = k \cdot n', k \in \mathbb{N}$) in the same way as before. To establish that every pitch number $(1 \ldots n)$ can be constructed by a linear combination of $m_{\frac{3}{2}}$, $m_{\frac{5}{4}}$ and $m_{\frac{6}{5}}$, representing the pitch numbers that approximate the fifth, major third and minor third in the n-tone system respectively, we have to demand that

$$\mathrm{GCD}[a \cdot m_{\frac{3}{2}} + b \cdot m_{\frac{5}{4}} + c \cdot m_{\frac{6}{5}}, n] = 1, \tag{3.44}$$

where $a, b, c \in \mathbb{Z}$. Using the Euclidean algorithm (see for example Ono 1987) this condition can be rewritten as:

$$\mathrm{GCD}[\mathrm{GCD}[\mathrm{GCD}[m_{\frac{3}{2}}, m_{\frac{5}{4}}], m_{\frac{6}{5}}], n] = 1, \tag{3.45}$$

which in turn can be written shorter as:

$$\mathrm{GCD}[m_{\frac{3}{2}}, m_{\frac{5}{4}}, m_{\frac{6}{5}}, n] = 1. \tag{3.46}$$

This condition together with the adding interval condition (3.43) result in the following values for n:

$$n = 3, 4, 5, 7, 8, 9, 10, 12, 15, 16, 18, 19, 22, 23, 25, 26,$$

$$\begin{aligned}&27, 28, 29, 31, 34, 35, 37, 39, 41, 42, 43, 45, 46, 47,\\&48, 49, 50, 53, 55, 56, 58, 59, 60, 61, 63, 65, 69, 70,\\&71, 72, 73, 74, 75, 77, 78, 79, 80, 81, 83, 84, 87, 88,\\&89, 90, 91, 94, 95, 96, 97, 99, ...\end{aligned} \tag{3.47}$$

Here we give only the values for n up to 100, but there is no limit to the value of n that satisfies these conditions. Comparing these results to (3.41), we see that a lot more divisions of the octave can be used if the restriction to use Western note-names is abandoned. However, condition (3.43) was constructed especially for Western music, since it is important to have triads (the building blocks of Western music) that are in tune. So, the n-tone equal tempered systems with n from (3.47) are to be used for Western music, but music written in the normal Western note-name system has to be translated to a system like Eitz's before it is possible to play it.

3.3.4 Summary and resulting temperaments

We have argued that two mathematical conditions should be satisfied for an n-tone equal tempered system suitable for keyboard application. These conditions have led to a number of values for n. Remarkably (but intuitively understandable) there is a certain maximum to the value of n (see eq. 3.41). As a consequence, there is a maximum bound to the closeness we can approximate just intonation with equal temperament, when using the Western notational system. It is clear that good divisions of the octave to approximate just intonation (as investigated in section 3.2) that do not satisfy the two required mathematical conditions, cannot be used in combination with the Western note-name system. If one still wishes to use such tone systems this has to be in combination with an extended note-name system like Eitz's system. Combining our results for good equal divisions of the octave with the general results using a goodness of fit approach, we conclude that possible divisions for a keyboard are $12, 19$ or 31 notes per octave. Indeed, keyboard systems with these octave divisions have been constructed (Partch 1974; Fokker 1955).

A suitable equal tempered system that maps onto Eitz's extended note-name system, is still required to satisfy certain conditions. The resulting sequence of possible values of the octave division is an infinite sequence (eq. 3.47). Using the conditions that apply to the traditional note-names, we did not find the values $n = 41, 53$ that resulted from a goodness of fit approach. However, these values are obtained as part of the results using Eitz's notation (eq. 3.47). The 53-tone system which was outstanding in the goodness-of-fit approach (fig 3.4) and obtained from eq. 3.47, has been realized in the 53-tone harmonium built in the 19th century (Bosanquet 1874b; Helmholtz 1863).

Summarizing, to find suitable n-tone ETs in which Western music could be played, the first demand may be that it should approximate the ratios from just intonation well, but an additional condition should be applied if the Western traditional notational system is used. In this section, we tried to give some insight in the possibilities for using the ETs that resulted from section 3.2. Furthermore, this approach can serve as a possible explanation for the historical choices of certain tone-systems.

Chapter 4

Well-formed or geometrically good pitch structures: (star-) convexity

In the previous chapters we have focused on equal-tempered approximations of just intonation and on limitations on these systems when the Western note name system is applied. This has led to a number of possible n-tone tempered systems, some of which can indeed be found in musical literature and practice. In this chapter we will focus on theories other than approximating just intonation ratios, that may serve as a principled basis for tonal music. These theories are concerned with the notion of well-formed tone systems, and strive to show that prominent musical objects are also in prominent mathematical positions, if a suitable mathematical context is chosen. Although this chapter on well-formed scale theories is separated from the previous chapter on equal temperament, this does not mean that the theories reviewed in section 4.1 cannot be used to develop or judge equal tempered scales: some theories indeed can.

After an overview of the literature on this topic and a review of two prominent theories, we present our convexity model. Where other theories are limited in that they do not account for 5-limit just intonation, the convexity model does. Furthermore, it can be applied to more than scales, as we focus on chords and harmonic reduction as well. Moreover, the notion of convexity has applications in an intonation and modulation finding model as we will see in chapters 5 and 6.

4.1 Previous approaches to well-formed scale theory

Are there general principles that govern the "well-formedness" of tonal pitch structures? For example, when is a sequence of notes a well-formed musical scale, chord or melody? General perceptual principles for musical structuring have been proposed. The Gestalt laws (Wertheimer 1923), refer to theories of visual perception that attempt to describe how people tend to organize visual elements into groups or unified wholes when certain principles are applied. These Gestalt laws have also been applied to music (Terhardt 1987; Tenney and Polansky 1980; Leman 1997). A musical structuring based on a preference rule system has been formalized by Lerdahl and Jackendoff (1983) and Temperley (2001). Different parts of this preference rule system account among others, for the organization of grouping (phrasing) and metrical structure. Even stochastic principles have been applied to music (Bod 2002), where manually annotated folksongs were used to train and test a memory based model for phrasing. More psychologically oriented research on the goodness or well-formedness of melodies has been carried out by Povel (2002). Furthermore, attempts have been made to simulate music-theoretical prominence in terms of mathematical prominence. Mazzola (2002) has discussed the consonance/dissonance dichotomy presented as two symmetric halves of the chromatic scale according to Vogel (1975) in the 3-dimensional Euler-lattice. Noll (1995, 2001) measured the morphological richness of chords in $\mathbb{Z}_{12}$ in terms of the number of transformation classes. However, none of the theories mentioned above applies to the goodness or well-formedness of scales.

Investigations on the mathematical properties[1] of diatonic scales have been done by Clough and Myerson (1985) and Agmon (1989). In both articles the diatonic scale is presented specifically as a partial set of integer classes mod 12, (for example $0, 1, \ldots, 11$), while generically a diatonic scale is interpreted as a full set of integer classes mod 7 $(0, 1, \ldots, 6)$. This notation is independent of intonation and temperament. From the indefinitely large number of specific diatonic systems, Agmon (1989) selects the familiar diatonic system (7 scale tones embedded in a system of 12 semitones) on the basis of the best approximation of the perfect fifth, and thus presents this as an explanation for the familiar diatonic scale. The familiar diatonic scale has furthermore been presented in a mathematical model by Lindley and Turner-Smith (1993).

In this section, we will focus on two theories that belong to the category of theories that present music-theoretical prominence in terms of mathematical prominence. We will review the theories of Carey and Clampitt (1989) and Balzano (1980) in more detail, since they are concerned with the goodness of scales and they both make predictions about the number of notes into which the octave can best be divided.

4.1.1 Carey and Clampitt's well-formed scales

Carey and Clampitt (1989) developed a theory about the well-formedness of musical scales. The well-formedness of a scale is a single structural principle that underlies the pentatonic, diatonic and chromatic scales, as well as the 17-tone

[1]For a summary on mathematical properties of scales, see Clough, Engebretsen, and Kochavi (1999).

Arabic and 53-tone Chinese theoretical systems. The scales are represented by $\mathbb{Z}_n$, such as the pentatonic scale $\mathbb{Z}_5$, the diatonic scale $\mathbb{Z}_7$ and the chromatic scale $\mathbb{Z}_{12}$. In their paper, Carey and Clampitt (1989) first provide what they call an "informal" definition of a well-formed scale.

4.1.1. Definition. Scales generated by consecutive fifths in which symmetry is preserved by scale ordering are called well-formed scales.

When the sequence of fifths is represented as n points regularly spaced around a circle, the tones can be connected by fifths or by scale order. If both resulting figures represent the same degree of rotational symmetry, the n tone scale is said to be well-formed. As an example, figure 4.1 shows the well-formed 7-tone scale in consecutive fifths and scale ordering. Both figures display seven degrees of rotational symmetry. In part II of their paper, Carey and Clampitt (1989) provide a formal definition of a well-formed scale.

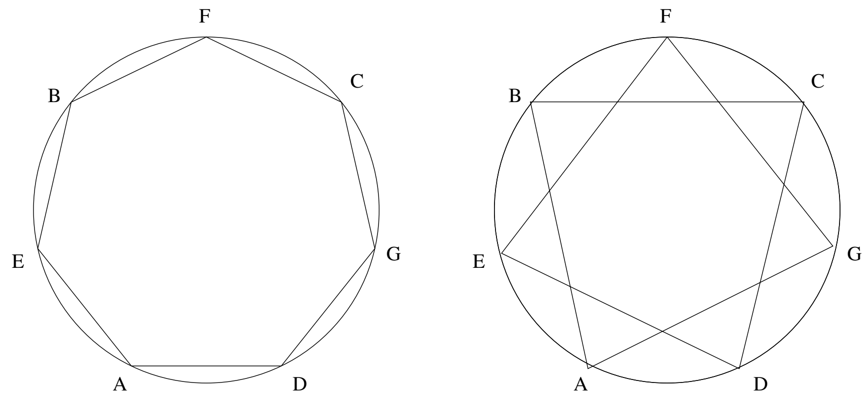


Figure 4.1: Well-formed 7-tone scale preserves symmetry by scale ordering.

4.1.2. Definition. Let $\mathbb{Z}_N = \{0, 1, \ldots, N-1\}$ represent a set of pitch classes of P produced by consecutive fifths. These pitch classes are the elements of a well-formed scale if there exists an automorphism which arranges $\mathbb{Z}_N$ in scale order.

Recall from chapter 2 that an automorphism is a special kind of isomorphism in which the set is projected onto itself. For example, there exists a permutation to rearrange the group $\mathbb{Z}_7$, which represents the diatonic set, into scale order. This automorphism of $\mathbb{Z}_7$ is given by multiplication of every element by 2 mod 7.

The theory of well-formed scales finds the sequence that starts with $n = 1, 2, 3, 5, 7, 12, \ldots$, of well-formed scales Z_n, and serves as a model explaining the existence of for example the pentatonic, diatonic and twelve tone scales. The theory is generalized by saying that the pitch classes of a scale do not need to be produced by fifths but by an interval μ. This μ is said to represent a formal fifth which may be fixed at any value where

$$2^{1/2} \leq \mu \leq 2. \tag{4.1}$$

This generalized Pythagorean system can then be represented by the set

$$P = \{2^a \mu^b | a, b \in \mathbb{Z}\}. \tag{4.2}$$

It turns out that "in a generalized Pythagorean system P, a scale with pitch classes $0, 1, \ldots, B-1$ is a well-formed scale if and only if B is the denominator in a convergent or semi-convergent A/B in the continued fraction representation of $\log_2 \mu$" (Carey and Clampitt 1989) (for continued fractions, see section 3.1.1). For example, for $\mu = 3/2$, representing Pythagorean tuning, the sequence of (semi) convergents A/B begins with $\frac{1}{1}, \frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \frac{4}{7}, \frac{7}{12}, \frac{10}{17}, \frac{17}{29}, \frac{24}{41}, \frac{31}{53}, \ldots$. The sequence $\mathbb{Z}_1, \mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_5$ etc. generically represent the well-formed scales; the specific generator is in this case $\log_2(3/2)$.

Equal tempered scales are said to be degenerate well-formed scales, because the asymmetry present in the general well-formed scale has been smoothed out in the symmetrical equal tempered scale. In well-formed scales, the intervals come in two sizes. For example, in the diatonic scale, the second, third, sixth and seventh have two sizes: major and minor; the fourth, fifth and unison have two sizes: perfect and diminished/augmented. This property has been called Myhill's property by Clough and Myerson (1985).

With their description of well-formed scale theory Carey and Clampitt provide a new approach to investigate the properties of tone systems. Furthermore the concept of a well-formed scale can serve as a principled basis for tonal music. However, the theory does only describe scales generated by μ and 2 (eq. 4.2), meaning that it could never describe a scale like for example the major diatonic scale in 5-limit just intonation[2]. Nevertheless, this well-formed scale theory gives a possible answer to the question as to what principles underlie the musical scales in the world, and is therefore important to consider here.

4.1.2 Balzano's group theoretical properties of scales

Also Balzano (1980) argued for another way of assessing the resources of a pitch system that is independent of ratio concerns. Balzano (1980) considered the individual intervals as transformations forming a mathematical group. He dealt with the group theoretical properties of 12-tone pitch systems and extended this to micro-tonal pitch systems. The octave divided into 12 semitones can be described by the set $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ (see also section 2.1.1). This set together

[2]We have seen in chapters 2 and 3 that tuning systems like 5-limit just intonation need 3 generators and is therefore referred to as a 3-dimensional system.

with the binary operation 'addition modulo 12', is a group. The identity element is 0 and the inverse of an element n is $12 - n$. This group is called C_{12} and is a cyclic group. In his paper, Balzano (1980) argues that every n-tone equal tempered system as well as every system of n ratios that can be approximated by an equal tempered system, possesses the structure of the so-called cyclic group of order n, C_n. The structure of C_{12} is examined and it turns out that this it possesses special properties. When represented in the circle of fifths, the diatonic scale is represented by a connected region (figure 4.2). When a diatonic scale is transposed a fifth up, let's say from C to G, it leads to a scale with all but one elements the same, the changed element altered by a semi-tone: $F \rightarrow F\sharp$. For more information on this specific property of the diatonic scale and its relation to well-formed scales, see Noll (2005). Every connected 7 tone set within the circle of fifths represents a diatonic scale, 12 different scales are possible.

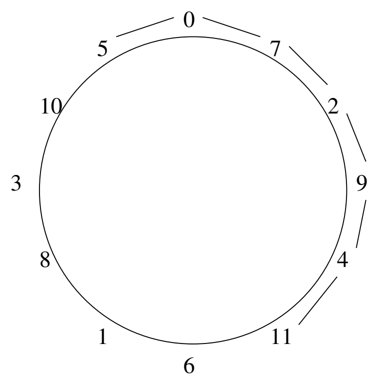


Figure 4.2: Circle of fifths with connected diatonic scale.

Balzano showed that there exists an isomorphism between C_{12} and the direct product of its subgroups $C_3 = \{0, 4, 8\}$ (augmented triad; built from major thirds) and $C_4 = \{0, 3, 6, 9\}$ (diminished seventh chord; built from minor thirds):

$$C_{12} \cong C_3 \times C_4. \tag{4.3}$$

This means there is a one to one correspondence between the elements of C_{12} and the elements of $C_3 \times C_4$. The elements of $C_3 \times C_4$ are describes as 2-tuples:

$$(a, b), \quad a \leftarrow \{0, 1, 2\}, \quad b \leftarrow \{0, 1, 2, 3\}.$$

The group $C_3 \times C_4$ has unit element $(0, 0)$, the inverse of (a, b) is $(3 - a, 4 - b)$ and the binary operation is given by:

$$(a, b) \circ (a', b') = ([a + a'] \bmod 3, [b + b'] \bmod 4).$$

The isomorphism (the mapping) between the two groups is given by

$$(a, b) \longleftrightarrow (4a + 3b)_{12}.$$

It means that every interval can be described in terms of major and minor thirds. For example, A perfect fifth can be broken down into one major third and one minor third, which is indicated by $(1, 1)$. In the same way a minor seventh can be broken down into one major third and two minor thirds $(1, 2)$. These elements of $C_3 \times C_4$ can be plotted in a two dimensional space, the minor thirds on the x-axis and the major thirds on the y-axis, see figure 4.3. Every point denoted by

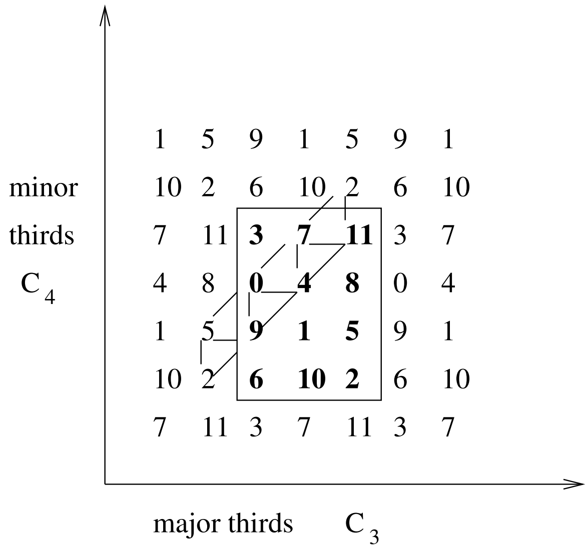


Figure 4.3: Representation of Balzano's thirds-space: $C_3 \times C_4$.

the same number n can be identified. Thus, the numbers in the square represent all 12 semitones within the octave. The numbers at the right of the square are identified with the leftmost column in the square. The row above the square is identified with the lowest row in the square. So in fact, a torus is obtained, see figure 4.4. In figure 4.3 the semitone space is represented by the line $y = -x + c$,

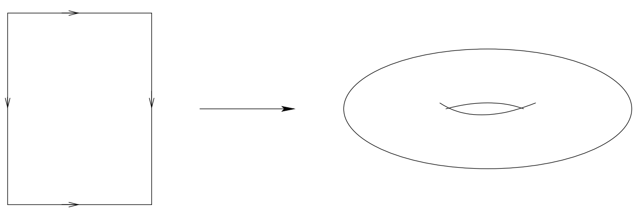


Figure 4.4: Construction of torus.

and the circle of fifths is represented by the line $y = x + c$, where c is a constant. Because the two axes in the figure consist of major and minor thirds, major and minor triads are easily identified. Every upward triangle represents a major triad, every downward triangle represents a minor triad. The full diatonic set $\{0, 2, 4, 5, 7, 9, 11\}$, indicated by the connected region contains three major and three minor triads. The diatonic scale that results is convex, compact, and spans a maximum amount of space along both axes. Therefore, the diatonic set has emerged as a unique pitch set. Generalizing this to n-fold systems, it turns out that groups that have the same structural resources as C_{12} are the ones that are of the form

$$C_n \cong C_k \times C_{k+1}, \qquad n = k(k+1), \tag{4.4}$$

for integer k. The 'diatonic' scale within this space contains $2k+1$ notes. This approach leads to octave divisions based on $20, 30, 42, \ldots$ tones. Studying these resulting tone systems, and emphasizing chordal structure, Zweifel (1996) argues that the only viable alternative to C_{12} is C_{20}. He argues furthermore, that within this 20-tone system, the eleven note scale is a better candidate for a scale than the nine note scale proposed by Balzano.

Balzano admits that the music from these alternate C_n's will probably "sound like nothing we have ever heard before". However, he claims that "the recurring triadic and diatonic set structures in changing environments will almost be surely distinguishable from random pitch changes and from C_{12} based wanderings", and "if our hearing facility cannot stretch beyond the C_{12} categories, then that is a problem of all micro-tonal systems". To question the importance of ratios in the origin of the 12-fold system, Balzano states that "it may well be that the group-theoretic properties [...] were the more perceptually important all along" (Balzano 1980).

With his group theoretic description of tone systems, Balzano gives a new and interesting view on the 12-tone system and generalizes this to n-tone systems. It may be that the group theoretic properties of the 12-fold system are at least as important as the just intonation ratios that are so well approximated by this system. However, the theory does not explain the origin of any other n tone systems than the 12-tone system. Furthermore, tone systems can only be described by cyclic groups C_n in terms of pitch numbers which means that the theory only applies to equal tempered system or tone systems that can be approximated by equal tempered systems. Yet, we will see that Balzano's approach, applied to the tone space of frequency ratios (instead of the tone space of pitch numbers) will lead to a new theory which we will now go into.

4.2 Convexity and the well-formedness of musical objects

In the current section which is based on Honingh and Bod (2004, 2005) we focus on empirical principles of "well-formedness" of a large number of tonal pitch structures, that include both 3-limit and 5-limit just intonation. The pitch structures we discuss range from ancient Greek scales to Chinese Zhou scales, and from the major triad to the eleventh and thirteenth chords. Together with other music-theoretical objects such as harmonic reductions we will see that there is a highly persistent principle holding for all these pitch structures: if represented in the tone space described in section 2.2, scales, diatonic chords and harmonic reductions form compact and convex or star-convex shapes. The convexity of harmonic reductions is due to the convexity of triads. Compactness, i.e. the extent to which elements of a set are close to the center of gravity of the set, is not a boolean valued property (i.e., it either has the property or not), but a continuum in which one object can be more compact than another. We think it is not suitable to use the term in this chapter (although Balzano 1980 used the term to describe the diatonic scale) where we consider properties of scales, chords and harmonic reductions. In chapters 5 and 6 we will come back to the notion of compactness. Instead, we focus on convexity.

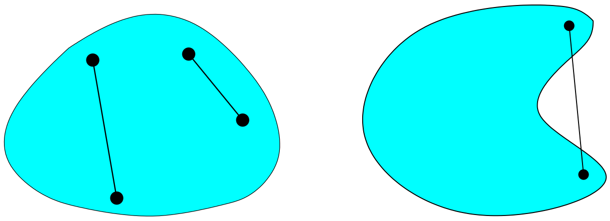
As we have seen, Balzano (1980) localized the scales in the tone space of pitch numbers. However, the pitch number tone space is not suitable for studying 3-limit and 5-limit just intonation scales, which will be our focus. Yardi and Chew (2004) have investigated the shapes of ragas from North Indian classical music in the Euler lattice. Longuet-Higgins and Steedman (1971) have noted the convexity of the major and minor scale, and used the specific form of the scales in a key finding algorithm, although the property of convexity was not necessary for the algorithm. [3] The convexity of triads in the spiral array tone space (see section 2.3) was noted earlier by Chew (2000, 2003), but the property of convexity it is not further used as a component in her theory. Moreover, neither Balzano, Longuet-Higgins and Steedman, nor Chew, have translated the property of convexity into a musical meaning.

4.2.1 Convexity on tone lattices

Convexity, as we use the term, is a notion from mathematical geometry. A set in the Euclidean space $\mathbb{R}^n$ is convex if it contains all the line segments connecting any pair of its points (see figure 4.5). Formally, a subset Y of $\mathbb{R}^n$ is said to be convex if $\alpha x + (1-\alpha)y$ is in Y whenever x and y are in Y and $0 \leq \alpha \leq 1$. Star-convexity is related to convexity. A subset X of $\mathbb{R}^n$ is star convex if there exists an $x_0 \in X$ such that the line segment from x_0 to any point in X is contained in X (see figure 4.6). A convex set is always star-convex but a star-convex object is not always convex.

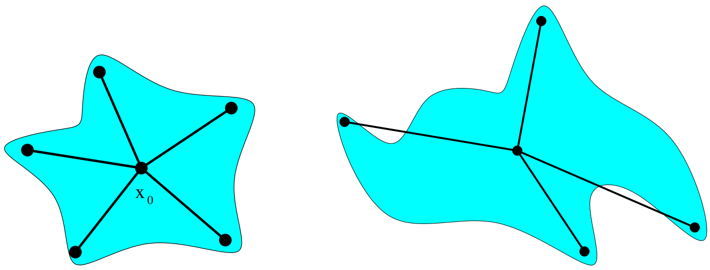
We will define a discrete convex set analogous to a convex set in continuous space, and we restrict ourselves in this thesis to discrete subsets of the lattice $\mathbb{Z}^2$. A discrete set is convex if, drawing lines between all points in the set, all elements which lie within the spanned area are elements of the set. Similarly, a discrete

[3] Convexity has also been observed in rhythm space (Desain and Honing 2003).



(a) convex (b) concave

Figure 4.5: Convex and concave set in two dimensional space.



(a) star-convex (b) not star-convex

Figure 4.6: Star-convex and non star-convex set in two dimensional space.

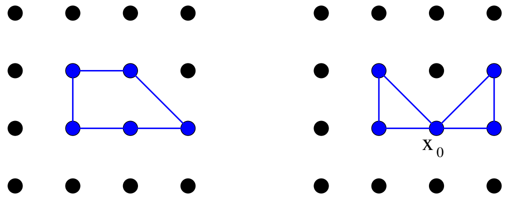
set is star-convex if there exists a point x_0 in the set such that all points lying on the line segment from x_0 to any point in the set are contained in the set (figure 4.7).[4]

These notions of convexity can be applied to music. In section 2.2 the concept of tone space was introduced. In this tone space tonal pitch structures like scales and chords can be found. For convenience, the three types of tone spaces are shown once more in figure 4.8. The definition of a discrete convex set applies to the tone space of frequency ratios[5]. In the next section, we will elaborate on the definition of convexity in the two other tone spaces.

Considering convex objects in the tone space of frequency ratios, it is important to understand how these sets transform under a basis-transformation of the

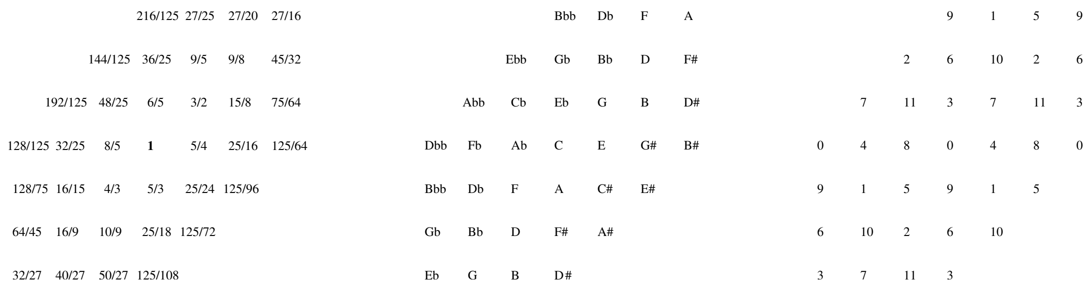
[4]The notion of a discrete line is not introduced here. We assume the lattice $\mathbb{Z}^2$ to be a subset of the continuous space $\mathbb{R}^2$ which makes the introduction of a line intuitively clear.

[5]Note that the 2-dimensional tone space of frequency ratios (fig 4.8a) is not simply equal to a plane from the 3-dimensional tone space (group P_3, that also represents the octaves), since the 2-D space was constructed by picking a representative lying within the interval $[1, 2)$ from every octave class in the 3-D space (see section 2.2 on the construction of this tone space). Therefore, a set of frequency ratios that is convex in the 2-dimensional space may be non-convex when considered in the 3-dimensional space.



(a) convex (b) star-convex

Figure 4.7: Convex and star-convex set in discrete two dimensional space.



(a) tone space of frequency ratios (b) tone space of note names (c) tone space of pitch numbers

Figure 4.8: Three representations of tone space: intervals space, note-name space, and space of pitch numbers. In figure b), the note names are chosen corresponding to the key of C.

tone space (see again section 2.2). When we choose basis-vectors for this space that are different from the major and minor third, a (star-)convex set will still remain a (star-)convex set. This can be proved as follows. Consider a linear basis-transformation T. A line between two points $\mathbf{x}$ and $\mathbf{y}$ is given by: $\alpha\mathbf{x} + (1-\alpha)\mathbf{y}$, with $\alpha \in [0, 1]$ for a continuous space, and $\alpha \in \{0, f_1, f_2, \dots, 1\}$ for a discrete space, with f_i representing the fractions of the line between $\mathbf{x}$ and $\mathbf{y}$ where other lattice points are situated. Under a transformation T it transforms into a line again:

$$\mathrm{T}(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) = \alpha\mathrm{T}(\mathbf{x}) + (1-\alpha)\mathrm{T}(\mathbf{y}). \tag{4.5}$$

Therefore, a convex set transforms under a basis transformation T into a convex set again. This property is important to ensure that convexity is a meaningful property and not just an artifact of the chosen basis. The area of these convex sets is also invariant under basis transformations since every convex set can be split into a finite number of triangles and the area of an arbitrary triangle is invariant under basis transformations, if the determinant of the transformation matrix equals 1 or -1. This can easily be verified by using the formula for the area A of a triangle given its coordinates (x_1, y_1), (x_2, y_2) and (x_3, y_3):

$$A = \frac{1}{2}\mathrm{Det}\begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{pmatrix}. \tag{4.6}$$

This formula can be worked out into:

$$A = \frac{1}{2}(x_2y_3 - x_3y_2 - x_1y_3 + x_3y_1 + x_1y_2 - x_2y_1). \tag{4.7}$$

After a basis-transformation using the transformation matrix

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}, \qquad \text{where } ad - bc = \pm 1, \tag{4.8}$$

the new coordinates x'_i and y'_i read:

$$\begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} ax_i + cy_i \\ bx_i + dy_i \end{pmatrix}. \tag{4.9}$$

The new coordinates can be filled in the formula for the area 4.6 resulting in the same value calculated in eq. 4.7, which shows that the area of a triangle is invariant under a basis-transformation with determinant 1 or -1.

4.2.2 Convex sets in note name space

In all three tone spaces from figure 2.3 pitch structures can be studied. Tonal pitch structures such as scales and chords can be expressed in terms of frequency ratios, note names or pitch numbers. The concept of convexity as explained in the previous section applies to the frequency ratio tone space. In this section we will discuss the definition of convexity in the note name space, which we will use in this chapter. In the tone space of note names, there exists more than one note named C, $D\sharp$, etc, which correspond with frequency ratios in the interval space that differ by a syntonic comma. Figure 4.9 illustrates this. When discussing a set of note names, ambiguity to as which location in the plane to consider, arises at this point, which has implications for the notion of convexity.

However, we have seen that the note names can be represented on the line $\mathbb{Z}$ (see chapter 2). In fact, the note name space as represented in figure 4.8b is isomorphic to $\mathbb{Z}$, representing the line of fifths[6] such that the ambiguity as discussed above is no longer a problem. The isomorphism with $\mathbb{Z}$ can be understood as follows. Recall from chapter 2 that other bases can be chosen for the tone space of frequency ratios. When choosing the valid basis vectors 3/2 (perfect fifth) and 81/80 (syntonic comma), the projection to the note name space looks like figure

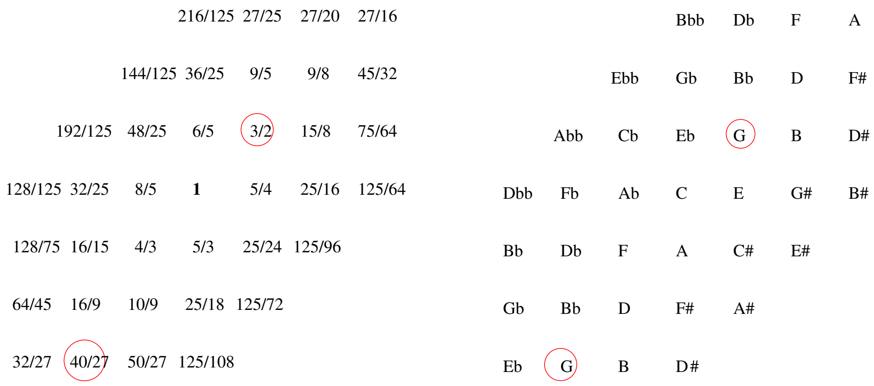


Figure 4.9: Example to show that, if all tones are labeled by their corresponding note name, multiple occurrences of the same note name appear. The corresponding interval ratios differ by a syntonic comma (81/80).

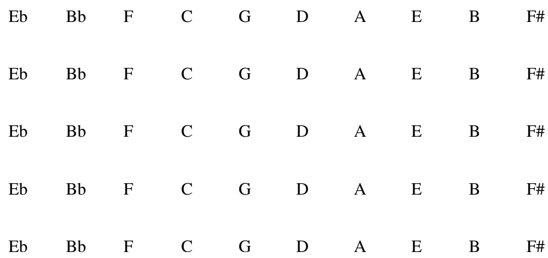


Figure 4.10: Note name space projected from frequency ratio space with basis vectors 3/2 and 81/80.

4.10. From figure 4.10 it may be clear that the note name space is isomorphic to the line of fifths:

$$\ldots - B\flat - F - C - G - D - A - E - B - F\sharp - \ldots \tag{4.10}$$

The definition of convexity as we have given it in the previous section is applicable to a discrete line $\mathbb{Z}$, so given a set of note names, it can be judged whether the set is convex or not. For example the set $C - G - D$ forms a convex set on the line of fifths. However, the set $C - E - G$ is not a convex set on the line of fifths, while in the frequency ratio space the set $1 - 5/4 - 3/2$ (representing the notes $C - E - G$) *is* a convex set. We would rather have a consistent notion of convexity through the possible tone spaces, such that a convex set of frequency ratios implies a convex set of note names, which implies in turn a convex set of pitch numbers. Therefore, we will introduce here the concept of *convex liftability*.

[6] Due to personal communication with Thomas Noll it became clear that the note name space is isomorphic to the line of fifths. He furthermore suggested to introduce the concept of convex liftability, as we will see.

Although the note name space is isomorphic to the line of fifths, we prefer to represent it here as the projection of the frequency ratio space as given in figure 4.8a. We can consider the projection of a set from the frequency ratio space to the unfolded note name space (fig. 4.8b). Then we define a set of note names to be a liftable convex set if there exists a convex set in the frequency ratio space which has these note names as a projection. Consider for example the set

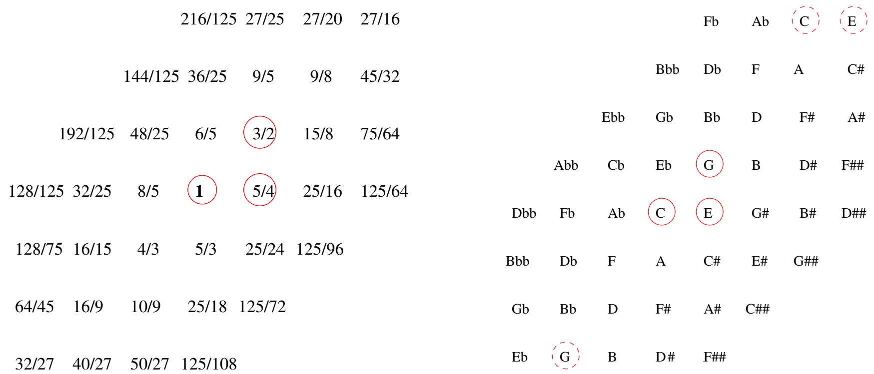


Figure 4.11: Convex projection of major triad.

C, E, G. This is a liftable convex set because it can be seen as the projection of the convex set $1, 5/4, 3/2$ in the frequency ratio space (see fig 4.11). Given a set of note names, a strategy to find the projection of a convex set (if it exists) from the frequency ratio space is to consider two points that have the same note name as two different points and to check all possibilities of compositions of a set of notes in the note name space. An example is given in figure 4.12 where a few possible compositions of the triad C, E, G are given. The first composition has a convex projection from the frequency ratio space and therefore we say that the triad C, E, G is liftable convex. To sum up: if one of the compositions of a set of

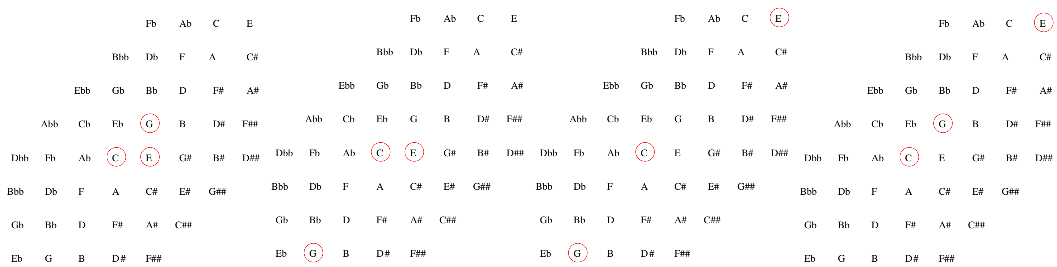


Figure 4.12: Possible compositions of the triad C, E, G.

note names is the projection of a convex set from the frequency ratio space, the set is said to be liftable convex. Since this is the notion of convexity that we will use for a set of note names in this thesis, from here we refer to convex liftability with the normal term convexity.

A similar definition can be chosen for a (liftable) convex set in the pitch number space: A set of pitch numbers is convex, if one of the compositions of a set of pitch numbers is the projection of a convex set from the frequency ratio space.

Musical interpretation of convexity

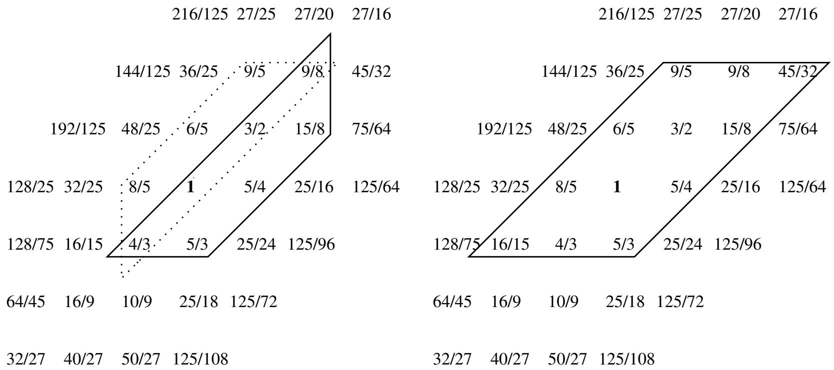
An important issue is what (star-)convexity actually means (for music). Formally, in words, convexity in the frequency ratio tone space means that all intervals lying on the line between two points are within the set. In terms of note-names it means that if two notes are in the set and the interval between these two notes can be composed from a multiple of another interval (modulo an octave), all other notes described by adding this interval (or a multiple hereof) to the lowest note, should be in the set. For example, the interval between a C and a $G\sharp$ is an augmented fifth. This interval can be composed from two major thirds. This means that if the C and $G\sharp$ are both present in a convex set, the note which is represented by a major third above the C (which is an E) should be in the set as well. From the above it appears that convexity has to do with connectivity of intervals. In a tonal context, it is natural to strive to move by the shortest number of consonances from the tonic to any other note in the scale. Therefore convexity may be a consequence of striving for maximizing connectivity, i.e. to get as many consonant intervals as possible within the notes defining the scale or chord. Star-convexity can be seen as a less strong notion: the consonance according to *one* tone is maximized.

4.2.3 Convexity of scales

Musical scales and chords can be described as sets of notes. A description of the notes in terms of frequency ratios (5-limit just intonation) provides more information than a description in terms of note names. The note names (in a specific key) can be induced from the ratios (by projecting figure 4.8a onto 4.8b), however it is not trivial to induce the frequency ratios from the note names.

For several scales, definitions in terms of 5-limit just intonation frequency ratios can be found. The major scale in 5-limit just intonation is defined as the scale in which each of the major triads I, IV and V is taken to have frequency ratios $4:5:6$ (see table 1.2). From this it can be calculated that the ratios of the scale are given by: $1/1, 9/8, 5/4, 4/3, 3/2, 5/3, 15/8$. These are indicated in figure 4.13a with solid lines.

In this figure, the (neutral) minor scale is indicated with dotted lines. These scales both turn out to form convex regions. The chromatic scale, as defined by



(a) major scale (solid line) and minor scale (dashed line)

(b) chromatic scale as defined by Vogel (1975)

Figure 4.13: Tone space of frequency ratios representing some Western scales.

Vogel (1975) and found in most textbooks[7], can also be found as a convex set in the tone space (fig 4.13b).

This result triggered us to investigate more 5-limit just intonation scales and to check whether they form a convex set in the frequency ratio space as well. From the Scala Home Page (`http://www.xs4all.nl/~huygensf/scala/`) a large collection of over 3000 scale files is available for downloads. From this collection we chose the 5-limit just intonation scales and determined whether they are convex in the tone space by plotting the scales onto the lattice. The 5-limit just intonation scales include the 3-limit just intonation scales. The other scales in the database include some 7-limit just intonation scales, equal tempered scales, mean-tone scales, and many more types of scales. The scales that are not in just intonation are represented in cents and are therefore difficult to represent in our tone space[8]. For the scales in n-limit just intonation with $n > 5$ we would need to expand the tone space to more dimensions. For example, for 7 limit, we need one extra dimension. However, relatively few scales are defined in 7 (or a higher prime) limit just intonation, and thus we have concentrated here on the scales including 3 and 5 limit just intonation. The scales that are investigated are listed in table 4.1 and besides the number of notes it is indicated whether these scales are convex and star-convex respectively.

Notice that all scales but four are convex. Differentiating between 'original' scales like the 'Ancient Greek Aeolic' and the 'Indian shruti scale', and 'con-

[7]In other definitions of the chromatic scale the minor seventh is sometimes defined as 16/9 instead of 9/5. In both cases the resulting scale forms a convex set.

[8]The equal tempered scales could be represented in a tone space in the way we did in chapter 3. The other scales (the ones that are neither in just intonation or equal temperament) may be represented in the tone space if they could be approximated by a just intonation or equal tempered scale. Furthermore, temperaments such as for example mean-tone temperament may be represented by introducing rational coefficients in the tone space, see Mazzola (1990).

name	description	No. of notes	convex	star-convex
aeolic.scl	Ancient Greek Aeolic	7	yes	yes
chin_5.scl	Chinese pentatonic from Zhou period	5	yes	yes
cifariello.scl	F. Cifariello Ciardi, ICMC 86 Proc. 15-tone 5-limit tuning	15	yes	yes
cluster.scl	13-tone 5-limit Tritriadic Cluster	13	yes	yes
cons_5.scl	Set of consonant 5-limit intervals within the octave	8	yes	yes
coul_13.scl	Symmetrical 13-tone 5-limit just system	13	no	yes
coul_27.scl	Symmetrical 27-tone 5-limit just system	27	yes	yes
danielou5_53.scl	Danilou's Harmonic Division in 5-limit, symmetrized	53	no	yes
darreg.scl	set of 19 ratios in 5-limit JI is for his megalyra family	19	no	yes
fokker-h.scl	Fokker-H 5-limit per.bl. synt.comma small & diesis, KNAW B71, 1968	19	yes	yes
fokker-k.scl	Fokker-K 5-limit per.bl. of 225/224 & 81/80 & 10976/10935, KNAW B71, 1968	19	yes	yes
harrison_5.scl	From Lou Harrison, a pelog style pentatonic	5	yes	yes
harrison_min.scl	From Lou Harrison, a symmetrical pentatonic with minor thirds	5	yes	yes
hirajoshi2.scl	Japanese pentatonic koto scale	5	yes	yes
indian_12.scl	North Indian Gamut, modern Hindustani gamut out of 22 or more shrutis	12	yes	yes
indian.scl	Indian shruti scale	22	yes	yes
ionic.scl	Ancient Greek Ionic	7	yes	yes
ji_13.scl	5-limit 12-tone symmetrical scale with two tritones	13	yes	yes
ji_19.scl	5-limit 19-tone scale	19	yes	yes
ji_22.scl	5-limit 22-tone scale	22	yes	yes
ji_31b.scl	A just 5-limit 31-tone scale	31	yes	yes
johnston_81.scl	Johnston 81-note 5-limit scale of Sonata for Microtonal Piano	81	no	yes
kayolonian.scl	19-tone 5-limit scale of the Kayenian Imperium on Kayolonia (reeks van Sjauriek)	19	yes	yes
kring1.scl	Double-tie circular mirroring of 4:5:6 and Partch's 5-limit tonality Diamond	7	yes	yes
lumma5.scl	Carl Lumma's 5-limit version of lumma7, also Fokker 12-tone just	12	yes	yes
mandelbaum5.scl	Mandelbaum's 5-limit 19-tone scale	19	yes	yes
monzo-sym-5.scl	Monzo symmetrical system: 5-limit	13	yes	yes
pipedum_15.scl	126/125, 128/125 and 875/864, 5-limit, Paul Erlich, 2001	15	yes	yes
turkish.scl	Turkish, 5-limit from Palmer on a Turkish music record, harmonic minor inverse	7	yes	yes
wilson5.scl	Wilson's 22-tone 5-limit scale	22	yes	yes
wilson_17.scl	Wilson's 17-tone 5-limit scale	17	yes	yes

Table 4.1: List of 5-limit just intonation scales from Scala archive

structed' scales we observe that all original scales are convex. Moreover, all scales are star-convex. The two scales with the highest number of notes (53 and

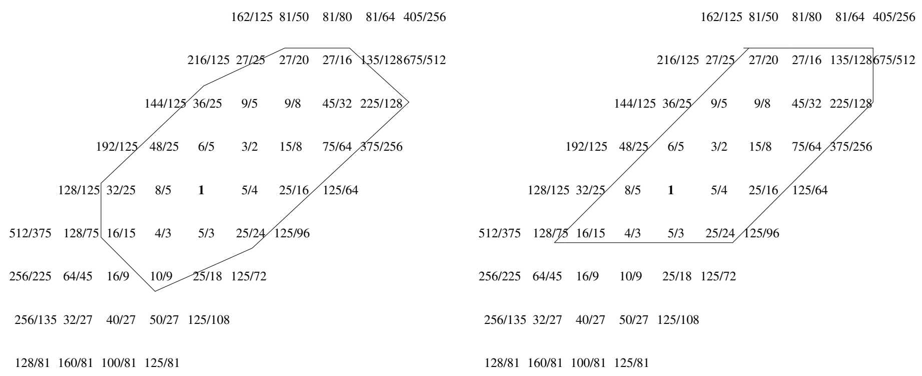
scale	convex	star-convex
Ramis' Monochord	yes	yes
Erlangen Monochord	yes	yes
Erlangen Monochord revised	yes	yes
Fogliano's Monochord no. 1	yes	yes
Fogliano's Monochord no. 2	yes	yes
Agricola's Monochord	yes	yes
De Caus's Monochord	yes	yes
Kepler's Monochord no. 1	yes	yes
Kepler's Monochord no. 2	yes	yes
Mersenne's Spinet Tuning no.1	yes	yes
Mersenne's Spinet Tuning no.2	no	yes
Mersenne's Lute Tuning no.1	no	yes
Mersenne's Lute Tuning no.2	yes	yes
Marpurg's Monochord no.1	yes	yes
Marpurg's Monochord no.3	no	yes
Marpurg's Monochord no.4	yes	yes
Malcolm's Monochord	yes	yes
Euler's Monochord	yes	yes
Montvallon's Monochord	yes	yes
Romieu's Monochord	yes	yes
Kinberger I	yes	yes
Rousseau's Monochord	yes	yes

Table 4.2: List of 5-limit just intonation scales from Barbour.

81) are not convex (but still star convex). Several of the scales from table 4.1 are symmetric around the prime interval 1/1, meaning that both an interval and its inverse are present in the scale. Two scales from table 4.1 are represented in figure 4.14 to give an idea of the typical shape of these scales. The scales that are not convex still have a similar shape, that is, a coherent object shaped around the diagonal (from bottom left to top right) of the lattice. The non-convex scales have only a few intervals not belonging to the scale which make the scale non-convex.

In his book, Barbour (1951) gives several examples of 5-limit just intonation 12-note systems. They are listed in table 4.2. Among the 26 scales, 23 are convex. Again, all of them are star-convex. The Scala archive together with Barbour's book give a balanced distribution of both traditionally and 'recently' constructed scales and therefore provide a versatile and well-documented test set.

As explained in section 4.2.2 convexity may be related to consonance. It can be understood that there is only a limited number of consonances and there are far more dissonances among all possible intervals. In creating a scale one usually aims to allow for using consonant intervals. All scales that are considered are star-convex. An interesting property of star-convexity is that the note representing the x_0 (see definition of star-convexity in section 4.2.1), is the note to which the consonance in the set is optimized. Of the scales that are star-convex but not convex, the note representing the x_0 represents the tonic. This means that in



(a) Wilson's 22-tone scale (b) Just intonation 19-tone scale

Figure 4.14: Two examples of convex scales.

these scales the consonance is optimized according to the tonic of the scale.

Thus, (star)-convexity seems to be a highly persistent property for scales, and we conjecture that it may even serve as a condition for the well-formedness of scales. In section 4.2.6 we will show that it is a non-trivial property for an item to form a convex set.

4.2.4 Convexity of chords

Now that we have investigated a number of scales in the tone space, we will look at smaller tonal pitch structures, like chords. In the area of Neo-Riemannian theory, chords in the 'Tonnetz' (a space closely related to our tone space, see section 2.3) have been studied (see for example Cohn 1998), although not with respect to the property of convexity.

A chord is a set of notes, and usually defined as a set of note names. The reason for this is that it is difficult to say something *a priori* about the intonation of the chords as there is no established theory about the intonation of all chords (see chapter 5). Therefore it is not possible to study chords in the frequency ratio tone space. We will look at chords in the note name tone space and consider convexity of the chords as explained in section 4.2.2. Contrary to the scales, we will here study chords from the Western diatonic scales only, as the note name system we use is a consequence of the Western music tradition. It is possible however, to study non-Western chords, provided that the elements hereof are given in frequency ratios or a format (like note names or pitch numbers) that can be represented in a tone space which is a projection of the tone space of frequency ratios.

Considering different kinds of chords, a distinction can be made between chords that are built from diatonic notes, which are notes that are present in the scale of the specific key, and chords that contain non-harmonic notes, the

harmonic chords	No. of notes	convex	star-convex
major triad	3	yes	yes
minor triad	3	yes	yes
diminished triad	3	yes	yes
augmented triad	3	yes	yes
dominant seventh chord	4	yes	yes
major seventh chord	4	yes	yes
minor seventh chord	4	yes	yes
half-diminished seventh chord	4	yes	yes
major-minor seventh chord	4	yes	yes
augmented seventh chord	4	yes	yes
diminished seventh chord	4	yes	yes
triad with added sixth	4	yes	yes
complete dominant ninth chord	5	yes	yes
tonic/dominant eleventh chord	6	yes	yes
tonic/dominant thirteenth chord	7	yes	yes

Table 4.3: Chords built from harmonic notes.

altered chords	No. of notes	convex	star-convex
non dominant diminished seventh chord	4	yes	yes
Neapolitan sixth	3	yes	yes
augmented sixth (Italian)	3	no	yes
augmented six-five-three (German)	4	no	yes
augmented six-four-three (French)	4	no	no
doubly augmented fourth	4	no	yes
chords with raised fifth - major	3	yes	yes
- minor	3	yes	yes
- with minor seventh	4	no	yes
dominant chord with lowered fifth	3	no	yes
-with seventh	4	no	no
dominant chord with lowered and raised fifth	4	no	yes

Table 4.4: Chords containing non-harmonic notes.

so-called altered chords. It turns out that all chords built from harmonic notes discussed by Piston and DeVoto (1989) are convex (and therefore also star-convex) in the note-name space. The chords are listed in table 4.3.

Altered chords are difficult to study since it is possible, through the process of chromatic alteration, to create a very large number of altered chords. Therefore, in this paper we reduce the number of these chords to the ones discussed by Piston and DeVoto (1989). In table 4.4 these altered chords are listed and it is indicated whether these chords are (star-)convex.

Remarkably, most altered chords are not convex, and some altered chords are not even star-convex. For most chords it can be checked immediately in figure 4.8 whether they are convex or not. For one chord, it is however not that obvious. The minor chord with raised fifth can be represented in note names as $C, E\flat, G\sharp$. On the tone space, it is most logically seen as the projection from the frequency ratios $1, 6/5, 25/26$, not forming a convex set. However, if it is seen as the projection from the ratios $1, 6/5, 125/81$, it does form a convex set. Since the $1, 6/5, 25/26$ set is a much more compact configuration, it may be strange to classify this chord as convex. In chapter 5 we will elaborate on this.

Precisely two chords are not star-convex, the French augmented sixth chord and the dominant chord with lowered fifth and minor seventh. These chords in fact consist of the same notes (but have a different function in harmony) and therefore describe the same shape in the tone space. Thus convexity roughly distinguishes between harmonic and altered chords, where the harmonic chords are all convex and the most altered ones are not convex and sometimes not even star-convex. Another surprising fact is, that of the star-convex (altered) chords the x_0 (see definition in section 4.2.1) does not represent the root of the chord in most cases[9]. In view of the meaning of star-convexity: the consonance is optimized according to one interval, this interval is not the root of the chord. Interpreting the property of convexity in terms of consonance, the fact that many altered chords are not convex can be related to the tension in the tension-resolution effect that is often found in music. Again, it is a non-trivial property for chords to be convex, as we will analyze in the discussion (section 4.2.6).

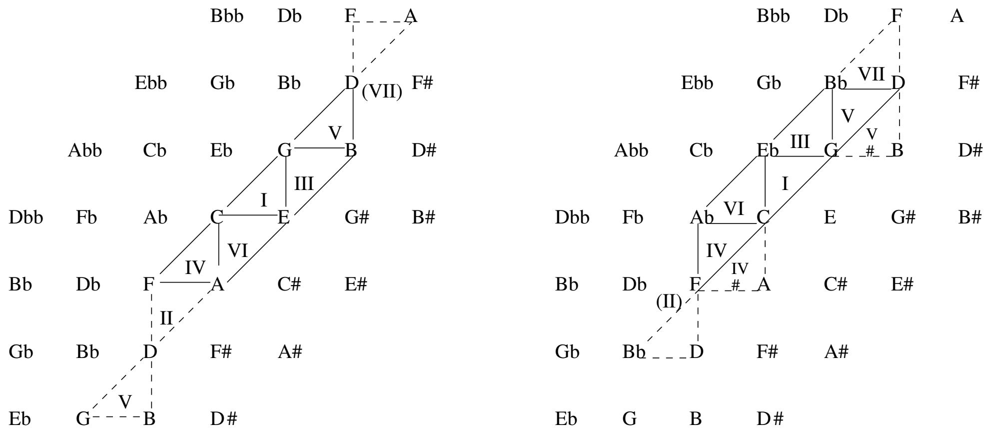
4.2.5 Convexity of harmonic reduction

Harmonic reductions of music are known to be useful for discovering the harmonic structure of a piece allowing for an easier analysis. In this process a score can be reduced to chords and ultimately to triads (Schenker 1906; Salzer 1962). Important theories of chord progressions include *Traité de l'harmonie* by Rameau (1722) and Hugo Riemann's (1914) theory of Tone images (Tonvorstellungen). Rameau postulated that harmonic progression is governed by the fifth and thirds connections of roots of triads. Riemann analyzed harmonic progressions in terms of chains of triads through his tone net (which is equivalent under a basis transformation to our tone space). This illustrates that the reduction of music into triads has a long tradition in analyzing Western tonal music.

We will now investigate the progression of triads in our tone space[10]. It is difficult to decide which tone space we should use, the frequency ratio tone space or the note name space. The major and minor diatonic scales are defined in frequency ratios which means that we could use the former space, however the tuning of the super-tonic triad (in C: $D - F - A$) is not unambiguous when preceded or followed by another chord (see section 1.3.1 about pitch drift) and

[9]However, the roots of altered chords are often ambiguous.

[10]Since we are using octave equivalence we are not taking into account the different inversions of a chord. Schenker (1906) and Salzer (1962) state that it depends on the bass whether a chord progression is a harmonic progression. Here we treat all triads as having an harmonic function.



(a) major scale (b) minor scale

Figure 4.15: (a) Triads from the major scale situated in the note-name space, VII is the diminished triad $B - D - F$. (b) Triads from the minor scale, II is the diminished triad $D - F - A\flat$.

therefore the note name space might be preferable to use. To deal with these problems the note name space is shown in figure 4.15a, where the major diatonic scale is projected from the frequency ratio space. The two possible super-tonic triads are indicated with dashed lines so that either one can be used. To summarize, the convexity definition in the frequency ratio space is used, but the scale is displayed in the note name space for convenience.

We saw in section 4.2.4 that all triads are convex. In the major scale only major and minor triads and one diminished triad are naturally present. It can be checked from figure 4.15a that every two triads following each other except for the progression $IV - V$ have the possibility to form a convex set (for sequences involving the super-tonic triad, using either the one or the other possibility, keeping notes that appear in adjacent triads in the same intonation.). The sequence $IV - V$ does not form a convex set (however it does form a star-convex set), we will see in chapter 5 that in this case convexity is overruled by compactness. Since the major scale consists of only 7 triads (8 if the two possibilities for the super-tonic triad is counted as 2), sequences of more than two triads have an even bigger chance of being convex. Apart from a sequence consisting (only) of the triads IV and V combined with II and/or VII, all sequences are convex. In practice, the sequence $IV - V$ is usually combined with the tonic triad I, resulting in a convex structure.

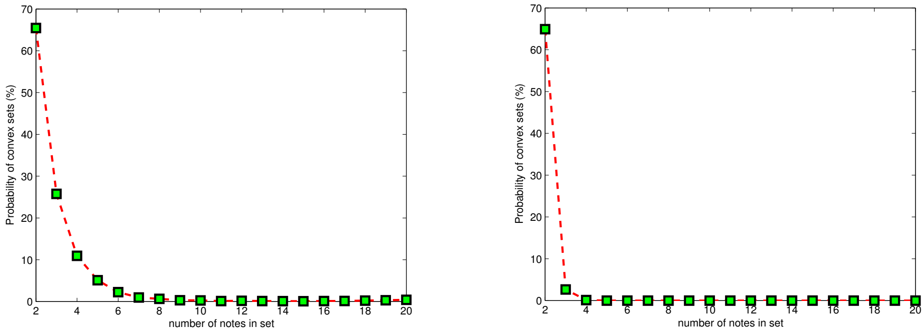
Thus, whatever segmentation of music in a major key used (segmentation per chord, bar, phrase, etc), the reduction of the music to triads most of the time represents a convex set. For music in a minor key the situation is somewhat more complicated. From figure 4.15b one can see that the neutral minor and the

Chord	is followed by	sometimes by
I	IV, V	VI
II	V	IV, VI
III	VI	IV
III (ma)	VII	
IV	V	I, II
V	I	IV, VI
VI	II, V	III, IV
VII (ma)	III	
VII (dim)	I	

Table 4.5: Chord progression in minor mode from Piston and DeVoto (1989); 'ma' indicates major chord, 'dim' indicates diminished chord.

harmonic minor scale form convex regions in the tone space. In the harmonic minor scale the seventh note of the scale is raised by a half tone, such that the III, V and VII triads are changed. In the ascending melodic minor scale the sixth and the seventh tone of the scale are raised. Consequently the triads on II, IV and VI are adjusted as well. Therefore, music in a minor key and the harmonic reduction thereof takes into account many more triads than music in a major key. Considering the harmonic progressions for the minor mode as given by Piston and DeVoto (1989) (see table 4.5), all progressions[11] form convex sets except for the progressions $IV - V$ as in the major scale, and the progression VII(diminished)-I. The VII(diminished)-I progression is not a convex set (but it does form a star-convex set) if the triads VII(dim) and I are chosen such that the triads themselves are convex. But the notes of the triads (in C minor: $B - D - F$, $C - E\flat - G$) can also be chosen in the note name space, such that this progression does form a convex set (choosing the F as a perfect fifth under the C). The way the location of the notes is chosen depends on the intonation of the notes as the mapping in figure 2.3 indicates. In just intonation, intervals are tuned to simple number ratios. The fact that intervals can be harmonic as well as melodic can cause some problems. It is not always possible to ensure that two adjacent chords are tuned to lowest number ratios in harmonic as well as melodic form. In the case of the VII(dim)-I progression, we could argue that the melodic just intonation overrules the harmonic intonation of the VII chord, and that therefore the F of the VII chord should not be tuned as 27/20 but as 4/3. In that case the $VII - I$ progression does form a convex set. Since the overall shape of notes present in the (melodic) minor scale does not represent a convex structure unlike the major scale (but does represent a star-convex structure) it is difficult to generalize the convexity of two note sequences to all possible sequences in a minor key. However, above we argued that most important progressions in the minor

[11] We have left out the progressions involving III(harmonic) and VI(ascending melodic), since these are rarely used according to Piston and DeVoto (1989).



(a) 5 × 5 lattice (b) 15 × 15 lattice

Figure 4.16: Probability of convexity for n element sets.

mode form convex regions, which means that most segmentations constitute a convex shape.

When music is harmonically analyzed, harmonic functions are assigned to groups of notes in accordance with people's perception. The music can be reduced to the triads corresponding to the harmonic functions. Therefore, it is believed that reduction represents the perception of a piece of music. Since we just argued that the reduction of a piece of music constitutes a convex body, we conjecture that this notion of convexity contributes to the perception of harmonic functions.

4.2.6 Discussion

A question that may arise is: how special is it for a set to be convex? One could think that the chance to obtain a convex set from randomly chosen points in our lattice is higher than average, and that convexity is therefore an artifact of the space we use. To check this, we wrote a program in Matlab which calculates the chance that a randomly chosen set of points is convex. We started with a 5 × 5 lattice to choose our sets from. For each number of elements n, a large number of randomly selected sets was chosen, and for each set it was calculated whether it was convex or not. One point is always chosen in the center of the lattice, since that is our reference point (1 in the ratio-space). Figure 4.16a shows that the probability of convex sets is a monotonically decreasing function.

To simulate a more realistic situation, we created another figure but now choosing sets from a 15 × 15 lattice. This seemed big enough (the real 'tone space' lattice is infinitely big) to cover all scales. From figure 4.16b we see that the percentage of convex 2-note sets is high (65%). This means that if one randomly chooses one note (the other is fixed in the center) on the lattice, the chance to obtain a convex set is 65%. For 3-note sets this percentage is around 5%, and for more note sets, the chance of choosing randomly a convex set is negligible.

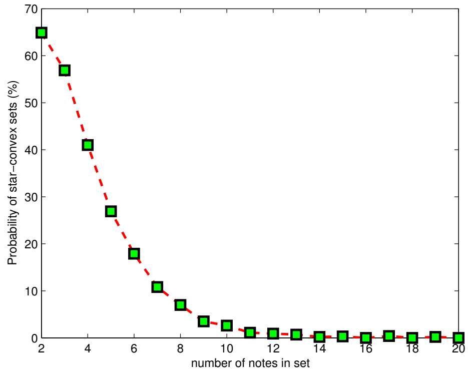


Figure 4.17: Probability of star-convexity for n element sets in a 15×15 lattice.

We also wrote a Matlab program to calculate the chance that a randomly chosen set is star-convex. The result is given in figure 4.17. Again, this is a monotonically decreasing function. The chances that randomly chosen sets appear to be star-convex are somewhat higher than in the case of convexity (which is what we expected), but still the probability to obtain a star-convex set consisting of seven notes or more is less than 20%, and for twelve notes or more the chances to get a star-convex set are negligible.

A second discussion point is the discrete lattice we used. One may wonder whether it is more convincing to study convexity in a continuous space instead of in a discrete space. However, such a study is not possible for our purposes. The tone space is built from points described by $\{2^p(\frac{5}{4})^q(\frac{6}{5})^r | p, q, r \in \mathbb{Z}\}$. In terms of coordinates, $(0, 0)$ indicates the frequency ratio 1, and $(1, 0)$ represents the frequency ratio $5/4$. Between these two there is, among other, the ratio of $6/5$ $(1 < 6/5 < 5/4)$ but this ratio can be found at the point with coordinates $(1, 0)$, and thus not between $(0, 0)$ and $(0, 1)$. Therefore, this lattice, although it is infinite in both directions, cannot be made into a continuous space.

Instead of checking whether a scale or chord is convex or star-convex one might propose to measure the *degree* of convexity. That is, a round object can be understood to be more convex than a stretched oval object. But, due to the different bases that are possible for the tone space, objects may change in their form. Having proved in (4.5) that a convex body is still convex in another basis, the degree of convexity can change. Therefore the distinction between convex and star-convex can be made, but a further division is impossible in this space. Still, a measure of convexity is possible in a different way. In table 4.1, the number of notes of every scale is also indicated. The more notes, the more possibilities of arranging these notes in the plane, the more special it is if these notes do form a convex set. Furthermore, a degree of compactness is possible, and we will focus on this in later chapters.

Number of tested pitch structures	percentage convex	percentage star-convex
53 scales	86.8%	100%
15 harmonic chords	100%	100%
12 altered chords	33.3%	83.3%

Table 4.6: Summary of results.

4.3 Concluding remarks on well-formedness

In the previous section we investigated the property of (star-)convexity as a general principle or condition for the well-formedness of tonal pitch structures. We noted that several pitch sets are (star-)convex. We discovered that all 5-limit just intonation scales and all chords built from harmonic notes are either convex or star-convex. Our results are summarized in table 4.6.

We saw in the discussion that it is highly unlikely for a randomly chosen set to be convex or star-convex. Therefore these results are far more surprising than one may think at first glance. Our results suggest an interesting hypothesis, namely that (star-)convexity serves as a condition for the 'well-formedness' of tonal pitch structures. Star-convexity is a less strict notion than convexity, but it is intriguing that nearly all the pitch structures discussed here follow this property. These two notions circumscribe a certain space of good chords and scales. We discussed the meaning of convexity and saw that it may be a consequence of maximizing the consonances in a musical scale or chord.

We have also shown that the tonal coherence which forms the harmonic reduction of a piece is usually a convex body. From this, we hypothesize that the notion of convexity may contribute to the perception of harmonic functions.

The convexity property can perhaps be best compared to the 'Good form' principle from Gestalt theory. This principle applies to visual cognition which prefers to group shapes that are symmetrical, completed, made of clean contours, and the like. Our results suggest that this principle can now also be applied to musical cognition by stating that musical objects prefer an intervallic structure in which consonance is optimized. We conjecture that the Good form principle can be formalized by means of convexity.

The convexity model differs from the well-formed scale theory of Carey and Clampitt (1989) and the group theoretical description of tone systems of Balzano (1980) in the sense that it specifically accounts for just intonation. All three theories can serve as a (partial) explanation of the origin of certain tone systems, including the Western 12-tone system with 7-tone diatonic scale. Contrary to the other two theories, the convexity model does not immediately predict tone systems of a specific size. Although we found that most tone systems form a convex set in the tone space, not every convex set in the tone space presents a possible tone system. Therefore, convexity might represent a necessary condition for a well-formed tone system, but not a sufficient condition. Still, convexity can help to make predictions about suitable n-tone systems when this condition is combined with another. Unfortunately, it is not easy to combine the three discussed theories. The theory of Carey and Clampitt (1989) only accounts for scales generated by the interval μ (and octave 2), which means that these scales form a trivial convex one-dimensional set if represented in a tone space[12]. The diatonic scales of $2k + 1$ elements embedded in the space of $k(k + 1)$ elements as predicted by Balzano are convex scales since that was one of the properties of the 12-tone scale from which the theory was generalized. In this perspective, the convexity condition supports the resulting n-tone systems suggested by Balzano (1980). Balzano's and Carey and Clampitt's predictions for n-tone systems have, besides $n = 12$ (which was the point of departure for Balzano), no other common values for n.

In this chapter, only scales in 5-limit just intonation are investigated with respect to the property of convexity. However, using the definition of convexity on the note name tone space from section 4.2.2 (which also applies to the pitch number tone space, since, similar to the note names, more than one pitch number corresponds to one frequency ratio) convexity of scales can also be studied in the tone spaces of note names and pitch numbers (the latter space having various appearances depending on the value for n in an n-tone ET as illustrated in chapter 3). Therefore, also equal tempered scales can be studied with respect to the property of convexity. However, the chance to obtain a convex set by randomly choosing points from an equal tempered pitch number space (with a finite number of pitches) is higher than doing this in the infinite frequency ratio space. Therefore, the property of convexity in the pitch number space is somewhat less interesting than in the frequency ratio space. Moreover, scales that are constructed from all the notes of the equal tempered system in which they are embedded (like the chromatic 12-tone scale in ET), form necessarily a convex region since they contain the whole (toroidal) pitch number space.

[12]The property of convexity *can* however be compared to the property of pair-wise well-formedness. A scale is pair-wise well-formed if, when any pair of step intervals is equivalenced, the resulting pattern is a well-formed scale (Clampitt 1997). Of all the convex scales discussed in this chapter, we have checked whether they are pairwise well-formed. It turned out that the majority of the convex scales is not pair-wise well formed. The other way around, we do not have any evidence pointing in the direction that pair wise well-formed scales induce convexity. There is nothing that makes a small step on the lattice preferable to a large step.

Chapter 5

Convexity and compactness as models for the preferred intonation of chords

In the previous chapter we have seen that a diatonic chord represents a convex and compact set in the tone space of note names. In this chapter this convexity and compactness of chords will be used as the basis of a model for the preferred intonation of chords. Parts of this chapter have been published as Honingh (2006a, 2006c). As explained in section 4.2.4, a chord is a set of notes, and usually defined as a set of note names. Therefore, in chapter 4 the convexity of chords has been considered in the tone space of note names (contrary to the space of frequency ratios). However, if one manages to make a suitable lift of set of note names to the space of frequency ratios, this set can be represented by frequency ratios as well which represents the intonation of the chord. We will present two hypotheses that deal with the preferred intonation of a chord, one that is concerned with convexity and one that is concerned with compactness. It will turn out that compactness is more indicative of consonance than convexity. We will try to relate the compactness of a set in the 2 and 3 dimensional tone space to the consonance measure proposed by Euler. Finally, in section 5.3 the compactness, convexity and consonance according to Euler, are calculated for all possible sets (chords) of 2, 3 and 4 notes within a bounded note name space, such that the relation between these three measures can be obtained.

5.1 Tuning of chords in isolation

The preferred intonation of an interval or chord in isolation (without a musical context) is usually given by the most consonant performance of the chord (see section 1.4.2 in the introduction). Many functions have been constructed to measure the consonance of an interval or chord, for example Helmholtz's (1863) roughness function, Euler's Gradus Suaventatis (Euler 1739; Fokker 1945), Parncutt's (1994) pitch distance or Sethares' (1993) dissonance curve based on Plomp and Levelt's

(1965) model. These functions can be used to put musical intervals in an order of most consonant to dissonant as we have seen in section 3.2.1. However, few of these functions have been used (and are difficult to use) to decide about different intonations of the same chord. As we have seen in section 4.2.2 different locations of a note name exist in the note name space, all giving rise to a different frequency ratio and therefore different intonation. For some chords, like for example a major triad, the intonation may be clear (see table 1.2), but for others there is no consensus. Consider for example a dominant seventh chord $C - E - G - B\flat$. It can be tuned choosing the ratios: $1, 5/4, 3/2, 9/5$ such that the minor seventh is tuned as minor third 6/5 above the fifth; or tuned as $1, 5/4, 3/2, 16/9$ such that the minor seventh is chosen to be two fourths above the tonic, and many other possibilities exist. This chapter is concerned with intonation based on frequency ratios from just intonation only. This means that the presented intonation model applies to either simple tones or tones whose pitch is determined by the frequency of the fundamental; the influence of other phenomena that apply to the pitch of a tone (as described in chapter 1) is not considered here.

Regener (1973) stated the ambiguity involving just intonation frequency ratios as follows: Each notated interval actually corresponds to an infinite number of frequency ratios, since multiplication of a frequency ratio by any integer power of 81/80 leaves the interval unchanged. Regener (1973) describes furthermore two criteria that are commonly used or assumed in determining which are the "preferential" frequency ratios in just intonation corresponding to a given interval:

1. Preferred ratios are those involving the lower numbers when in lowest terms.

2. Preferred ratios are those that can be derived by linear combination from known preferred values for other intervals (beginning with the ratios 3/2 for a perfect fifth and 5/4 for a major third), possibly with a certain use of intervals in mind from some musical context.

It may be clear that these two criteria are not always in agreement and do not constitute a full intonation theory for chords in isolation.

5.1.1 A model for intonation

In chapter 4 we have seen that the major and minor diatonic scale as well as all harmonic chords form compact and convex items in the note name space. From the projection of the frequency ratio space to the note name space it became clear that the difference between two ratios having the same note name is a factor 81/80 (or multiples hereof), which is known as the syntonic comma. It can therefore be understood that a chord defined as a set of note names has several possibilities for intonation. In terms of the tone spaces of frequency ratios and note names, the problem is to choose the right locations of the notes in the tone space in order to reflect the right frequency ratios that can be projected from one space to the other. It has been explained in chapter 4 that convexity can be interpreted in terms of consonance. Therefore, we can hypothesize that a convex set represents the preferred intonation of a chord in the frequency ratio space. As we explained in 4.2 we did not quantify compactness, but it may be useful to elaborate on it now to be able to distinguish between several convex sets. As a matter of fact, a set of notes can have more than one convex *configuration* in the note name space. By the configuration of a set, we mean the locations of all elements in this set, indicated in the tone space. The configuration of a set of note names can be changed by moving one or more elements of the set by a syntonic comma (=81/80). The ratio 81/80 is exactly the difference between two ratios with the same note name. Therefore, multiplying one of the ratios in a chord with this, gives a different tuning of the same chord. For example, the two-note set $C - G$ tuned as $1 - 3/2$ represents a convex configuration of this set. However, tuning the set as $1 - 40/27$, gives a convex configuration as well (see figure 4.9 in the previous chapter). In cases like this, a choice has to be made between the two configurations to present the preferred tuning, and a possibility is to choose the most compact one. Compactness is intuitively understood as the extent to which elements of a set are close to the center of gravity of the set. In a three dimensional space the most compact object would be shaped like a ball. In this thesis, we will define the compactness of a set as the sum of the euclidean distances between all pairs of points in the set. The decision to choose the most compact set is not a random choice. If two notes are close together in the tone space, they have many prime factors in common, as the tone space was built from powers of the primes $2, 3$ and 5. Therefore, the closer together two notes are in the tone space, the smaller are the integers forming the ratio that represents the interval between the two notes. According to just intonation, ratios with small integer ratios are preferred. Generalizing this for chords consisting of more than two notes, the intonation of a chord whose notes are the most close together in the tone space should be preferred. Unlike convexity, compactness is not independent of the lattice. If a compact round object is viewed in a lattice with other basis-vectors, the degree of compactness will change. If we want compactness of a set of notes to correspond as well as possible to the lowest powers of primes, the tone space should be constructed using the factors $2, 3$ and 5 as basis vectors. Since the tone space represents intervals modulo the octave (factors of 2) this means that we use basis vectors 3/2 and 5/4 representing the perfect fifth and the major third. The resulting tone space is shown in figure 5.1 which presents the space as described in chapters 2 and 4 under a basis-transformation.

Now we have motivated why compactness is a good indication to decide which of the possible convex sets represents the preferred intonation, we can actually make two hypotheses:

1. the preferred intonation of a chord is represented by the most compact set of the possible convex configurations of that chord.

	25/18	25/24	25/16	75/64	225/128		
40/27	10/9	5/3	5/4	15/8	45/32	135/128	
32/27	16/9	4/3	1	3/2	9/8	27/16	81/64
256/135	64/45	16/15	8/5	6/5	9/5	27/20	81/80
	256/225	128/75	32/25	48/25	36/25	27/25	

	F#	C#	G#	D#	A#		
G	D	A	E	B	F#	C#	
Eb	Bb	F	C	G	D	A	E
Cb	Gb	Db	Ab	Eb	Bb	F	C
	Ebb	Bbb	Fb	Cb	Gb	Db	

Figure 5.1: Tone space of frequency ratios constructed from projection $\phi : 2^p(\frac{3}{2})^q(\frac{5}{4})^r \rightarrow (q,r)$. The tone space of note names is obtained from the former space as explained in chapter 2.

2. the preferred intonation of a chord is represented by the most compact configuration of that chord.

Note that these hypotheses contain an empirical component, since "preferred intonation" applies to the perception of humans. However, we will follow here the path of investigating the correlation between the hypotheses and an established consonance measure. Thus, the hypotheses will be applied them to a number of chords, and the result will be compared with an existing consonance measure. The consonance measure we will use is Euler's Gradus function, since it applies, similar to the hypotheses, to frequency ratios of chords in isolation. Although nowadays Helmholtz's (1863) consonance theory which is based on the beating of partials, seems to be most supported (Terhardt 1974; Plomp and Levelt 1965; Sethares 1993; Kameoka and Kuriyagawa 1969a, 1969b), the difference between Helmholtz's and Euler's theory is small in view of our purposes in this chapter. The order from the most to least consonant interval according to Euler or Helmholtz (see section 3.2.1) may differ slightly, but the question we will address in this paper is about the most consonant frequency ratios given a chord or interval. From all possibilities, Helmholtz and Euler's theories will choose virtually always one and the same set of frequency ratios. For example, given the interval $C - E$, what is the frequency ratio that makes this interval as consonant as possible? Both Euler and Helmholtz rate 5/4 as the most preferred intonation.

Euler developed his Gradus Suaventatis (degree of softness) Γ. The function is defined as a measure of the simplicity of a number or ratio. Any positive integer a can be written as a unique product $a = p_1^{e_1} \cdot p_2^{e_2} \ldots p_n^{e_n}$ of positive integer powers e_i of primes $p_1 < p_2 < \ldots < p_n$. Euler's formula is then defined as:

$$\Gamma(a) = 1 + \sum_{k=1}^{n} e_k(p_k - 1), \tag{5.1}$$

$\Gamma(a)$ is a number that expresses the simplicity of a. The lower the number the simpler is a. For intervals and chords, a so-called *exponent* needs to be calculated to obtain Γ from. For an interval x/y the exponent is the ordinary product $x \cdot y$ so $\Gamma(x \cdot y)$ expresses the simplicity of the interval x/y. For chords where the frequency ratios are expressed as $a : b : c$, the exponent is given by the Least Common Multiple (LCM) of these a, b and c. The Gradus Suaventatis is then calculated as $\Gamma(LCM(a, b, c))$. Euler connected the simplicity of chords and intervals with the consonance thereof. This can be understood by thinking in terms of frequency periodicity. If one hears for example a tone of 300 ($= 5 \times 60$) Hz and one of 420 ($= 7 \times 60$) Hz, then per second 60 repeated patterns can be heard in which each pattern can be subdivided in $5 \times 7 = 35$ pieces. The more 'repetition' can be heard, the simpler or the more consonant is the sound, was the argument by Euler. Therefore, the lower the value $\Gamma(LCM(a, b, c))$, the more consonant is the chord $a : b : c$. Here is an example to calculate the Gradus Suaventatis. A major triad $1 : 5/4 : 3/2$ can be written as $4 : 5 : 6$ which can in turn be written as $2^2 : 5 : 2 \cdot 3$ (to make the calculation of the LCM easier). The LCM of these numbers is then $2^2 \cdot 3 \cdot 5 = 60$ and the Gradus Suaventatis of 60 is $\Gamma(60) = 1 + 2 \cdot 1 + 1 \cdot 2 + 1 \cdot 4 = 9$. According to the tonal space that maps frequency ratios to note names (see figure 5.1), this chord can also be tuned differently, for example as $1 : 5/4 : 40/27$, the fifth of the triad is then changed by the syntonic comma (81/80): $\frac{3}{2}/\frac{81}{80} = \frac{40}{27}$. The ratios $1 : 5/4 : 40/27$ can be written differently as $108 : 135 : 160 = 2^2 \cdot 3^3 : 3^3 \cdot 5 : 2^5 \cdot 5$. Then the LCM equals $2^5 \cdot 3^3 \cdot 5 = 4320$ which results in $\Gamma(4320) = 16$. This is obviously higher than the value for the $4 : 5 : 6$ chord and this means that this chord is less consonant than the $4 : 5 : 6$ chord according to this function.

5.1.2 Compositions in the tone space indicating the intonation

In the same way we can compare other chords in different tuning to see which tuning is most preferable. We compare different configurations of a chord. As we have seen, the configuration of a set of note names can be changed by moving one or more elements of the set by a syntonic comma (=81/80). In the tables 5.1, 5.2, and 5.3, the diatonic chords are listed with a number of possibilities for tuning.

Since the tone space is infinitely big, there are infinitely many tunings for a chord, however only some musically logical ones are listed here to give an example. In the first column of every table the name of the chords with corresponding note names is given. In the other columns different tunings and their compositions in the plane are given. The tones are indicated by circles, the black circle being the root of the chord (C was chosen to be the root in all cases). For every chord, the Gradus Suaventatis is calculated and given in the tables. We can test our first hypothesis which says that the convex composition (and if there is ambiguity, the most compact convex composition) represents the preferred intonation. One can

3-note chords	convex	
major triad C-E-G	$1 - 5/4 - 3/2$ · ⊙ · · ● ⊙ · · · $\Gamma = 9$	$1 - 5/4 - 40/27$ ⊙ · · ⊙ · · · ● $\Gamma = 16$
minor triad C-E♭-G	$1 - 6/5 - 3/2$ · · · · ● ⊙ · · ⊙ $\Gamma = 9$	$1 - 32/27 - 3/2$ · · · · · ⊙ · · ● ⊙ · · · · · $\Gamma = 15$
diminished triad C-E♭-G♭	$1 - 6/5 - 36/25$ ● · · · ⊙ · · · ⊙ $\Gamma = 15$	$1 - 32/27 - 64/45$ ⊙ · · ● · ⊙ · · $\Gamma = 17$
augmented triad C-E-G♯	$1 - 5/4 - 25/16$ · ⊙ · · ⊙ · · ● · $\Gamma = 13$	$1 - 5/4 - 125/81$ ⊙ · · · · · · · · · · · · · ⊙ · · · · ● $\Gamma = 23$

Table 5.1: Harmonic chords consisting of 3 notes. Of each chord, the convex configuration is given, together with another possible configuration. More configurations (intonations) are possible but only one is given here. The circles represent the notes in the frequency ratio space, the black circle representing the tonic C of the chord.

see that for almost every chord the convex composition of it in the tone space is more consonant according to Euler (i.e., lower value for Γ) than the other. There are two exceptions to this which are the diminished seventh chord and the dominant eleventh chord. The diminished seventh chord can be tuned in various ways to give the same 'consonance value'. This can perhaps be explained from the fact that this chord is a dissonant chord and therefore changing one of the intervals by a comma (81/80) has less impact than doing this with a more consonant chord[1].

The dominant eleventh chord is in a sense a reduction of the dominant thirteenth chord, only one note is missing. Filling in the missing note in the composition that is most favored, one obtains the most consonant thirteenth chord which is convex as well. In this way, we can understand why this particular composition for the eleventh chord is more consonant than the convex one. However, this second composition is more compact than the first one, supporting hypothesis number 2, which says to prefer the most compact configuration. This hypothe-

[1]The syntonic comma $81/80 = 3^4/(2^4 \cdot 5)$ has factors of 2 3 and 5 in it. Therefore, if the Least common multiple of a chord is already high, the chance that it changes a lot after one of the intervals is multiplied by a comma is low, since the LCM is constructed by multiplying the highest powers of 2, 3 and 5 form the intervals.

4-note chords	convex		
dominant seventh chord C-E-G-B♭	1:5/4:3/2:9/5 ⊙ · · ● ⊙ · · · ⊙ $\Gamma = 15$	1 : 5/4 : 3/2 : 16/9 · · ⊙ · ⊙ · ● ⊙ · · · · $\Gamma = 17$	
major seventh chord C-E-G-B	1 : 5/4 : 3/2 : 15/8 · ⊙ ⊙ · ● ⊙ · · · $\Gamma = 10$	1 : 5/4 : 3/2 : 50/27 ⊙ · · · · · · · ⊙ · · · · ● ⊙ $\Gamma = 18$	
minor seventh chord C-E♭-G-B♭	1 : 6/5 : 3/2 : 9/5 · · · ● ⊙ · · ⊙ ⊙ $\Gamma = 11$	1 : 6/5 : 3/2 : 16/9 ⊙ · ● ⊙ · · · ⊙ $\Gamma = 16$	
half-diminished seventh chord C-E♭-G♭-B♭	1 : 6/5 : 36/25 : 9/5 ● · · · ⊙ ⊙ · · ⊙ $\Gamma = 15$	1 : 6/5 : 36/25 : 16/9 ⊙ · ● · · · · · ⊙ · · · · · ⊙ $\Gamma = 19$	1 : 6/5 : 64/45 : 16/9 ⊙ · ● · ⊙ · · ⊙ $\Gamma = 17$
major-minor seventh chord C-E♭-G-B	1 : 6/5 : 3/2 : 15/8 · ⊙ ● ⊙ · ⊙ $\Gamma = 15$	1 : 6/5 : 3/2 : 50/27 ⊙ · · · · · · · · · · · · ● ⊙ · · · · ⊙ $\Gamma = 23$	
augmented seventh chord C-E-G♯-B	1:5/4:25/16:15/8 ⊙ · ⊙ ⊙ ● · $\Gamma = 15$	1:5/4:125/81:50/27 ⊙ · · · · · ⊙ · · · · · · · ⊙ · · · · ● $\Gamma = 25$	1:5/4:25/16:50/27 ⊙ · · ⊙ · · · ⊙ · · · ● $\Gamma = 20$
diminished seventh chord C-E♭-G♭-B♭♭	1:6/5:36/25:216/125 ● · · · · ⊙ · · · · ⊙ · · · · ⊙ $\Gamma = 22$	1:6/5:64/45:128/75 · · ● · ⊙ · · ⊙ · ⊙ · · $\Gamma = 22$	1:6/5:36/25:128/75 · · · · · ● · · · · ⊙ · ⊙ · · ⊙ $\Gamma = 22$
major triad with added sixth C-E-G-A	1 : 5/4 : 3/2 : 5/3 ⊙ ⊙ · · ● ⊙ $\Gamma = 11$	1 : 5/4 : 3/2 : 27/16 ⊙ · · · ● ⊙ · ⊙ $\Gamma = 13$	
minor triad with added sixth C-E♭-G-A♭	1:6/5:3/2:8/5 · · · · ● ⊙ · ⊙ ⊙ $\Gamma = 11$	1:6/5:3/2:81/50 ● ⊙ · · · · ⊙ · · · · · · · ⊙ $\Gamma = 19$	

Table 5.2: Harmonic chords consisting of 4 notes.

5/6/7-note chords	convex	
dominant ninth chord C-E-G-B♭-D	1:5/4:3/2:9/5:9/8 ⊙ · · ● ⊙ ⊙ · · ⊙ $\Gamma = 16$	1:5/4:3/2:16/9:10/9 ⊙ · ⊙ · ⊙ · ● ⊙ $\Gamma = 17$
dominant eleventh chord C-E-G-B♭-D-F	1:5/4:3/2:9/5:9/8:27/20 ⊙ · · · ● ⊙ ⊙ · · · ⊙ ⊙ $\Gamma = 18$	1:5/4:3/2:16/9:10/9:4/3 ⊙ · ⊙ · ⊙ ⊙ ● ⊙ · · · · $\Gamma = 17$
dominant thirteenth chord C-E-G-B♭-D-F-A	1 : 5/4 : 3/2 : 9/5 : 9/8 : 27/20 : 27/16 ⊙ · · · ● ⊙ ⊙ ⊙ · · ⊙ ⊙ $\Gamma = 19$	1 : 5/4 : 3/2 : 16/9 : 10/9 : 4/3 : 5/3 ⊙ ⊙ ⊙ · ⊙ ⊙ ● ⊙ · · · · also convex! $\Gamma = 17$

Table 5.3: Diatonic chords consisting of 5, 6 or 7 notes.

sis was also validated in all other cases except for the diminished seventh chord. The second configuration of the diminished seventh chord listed in table 5.2 is the most compact one. The compactness of a configuration may be difficult to judge at first sight. In the next section we will present a mathematical formula to calculate compactness. We checked (with a Matlab program, as we will see) all other tuning possibilities of these chords by multiplying one or more of the ratios with $(81/80)^n$ and verifying whether this resulted in a lower value for Γ. Note that both listed dominant thirteenth chords are convex. The second one listed is the preferred one according to both hypotheses, since it is more compact. This is also the configuration which is preferred by Euler's function.

To sum up, we proposed two hypotheses in order to present the best intonation, the first saying to prefer (the most compact) convex configuration, and the second saying to prefer the most compact configuration. The values of consonance of the chords were calculated using Euler's Gradus function. Of the 16 chords, for 14 of them hypothesis 1 was validated. For 15 of them, hypothesis 2 was validated. The exceptions can be explained from music theory and from the convexity theory itself. In section 5.3 the correlation between the concepts of consonance, convexity and compactness will be investigated further. Moreover, we want to stress that "preferred tuning" in this case is only based on the sound of the chord in isolation. In musical practice, there can be more than one choice for the intonation of a chord depending on its musical function in the chord sequence. However, this can still be a very useful measure because it can serve as the beginning of a full tuning theory.

5.2 Compactness and Euler

In the above sections we have shown that the convex and compact configuration of a set of notes may give an indication of the most consonant sound, as predicted by the measure of Euler. In this section we will try to formalize this in a mathematical way. Looking at the formula for Euler's Gradus function (eq. 5.1), we understand that the value for Γ becomes bigger when there are more factors of $2, 3$ and 5 in the LCM of the chord. If two notes are close together in the tone space, they have many prime factors in common, as explained in the previous section. This suggests that the Gradus function is related to the compactness of a set. We will therefore try to formalize the relation between compactness and consonance according to Euler's function.

Since our 2-dimensional tonal space (figure 5.1) neglects all factors of 2, we first consider the 3-dimensional tonal space that allows also all octave transpositions. In this way, we consider all axes representing the powers of primes, that is the x-axis representing the powers of 2, the y-axis representing the powers of 3 and the z-axis representing the powers of 5. In this coordinate system, a point with coordinates $(2, 4, 3)$, represents the number $2^2 \cdot 3^4 \cdot 5^3$.

5.2.1 Compactness in 3D

We want to consider a set of points (representing a chord) in the 2-3-5-space (octave-fifth-third-space) and measure its compactness. Several definitions of compactness are possible. The most intuitive way to measure compactness for our purposes is to sum the distances between all pairs of points[2]. The lower the value of the sum, the more compact the set is. The compactness C of a set of notes is then defined as follows:

$$C = \sum_{1 \leq i,j \leq n} |\vec{x_i} - \vec{x_j}| = \sum_{1 \leq i,j \leq n} \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + (x_{i3} - x_{j3})^2}, \quad (5.2)$$

where $\vec{x} = (x_1, x_2, x_3)$ defines the coordinates of a tone in the tone space. The term $x_{i1} - x_{j1}$ now defines the difference in the factor 2, the $x_{i2} - x_{j2}$, the difference in the factors 3, and $x_{i3} - x_{j3}$ the difference in the factors 5.

Each tone (note name) has more than one position in the tone space, which means that each chord has several compositions in tone space as we saw in the previous section. The factor that changes the frequency ratio but keeps the note name constant is $81/80 = 2^{-4}3^4 5^{-1}$. Multiplying a frequency ratio by this factor means moving a point in the $2, 3, 5$-coordinate system over $(-4, 4, -1)$. Given a set of points $\vec{x_1}\vec{x_2}\ldots\vec{x_n}$, every $\vec{x_i}$ has a number of possible coordinates such that

[2]Note that this concept of compactness is different from the concept of a 'compact set' in topology.

the point represents the same note name:

$$\vec{x_i} = \begin{pmatrix} x_{i1} - 4k_i \\ x_{i2} + 4k_i \\ x_{i3} - k_i \end{pmatrix}, \qquad k \in \mathbb{Z}. \tag{5.3}$$

The compactness C can then be written as:

$$\begin{aligned} C &= \sum_{1 \le i,j \le n} \sqrt{X_1 + X_2 + X_3}, \\ X_1 &= (x_{i1} - x_{j1} - 4(k_i - k_j))^2 \\ X_2 &= (x_{i2} - x_{j2} + 4(k_i - k_j))^2 \\ X_3 &= (x_{i3} - x_{j3} - (k_i - k_j))^2 \end{aligned} \tag{5.4}$$

and the most compact configuration of a set $\vec{x_1}\vec{x_2}\dots\vec{x_n}$ is given by the[3] $k_2, k_3, \dots, k_n$ for which C has a minimum.

The value for Euler's Gradus function can now be calculated for a certain configuration of points. Therefore, we first need to find the Least Common Multiple (LCM) of the chord. To be able to find the LCM of a chord we have to write the chord in the form $a : b : c$ such that a, b, c are integers (just like we did in section 2 were the chord $1 : 5/4 : 3/2$ was written as $4 : 5 : 6$). Since the point x_{ij} represents the j^{th} coordinate (meaning the multiples of 2,3 or 5) of note x_i, a whole frequency ratio is expressed as $2^{x_{i1}-4k_i} \cdot 3^{x_{i2}+4k_i} \cdot 5^{x_{i3}-k_i}$. A (3-note) chord $a : b : c$ can therefore be written as

$$\begin{aligned} a : b : c = \quad & 2^{x_{11}-4k_1} \cdot 3^{x_{12}+4k_1} \cdot 5^{x_{13}-k_1} : \\ & 2^{x_{21}-4k_2} \cdot 3^{x_{22}+4k_2} \cdot 5^{x_{23}-k_2} : 2^{x_{31}-4k_3} \cdot 3^{x_{32}+4k_3} \cdot 5^{x_{33}-k_3}. \end{aligned} \tag{5.5}$$

If the chord is already in a form such that a, b, c are (positive) integers, the LCM of 5.5 can be found as follows:

$$\mathrm{LCM} = 2^{\max\{x_{11}-4k_1,\dots,x_{n1}-4k_n\}} \cdot 3^{\max\{x_{12}+4k_1,\dots,x_{n2}+4k_n\}} \cdot 5^{\max\{x_{13}-k_1,\dots,x_{n3}-k_n\}} \tag{5.6}$$

where $\max\{a_1, \dots, a_n\}$ picks the largest of numbers a_1 to a_n.

When a, b, c are not integers, the expression for the LCM looks a bit different. To write the chord $a : b : c$ in a form such that it is represented by integers, a, b, c should be multiplied by the Least Common Multiple (LCM) of the denominators of a, b, c (for example to write the chord $1 : 5/4 : 3/2$ as $4 : 5 : 6$ each ratio was multiplied by $\mathrm{LCM}(1, 2, 4) = 4$). The fact that a, b and c are split into powers of $2, 3$ and 5 makes this process easier. Instead of finding the LCM of the

[3]One of the k_i is fixed (k_1 in this case) and set to zero because the set needs to have a reference point. If all k_i were to be chosen freely, many sets with the same compactness but different locations may exist.

denominators we just need to find the maximum of all factors of $2, 3$ and 5 in the denominators. The LCM (5.5) then changes as follows:

$$\mathrm{LCM} = 2^{v_1} 2^{\max\{x_{11}-4k_1,\ldots,x_{n1}-4k_n\}} . 3^{v_2} 3^{\max\{x_{12}+4k_1,\ldots,x_{n2}+4k_n\}} . 5^{v_3} 5^{\max\{x_{13}-k_1,\ldots,x_{n3}-k_n\}} \tag{5.7}$$

where

$$v_j = \max A, \qquad A = \{-z | z \in B_j \,\&\, z < 0,\ j = 1, 2, 3\} \tag{5.8}$$

and

$$B_j = \bigcup_{1 \le i \le n} \{z_{ij}\} \quad \text{where} \quad \begin{cases} z_{i1} = x_{i1} - 4k_i \\ z_{i2} = x_{i2} + 4k_i \\ z_{i3} = x_{i3} - k_i \end{cases} \tag{5.9}$$

and n is the number of notes in the chord. Finally, the value for Γ (defined in eq. 5.1) is given as follows:

$$\begin{aligned} \Gamma &= 1 + v_1 + \max\{x_{11} - 4k_1, \ldots, x_{n1} - 4k_n\} \qquad (5.10) \\ &+ 2 * (v_2 + \max\{x_{12} + 4k_1, \ldots, x_{n2} + 4k_n\}) \\ &+ 4 * (v_3 + \max\{x_{13} - k_1, \ldots, x_{n3} - k_n\}). \end{aligned}$$

We now have expressions for C and Γ and we would like to see that the $k_2 \ldots k_n$ that make C minimal, also make Γ minimal. This would prove our hypothesis: the configuration of a chord that is most compact is also most consonant. This turns out to be quite hard to solve. To minimize a function one normally takes the derivative and sets this to zero. However, the function Γ is of such a complicated form (the function 'max' gives Γ a discrete character) that it is not possible to take the derivative to k. Also, for $n > 2$, it is computationally very intensive to solve the equations $\partial k_i C = 0$. For $n = 2$, taking one point at $(0, 0, 0)$ with $k_1 = 0$ and one at (a, b, c) with $k_2 = k$:

$$C = \sqrt{(a - 4k)^2 + (b + 4k)^2 + (c - k)^2}, \tag{5.11}$$

$$\Gamma = 1 + |a - 4k| + 2|b + 4k| + 4|c - k|. \tag{5.12}$$

We can find the value for k that minimizes C by taking the weighted mean of the k's that minimize the separate parts. The minimum for the term $(a - 4k)$ is at $k = 1/4 \cdot a$ and should be weighed with a factor 4^2. The total is divided by $4^2 + 4^2 + 1 = 33$. The k that minimizes C is

$$k_c = 1/33(4a - 4b + c). \tag{5.13}$$

The k that minimizes Γ is calculated in the same way and equals

$$k_\Gamma = 1/14(a - 2b + 4c). \tag{5.14}$$

Of course k should be an integer, but simply rounding off the values given by 5.13 and 5.14 has turned out not to be correct. Still, from these equations we can see that the values for k are not the same: the value for k that minimizes C is not always equal to the value for k that minimizes Γ.

Therefore, a Matlab program is written that can calculate the values for C, Γ and the values for k that make both equations minimal. This is done by varying the coordinates of point 2 over all points of the 3-Dim space in which every coordinate runs from -4 to 4. Point 1 is taken at the origin. It turns out that in 86.5% of the cases both C and Γ have a minimum for the same k. For the case $n = 3$ can in the same way also be calculated if the same value for k_i makes the C and Γ minimal. It turns out this is true for 70.1% of the cases. For $n > 3$ the problem becomes computationally very intensive[4]. However, from this we can conclude that the hypothesis: the more compact, the more consonant, is not always true, but it is true in the majority of the cases.

5.2.2 Compactness in 2D

In the 2-D space where all chords are projected, the frequency ratios are considered under octave equivalence. Instead of considering the x_{i1}, x_{i2} and x_{i3} component (2,3 and 5 component) we only consider the x_{i2} and x_{i3} component (for convenience we kept these names) in the 2D space. Thus the expression for C becomes simpler:

$$C = \sum_{1 \leq i,j \leq n} \sqrt{(x_{i2} - x_{j2} + 4(k_i - k_j))^2 + (x_{i3} - x_{j3} - (k_i - k_j))^2}. \tag{5.15}$$

The expression for Γ however, becomes more complicated. The first term $\max\{x_{11} - 4k_1, \ldots, x_{n1} - 4k_n\}$ changes. A point in the plane now only is specified by its 3- and 5-components: $3^{x_{i2}+4k_i}5^{x_{i3}-k_i}$. The factor 2^n that together with this specifies the whole frequency ratio: $2^n \cdot 3^{x_{i2}+4k_i} \cdot 5^{x_{i3}-k_i}$, only serves to keep the frequency ratio within the interval $[1, 2)$. Therefore, to find an expression for n, we need to solve:

$$1 \leq 2^n \cdot 3^{x_{i2}+4k_i} \cdot 5^{x_{i3}-k_i} < 2. \tag{5.16}$$

From this n can be analytically solved:

$$-(x_{i2}+4k_i)\cdot\log_2 3-(x_{i3}-k_i)\cdot\log_2 5 \leq n < 1-(x_{i2}+4k_i)\cdot\log_2 3-(x_{i3}-k_i)\cdot\log_2 5. \tag{5.17}$$

Since n should be an integer, this makes:

$$n = \lceil -(x_{i2} + 4k_i) \cdot \log_2 3 - (x_{i3} - k_i) \cdot \log_2 5 \rceil, \tag{5.18}$$

[4] The number of possible configurations of a set consisting of n points, increases with n as $\binom{728}{n-1}$, since this expresses the number of possibilities to choose n points from a $9 \times 9 \times 9$ lattice where one note is fixed in the origin.

where $\lceil x \rceil$ is the smallest integer greater or equal to x. We can therefore understand that the first term in Γ can now be replaced by $\max\{\lceil -(x_{12}+4k_1)\cdot\log_2 3-(x_{13}-k_1)\cdot\log_2 5\rceil, \lceil -(x_{22}+4k_2)\cdot\log_2 3-(x_{23}-k_2)\cdot\log_2 5\rceil, \ldots, \lceil -(x_{n2}+4k_n)\cdot\log_2 3-(x_{n3}-k_n)\cdot\log_2 5)\rceil\}$, thus Γ becomes:

$$\begin{aligned} \Gamma &= 1+v_1+\max\{\lceil -(x_{12}+4k_1)\cdot\log_2 3-(x_{13}-k_1)\cdot\log_2 5\rceil, \ldots \\ &\quad \ldots, \lceil -(x_{n2}+4k_n)\cdot\log_2 3-(x_{n3}-k_n)\cdot\log_2 5\rceil\} \\ &+ 2*(v_2+\max\{x_{12}+4k_1,\ldots,x_{n2}+4k_n\}) \\ &+ 4*(v_3+\max\{x_{13}-k_1,\ldots,x_{n3}-k_n\}), \end{aligned} \tag{5.19}$$

with v_1 now given by:

$$\begin{aligned} v_1 &= \max\{-z|z\in B\,\&\,z<0\}, \\ B &= \bigcup_{1\le i\le n}\{\lceil -(x_{i2}+4k_i)\cdot\log_2 3-(x_{i3}-k_i)\cdot\log_2 5\rceil\} \end{aligned} \tag{5.20}$$

and v_2, v_3 as given in eq 5.8,5.9. Using these expressions, we calculate the number of cases for which the value of k that makes C minimal also makes Γ minimal. In table 5.4 all percentages are given. Surprisingly, the percentages for the 2-D lattice are higher than for the 3-D lattice. Using the Matlab program we

lattice	number of notes	percentage correct
3-D	2	86.5 %
3-D	3	70.1 %
2-D	2	97.5 %
2-D	3	85.4 %
2-D	4	76.8 %

Table 5.4: Results of testing the hypothesis: the configuration of a chord that is most compact is also most consonant.

have also checked all chords that are listed in tables 5.1, 5.2 and 5.3, to be sure that we indeed listed the most compact configurations in these tables. It indeed turns out that the hypothesis "the configuration that is most compact, is the most consonant according to Euler's value" is true for all chords except for the diminished seventh chord.

Interpretation of results

How can these results be explained, and can they perhaps be related to convexity? We have tested the hypothesis "the more compact, the more consonant" for all possible 2,3,4-tone sets within a 2-D 9×9 lattice or a 3-D $9\times 9\times 9$ lattice. It turns out that for neither space the hypothesis is 100 percent true, but the correct percentages are nevertheless reasonably high. In the 2-D space we gained a little higher percentage than in the 3D space. The fact that the relationship between consonance and compactness is not a one to one correspondence has to do with the weights in the definition of Γ and with the measurement of the syntonic comma in the 2-D and 3-D space. The weights $1, 2$ and 4 in the definition for Γ (eq. 5.10) cause Γ to change more due to a shift in the 3-coordinate than to a shift in the 2-coordinate (factor 2). Γ is changed most due to a shift in the 5-coordinate (factor 4). Therefore we can also understand that the percentages decrease as the number of notes increase: the more notes, the more directions in the lattice are involved. In the expression for C (eq. 5.4) we see two times a factor 4 (which comes from the syntonic comma $2^{-4}3^{4}5^{-1}$) in the terms that are concerned with the distances in the 2 and 3 direction. This means that the compactness C is more influenced by changes in the 2 and 3 direction than by changes in the 5 direction. We therefore understand that there cannot be a one to one correspondence between the compactness C and the consonance measure Γ.

It was already intuitively clear that consonance Γ could be related to compactness C in a general way. We have seen that for a high percentage of sets, the most consonant configuration is also the most compact one. Consider a convex and highly compact set in the 3-D space centered around the origin. If one of the elements of the set is moved by the syntonic comma it is moved by the vector $(-4, 4, -1)$ (since $81/80 = 2^{-4}3^{4}5^{-1}$). The new set is always less compact than it was if the size of the set is less than a certain number of elements, since the vector $(-4, 4, -1)$ then takes the element outside the area spanned by the other elements (which makes the set less compact). For a set in the 2-D space the syntonic comma is represented by the vector $(4, -1)$ which yields the same conclusion. Consider now a convex and highly compact set centered in the lattice, and imagine what happens with the consonance Γ if one or more elements are shifted by a syntonic comma. Again if the number of elements is within a certain range, a shift by $81/80 = 2^{-4}3^{4}5^{-1}$ will increase the LCM of the chord and the new chord will be less consonant. Since here we have only observed sets consisting of $2, 3$ and 4 elements it is understandable that shifting one or more elements of the set by a syntonic comma makes the set less compact and less consonant.

Now we want to make a connection to convexity. If we look at table 5.3, the last column represents chords that are all introduced as alternative intonations of the chords mentioned. It is remarkable that these configurations all have the same value for Γ, namely $\Gamma = 17$. By looking at the configurations, we understand that it doesn't matter if the inner notes are filled, the value of Γ just depends on the boundary notes. This is understandable since the value of Γ only depends on the Least Common Multiple of the frequency ratios (written in integers as we have shown) instead of depending on all frequency ratios. The LCM picks the highest factors of 2,3 and 5, which precisely indicate the boundaries of the chord. Therefore the value of Γ of a chord equals the value of Γ of the chord that represents the convex hull. The *convex hull* of a set of points S is the intersection of all convex sets containing S. For example, the rightmost chord at the bottom of table 5.3 represents the convex hull of the two chords above this chord. It can now be seen that whenever there is a possibility of a convex configuration of a chord, this will often be the most compact one. How often that is, we will investigate in the next section.

5.3 Convexity, compactness and consonance

In the previous experiments we varied the coordinates of a (2-D or 3-D) space to represent sets of notes for which we wanted to calculate whether the most compact configuration corresponded to the most consonant configuration. We now want to know whether these sets do also correspond with a convex configuration. More precisely: which percentage of the sets of notes that have a possible convex configuration, have a convex configuration that corresponds with 1) the most compact configuration, and 2) the most consonant configuration. For some chords, there is no possible intonation such that the notes form a convex set in the tone space. For these chords, only the compactness can say something about the preferred intonation. Figure 5.2 illustrates what percentages we are looking for. In the figure, the set S consists of all configurations of all possible chords con-

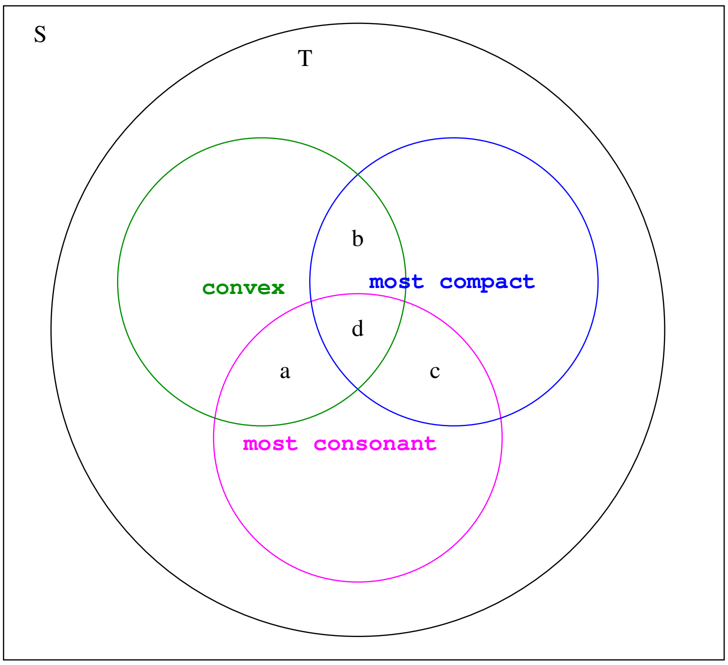


Figure 5.2: Illustration of overlap of convex, most compact and most consonant configuration when trying to find the preferred intonation of a chord.

sisting of n notes. The set T consists of all configurations of the chords that have a possible convex configuration. Within the set T, the set '`convex`' represents all convex configurations. Then, '`most compact`' is the set consisting of every most compact configuration of each chord (in T). Similarly, '`most consonant`' is the set consisting of every most consonant configuration of each chord (in T). The intersection of sets of our interest are given in equation 5.21.

$$\begin{aligned} a \cup d &= \texttt{convex} \ \cap \ \texttt{consonant} \\ b \cup d &= \texttt{compact} \ \cap \ \texttt{convex} \\ d \cup c &= \texttt{compact} \ \cap \ \texttt{consonant} \\ d &= \texttt{compact} \ \cap \ \texttt{convex} \ \cap \ \texttt{consonant} \end{aligned} \tag{5.21}$$

The percentages that we are looking for, are obtained by dividing the number of elements of the sets given in eq. 5.21, by the number of chords that have a possible convex configuration. Note that the latter value is not equivalent to the number of elements in T, since this set represents the number of configurations instead of the number of chords.

There could be more than one convex possibility per set (as is the case with the dominant thirteenth chord in table 5.3). Also, it is possible that more than one configuration has the same (lowest) value for the compactness or consonance (although this rarely happens). This only means that some solutions are not unique, but since we count the number of chords and not the number of configurations, this does not change the obtained percentages.

We have written a program in Matlab that finds all possible $2, 3$ and 4 note sets in a 9×9 (coordinates run from -4 to 4) 2-dimensional lattice, and calculates for each set 1) whether it has a convex configuration and which configurations are convex, 2) the configuration that is most compact, and 3) the configuration that is most consonant. To distinguish between the configurations, a variable k is used in the same way as introduced in equations 5.15 and 5.19 for the expressions for compactness and consonance. The convexity of a set is calculated from the coordinates of the elements in the set; the coordinates change with k as in equation 5.3. The variable k is varied from -2 to 2. If a wider range was chosen, the obtained points (5.3) would lie outside the lattice. From short test-runs it was concluded that it is sufficient to work with a 9×9 2-dimensional lattice, a bigger lattice did not significantly change the percentages. This conforms to our intuition, since we found a high correlation between consonance and compactness (table 5.4), and the more compact a set is, the better it fits into a smaller lattice. One point is chosen in the center $(0, 0)$, so for $n = 2$ only one point is varied, for $n = 3$ two points, and so on. To ensure that some sets are not counted twice, point 3 is varied over the points that point 2 has not been varied over[5] and so on for the points thereafter (point 1 is fixed). The number of possible sets for

[5] In pseudo code:

n points is then calculated as follows: the lattice contains $9 \times 9 = 81$ points. One note is fixed at the origin so there are 80 points left for notes 2 to n to vary over. The number of possibilities[6] to choose $n-1$ point from 80 points is $\binom{80}{n-1}$. Table 5.5 shows the number of sets that can be chosen from the lattice for the number of notes varying from 2 to 4. Observing that this number increases very

n	number of possible sets for n points
2	$\binom{80}{1} = 80$
3	$\binom{80}{2} = \frac{80 \cdot 79}{2!} = 3160$
4	$\binom{80}{3} = \frac{80 \cdot 79 \cdot 78}{3!} = 82160$

Table 5.5: Number of possible sets of n point in a 9×9 lattice with point 1 fixed in the origin.

fast as a function of n, one can understand that it is computationally impossible to go much beyond $n = 4$. These numbers of possible sets would be the numbers that are examined with our algorithm if all these sets have a convex possibility. This turns out not to be the case, so the number of examined sets is reduced. The results of the Matlab program are shown in table 5.6. In this table it is also

percentage	n=2	n=3	n=4
compact & consonant	97.5 %	85.4 %	85.6 %
convex & consonant	16.3 %	41.4 %	41.2%
convex & compact	11.3 %	40.8 %	36.0%
compact & convex & consonant	11.3 %	37.3 %	34.1%
number of sets examined	80	1590	14810

Table 5.6: Results of the percentages as indicated in figure 5.2.

indicated how many sets have a convex possibility and are therefore examined. We see that for $n = 2$ all 80 sets have a possible convex configuration. For $n > 2$, this is not the case anymore. For example for $n = 3$, only 1590 sets of the 3160 possible sets have a possible convex configuration.

Observing the results, one can see that the biggest correlation can be found between the most compact and consonant sets, as we expected. The correlation

```
for i: from 1 to total number of points do
    vary point 2
    for j: from i to total number of points do
        vary point 3
    end
end
```

[6] $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$

between convexity and the other items is very low for $n = 2$ and gets higher as the number of notes increases. This agrees with our intuition too, since our understanding of the relation between convexity and consonance was through the notion of compactness (see the end of the previous section). When considering only 2 notes, the notion of convexity differs a lot from the notion of compactness, since two notes form a convex set if a line can be drawn between the two notes on which no other notes lie. Therefore it is not easier for two notes to form a convex set if the notes lie close to each other than when the notes lie far from each other, as can be seen from the low correlation between convexity and compactness for $n = 2$. However, for increasing n, the correlation between convexity and compactness increases as well. Note that regions a and d are really small, especially for small n (for $n = 2$, $a = 0$). This means that when the most consonant configurations are also convex, they are most likely to be also the most compact configurations (a); and when the most compact configurations are also convex, they are most likely to be also the most consonant configurations (b). The results from table 5.4 differ from the results “compact & consonant” in table 5.6. This difference is due to the difference in sets that is taken into account. In the experiment leading to the results of table 5.6 only the sets that have a possible convex configuration were taken into account. At the end of section 5.2.2 we explained why the increase of n causes a decrease of the percentage “compact & consonant”. Remarkably, there is no decreasing percentage if n increases from 3 to 4, when sets with a convex possibility are considered (table 5.6).

The reason that we only calculated these percentages on the 2-dimensional lattice, is that our convexity routine only works for 2-dimensional sets. We intend to generalize this routine to enable it to work for three and higher dimensional lattices.

5.4 Concluding remarks on compactness and convexity

In this chapter we have investigated whether the notions of convexity and compactness can be used in an intonation model for chords in isolation. It has turned out that the most compact configuration of a chord corresponds up to a high percentage to the most consonant configuration of a chord when using the consonance measure given by Euler’s Gradus function.

Remarkably, the correspondence of convexity and consonance according to Euler is only around 40 percent (table 5.6). This means that, although we found in section 5.1.2 a correspondence of 14 out of 16 for the diatonic chords, convexity is not as good a measure of consonance as initially thought. The correspondence between convexity and compactness, measured as the percentage of most compact sets that are also convex, is even less (table 5.6).

In section 4.2.2 in chapter 4, a definition for a convex set on a note name space was given. Since there is ambiguity among the locations of the note names in the tone space, we decided to define a set in the note-name space to be convex if there exists a convex configuration that can be projected from the frequency ratio space corresponding to this set of note names. Since in the current chapter it became clear that compactness is more related to consonance than convexity is, we might want to change the definition of a convex set in the note name space into: a set in the note name space is convex if the most compact configuration of this set is convex. To this end, let us consider the sets from table 4.4 with this new definition. We noted that the minor chord with raised fifth: $C, E\flat, G\sharp$ was a convex set using the old definition, but in the new definition it is not. The most compact configuration[7] of this chord is the set of ratios $1, 6/5, 25/26$ projected on the note name space, which does not form a convex set. This is in line with our intuition and the non-convexity of this chord fits into the low percentage of convex altered chords observed in section 4.2.4. Furthermore, we recall that the $IV - V$ progression in section 4.2.5 did not form a convex set. However, it does form the most compact configuration of possible $IV - V$ configurations projected from the frequency ratio tone space.

We can conclude that although convexity is a useful property to characterize scales in the frequency ratio tone space, the definition of a convex set in a space that can be represented as a cylinder (such as the note name space) is open to adjustments to be able to characterize sets in this space as well as possible. Moreover, in a model for consonance, compactness turns out to be a more useful measure. In the next chapter, we will use both convexity and compactness for two computational applications. Before incorporating the notion of convexity we have to decide between the two alternative definitions.

[7] Although compactness should be considered in the tone space having 3/2 and 5/4 as basis vectors, in chapter 4 for this chord it does not make a difference.

Chapter 6

Computational applications of convexity and compactness

In this chapter we will look at two computational applications of the notions of convexity and compactness, one which turns out to be only moderately successful (modulation finding) and the other turns out to be much more successful (pitch spelling). Both problems have been previously dealt with by various authors (Chew 2002; Longuet-Higgins 1987a; Temperley 2001; Meredith 2003, 2006; Cambouropoulos 2003; Chew and Chen 2005). In the first part of this chapter we deal with modulation finding. We have seen that convexity is a property that many pitch structures like scales, chords and harmonic reductions possess. Therefore, among chords or sets of notes within a certain scale, it may be special if a chord or set is non-convex. We will study the meaning of some non-convex sets and will see that it can indicate a modulation in music.

In the second part of this chapter of which a part has been published as Honingh (2006b), we study pitch spelling. In the previous chapter, we used compactness to make a projection from the note name space (containing no information about intonation) to the frequency ratio space (which does contain information about intonation). In turn, one can wonder about the projection from the pitch number space (containing no information about the note names) to the note name space. The problem of finding such a projection is known as the pitch spelling problem. We attempt to formalize this projection by using compactness.

6.1 Modulation finding

When a piece of music is said to be in a specific key, we usually mean that the piece starts and ends in this key. It rarely occurs that the piece is entirely in the same key, other keys can occur at several places in the music. A modulation is the act or process of changing from one key to another. In the research of key finding (see for example Krumhansl 1990; Temperley 2001; Longuet-Higgins and

Steedman 1971), the most difficult part of the analysis is usually formed by the modulations. The input for a key finding algorithm consists usually of the note names of a piece. The output of such an algorithm contains the key of the piece and sometimes local keys of smaller parts of the piece. At the location of a modulation, it can be preferred to find two keys being equally present at the point where one is changing into the other. However, sometimes, it is not possible to find a key at all at a certain point, since for example the modulation between two keys consists only of a chromatic melody. The key finding algorithms that are designed to determine the local key of segments in the music, have usually the most problems at the points where modulations occur. With their key-finding model, Vos and van Geenen (1996) detected only two of the six modulations that were analyzed by Keller (1976), when the model was tested on the 48 fugue subjects of Bach's Well-Tempered Clavier. Furthermore, it also found modulations in 10 other cases in which there was no modulation. Temperley (2001) tested his model on the same corpus and found two of the modulations correctly. Therefore, a specially designed program to indicate the modulations in a piece of music would be a helpful tool to implement in several key finding models. Chew (2002, 2006) has described a method specifically for determining key boundaries and points out that these kind of models are furthermore important for computer analysis of music, computational modeling of music cognition, content-based categorization and retrieval of music information and automatic generation of expressive performance.

6.1.1 Probability of convex sets in music

It has been observed (in chapter 4) that the major and minor scales as a whole form a convex set. Furthermore, the subsets of these scales constituting the diatonic chords form convex sets as well. This might suggest that non convex subsets of the diatonic scales are not so common. If this is indeed the case, a non-convex set within a piece of music may indicate that this specific set is not part of a diatonic scale which could indicate a modulation in the music. To verify the correctness of this reasoning, we need to investigate the convexity of all possible subsets of the diatonic scales. Hence, we will address the question "what is the chance for a set of notes within a piece of music to be convex?". Assuming a certain piece is in one and the same key, this means calculating the chance that a set from one scale is convex.

Since music is usually written in terms of notes (as opposed to frequency ratios or pitch numbers), we need to use the definition of convexity that applies to note names. In chapter 4 (see section 4.2.2) we gave the following definition, which we will call definition 1:

6.1.1. Definition. If (at least) one of the configurations of a set of note names has a convex projection from the frequency ratio space, the set is said to be

number of notes in the set	percentage convex according to def. 1	percentage convex according to def. 2
2	100 %	95.24 %
3	94.29 %	68.57 %
4	94.29 %	74.29 %
5	100 %	66.67 %
6	100 %	100 %

Table 6.1: Percentage of n-note sets that are convex if chosen from a major scale. See text for details about definition 1 and 2.

convex.

However, in chapter 5 (see section 5.4) it was proposed that the definition might be changed to the more intuitive definition 2:

6.1.2. Definition. if the most compact configuration of a set of note names has a convex projection from the frequency ratio space, the set is said to be convex.

To see which definition is most useful for our purposes, we calculate the percentages of convex sets of n notes in a certain scale according to both definitions. A Matlab program was written for this purpose. The results are displayed in table 6.1. The values 1 and 7 are left out because the convexity of one note does not mean anything, and there is only one configuration for 7 notes within one scale which is the whole scale and which is necessarily convex. Let us shortly explain how these values were calculated. The computer program dealt with a 9×9 lattice that represents the note name space. This is big enough to contain all note sets that we wanted to consider and it contained also enough compositions of a set to calculate whether it is a convex set or not. Given a set of note names to the computer program, the program computed every composition in the 9×9 lattice. In case of definition 1 it was calculated for every composition whether it is convex or not (in practice the program stops when it has calculated a composition that is convex, because according to definition 1 the set is convex if there is at least one convex projection from the frequency ratio space); for definition 2, it was calculated whether the most compact configuration of a set is convex. In the 9×9 plane, every note name has 2 or 3 possible configurations. Therefore, if a set consists of n notes, the number of possible compositions lies between 2^n and 3^n.

Observing the percentages in table 6.1, we can choose between the two definitions of convexity of note names. We see that the percentages of convex sets according to definition 1 are higher than the percentages of convex sets according to definition 2. Therefore, using definition 1, the occurrence of a non-convex set is with high certainty an indication of something different than a diatonic subset (for example, a modulation). Using definition 2, the chance for a non-convex set to be just a diatonic subset is greater than when using definition 1. Therefore, for the purpose of finding modulations, definition no. 1 is most useful, and we will continue to use this definition.

It can be observed that the notes from a piece of music in one key do not only come from one scale, even if the piece of music is in one and the same key. There are often more notes that are used in a piece of music than only the notes from the scale of the tonic. For example, in the first fugue from the Well-tempered Clavier of Bach, which is written in C major, the notes that appear throughout the piece are the notes from the major scale in C plus the additional notes $F\sharp, B\flat, C\sharp$ and $G\sharp$ (together forming a convex set). The idea that the key contains more notes than the scale of the tonic is formalized in the book “De fis van Euler” (Van de Craats 1989). There, Van de Craats claims that in a major key, the augmented fourth is often used and should therefore be included in the scale. This means that in C major, the scale would contain the notes (given in a fifth sequence): $F, C, G, D, A, E, B, F\sharp$. A piece of music in C minor can contain the notes (given in a sequence of fifths): $D\flat, A\flat, E\flat, B\flat, F, C, G, D, A, E, B, F\sharp$, according to Van de Craats. In accordance with the latter claim, Longuet-Higgins (1987a) states that “a note is regarded as belonging to a given key if its sharpness relative to the tonic lies in the range -5 to +6 inclusive”. Results by other researchers (Youngblood 1985; Knopoff and Hutchinson 1983; Krumhansl and Kessler 1982) are in agreement with Longuet-Higgins’ and Van de Craats’ suggestions. Statistical analyses were accomplished by Youngblood (1985) and Knopoff and Hutchinson (1983). Their papers contained tables giving the total frequencies of each tone of the chromatic scale in a variety of compositions. Table 6.2 represents a summary of these studies which is taken from Krumhansl (1990). In the study by Youngblood, music from Schubert, Mendelssohn and Schumann was analyzed. Knopoff and Hutchinson analyzed three complete song cycles by Schubert and pieces by Mozart, Hasse and R. Strauss. Table 6.2 shows the total number of times that each tone of the chromatic scale was sounded in the vocal lines of the pieces. All keys have been transposed to C major or minor. Krumhansl and Kessler (1982) obtained similar results when measuring the degree to which each of the 12 chromatic scale notes fit in the particular key. In their experiment, the task was to rate how well the final probe tone “fit with” the context in a musical sense. This experiment resulted in key profiles for major and minor keys. The correlations between tone distributions (Youngblood 1985; Knopoff and Hutchinson 1983) and perceived tonal hierarchies (Krumhansl and Kessler 1982) are strong for both major and minor keys (comparison made in Krumhansl 1990).

We have seen that the notes from a piece of music in one key do not come from one scale. However, from the above listed results it is difficult to formalize a key content. The results of Youngblood (1985), Knopoff and Hutchinson (1983) and Krumhansl and Kessler (1982) do not directly suggest a dichotomy between tones belonging to a specific subset of the chromatic scale and the rest of the scale

Tone	Pieces in Major	Pieces in Minor
C	16.03 %	18.84 %
$C\sharp/D\flat$	0.97 %	2.14 %
D	14.97 %	11.43 %
$D\sharp/E\flat$	1.76 %	11.73
E	15.52 %	2.58 %
F	9.71 %	8.94 %
$F\sharp/G\flat$	2.77 %	2.43 %
G	18.04 %	21.66 %
$G\sharp/A\flat$	1.74 %	7.13 %
A	9.18 %	2.08 %
$A\sharp/B\flat$	1.80 %	5.38 %
B	7.50 %	5.65 %
Total number of notes	20,042	4,810

Table 6.2: Percentage of occurrences each tone in the chromatic scale in a variety of compositions (see text for details) published with tone occurrences (instead of percentages) in Krumhansl (1990), source taken from Youngblood (1985) and Knopoff and Hutchinson (1983). All notes within a key have been scaled to C, such that the percentages are shown with reference to tonic C.

(like Van de Craats does when claiming that $F, C, G, D, A, E, B, F\sharp$ is a special subset of the chromatic scale).[1] However, the data (both table 6.2 as well as the key profiles obtained by Krumhansl and Kessler 1982) does not contradict the claims of Van de Craats. If the tones of table 6.2 are represented in order of decreasing tonal frequencies, the set representing the major scale proposed by Van de Craats represents a coherent subset of the chromatic scale. For the minor scale this is trivial, since the minor scale proposed by Van de Craats embodies 12 elements which is as much as the chromatic scale when enharmonic equivalence is used. Since the data represented in table 6.2 does not distinguish between enharmonically equivalent notes, it is more convenient to use the diatonic scales that Van de Craats proposed, to get information about convex subsets of these scales. These 'scales' of 8 and 12 notes respectively can be used as input for our Matlab program, to calculate the percentages of sets that are convex. The results can be found in tables 6.3 and 6.4.

Of course, the bigger the total set of notes to choose from, the higher the percentages of non-convex subsets. Therefore, in table 6.4 the percentages of convex sets decrease to a minimum of 49.03% at $n = 6$, meaning that there

[1] If any, the subsets C, D, E, G or the major diatonic scale C, D, E, F, G, A, B represent subsets of the chromatic scale having all substantial high tonal frequencies.

number of notes in the set	percentage convex
2	100 %
3	92.86 %
4	88.57 %
5	92.86 %
6	100 %
7	100 %

Table 6.3: Percentage of n-note sets that are convex if chosen from the set of notes representing the C major scale with an additional $F\sharp$.

number of notes in the set	percentage convex
2	100 %
3	80.91 %
4	59.80 %
5	52.02 %
6	49.03 %
7	51.01 %
8	58.79%
9	71.82%
10	89.39 %
11	100%

Table 6.4: Percentage of n-note sets that are convex if chosen from the set of notes representing the C neutral minor scale with additional $D\flat, A, E, B, F\sharp$.

is a reasonable chance of finding a 6 note set that is non-convex in a piece of music written in a minor key. From both tables 6.3 and 6.4 we see that the highest percentages of convex sets are for the smallest and biggest possible sets in the key. Therefore the smallest and the biggest non-convex sets are the best indicators of modulations. We recall from the previous chapter that most altered chords form non-convex sets. Those chords form an additional difficulty in finding modulations by indication of non-convex sets: a non-convex set can also indicate an altered chord.

From the above results we learn that if we choose randomly a set of notes from one key, there is a high chance for the set to be convex. Therefore, we hypothesize that, if we analyze a piece of music by dividing it into sets of n notes, most of the sets are convex. It is thus more special in a piece of music for a set to be non-convex than convex. And because we have seen that sets from one key tend to be convex, a non-convex set within a piece could point to a change of key or modulation.

6.1.2 Finding modulations by means of convexity

A Matlab program is written that finds modulations in a piece of music by localizing non-convex sets. The idea is that a high percentage of sets that are taken from a specific key, is convex (as we saw in tables 6.1, 6.3 and 6.4). Therefore we expect something special in the music at locations where sets are not convex. The more sets that are not convex around a certain location, the stronger the indication of a change of key. From tables 6.1, 6.3 and 6.4 we see that both a low number of notes and a high number of notes have the highest chance to be convex. This means that in a piece of music, for these numbers of notes in a set it is more special to be non-convex, so this would be a stronger indication to a change of key, than for another number of notes. However, when segmenting a piece of music into a specific number of notes, we do not know if there are any notes represented twice in one set. The numbers in tables 6.1, 6.3 and 6.4 are only valid for sets of notes in which every note is different.

To be able to judge all n-tone sets on convexity, we introduce a sliding window of width n moving over the piece. We start with a window of width 2 after which we enlarge it to 3, etc. We stop at a width of 7 notes, since non-convex n-tone sets with $n > 7$ rarely occur for a major key. Furthermore, for $n > 7$, the computation gets highly intensive since all possible compositions (which is a number between 2^n and 3^n) should be checked. For each non-convex set a vertical bar is plotted at the position of the notes in the piece that it affects. The width of the bar is the number of notes in the (non-convex) set. For each n, a sliding window is moving over the piece resulting in a histogram. These histograms belonging to $n = 2$ to 7 are plotted in the same figure such that the result is one histogram presenting all non-convex sets in a piece of music. The music that we tested the model on is from the Well-tempered Clavier of Bach. Data files containing the notes and other information from all Preludes and all Fugues in first book of J. S. Bach's Well-tempered Clavier (BWV 846 - 869) was made available by Meredith (2003). We used these files as input for our program. Of all the information contained in the files, we were only interested in the notes under octave equivalence. This means for example that we judged a C and a C one octave higher as the same note. Furthermore, all information on note length was neglected. Altogether, the only input used by our model are the note names, so no rhythm, meter, note length, key information etc. were involved.

As an example we consider the third prelude from book 1 of the Well-tempered Clavier. The bars in figure 6.1 show the position of the non-convex sets in the piece. The x-axis represents the number of notes, the prelude contains around 800 notes.

The values on the y-axis do not have any musical meaning, but relative to each other they form an indication for the amount of notes in a set. The length of the bars are scaled as $1 - n/8$, with n indicating the number of notes in a set. Therefore, the bars indicating the (non-convex) 3-note sets have a length

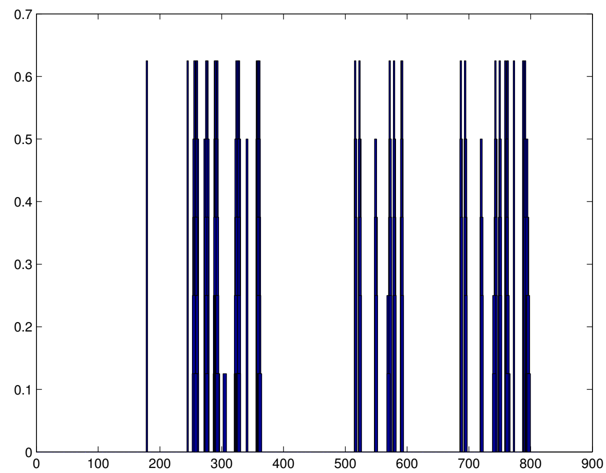


Figure 6.1: Histogram of non convex sets in the third prelude from the Well-tempered clavier. On the x-axis is the number of notes from the piece - the piece ends at note 800, the y-axis is arbitrary.

of $1 - 3/8 = 0.625$, bars indicating 4-note sets have a length of 0.5 etc., until bars indicating 7-note sets have a length of 0.125. Looking at figure 6.1 we see three regions in the music in which a lot of non-convex sets appear. We will now see how these regions relate to the structure of the piece. In "J.S. Bach's Well-Tempered Clavier: In-depth Analysis and Interpretation" (Bruhn 1993) an analysis of the third prelude can be found. The analysis states that from bar 31 to 35 there is a modulation from $A\sharp$ minor to $D\sharp$ minor, from bar 35 to 39 a modulation from $D\sharp$ minor to $G\sharp$ major, from bar 39 to 43 a modulation from $G\sharp$ major to $C\sharp$ major and from bar 43 to 47 a modulation from $C\sharp$ major to $F\sharp$ major. Bars 31 to 46 correspond to notes 250 to 365, which means that this region of modulation is precisely indicated by the first cluster of bars in figure 6.1. Looking to the second cluster of bars in figure 6.1 (notes 520 to 590), one can see that this pattern is repeated a bit later at notes 690 to 760. These two regions correspond to two (similar) pieces in $G\sharp$ having a pedal on the tonic. There are no modulations involved but the notes of the seventh chords are melodically laid out in a way that in forming sets often the fifth is omitted and therefore some sets are non-convex. The last region of bars in figure 6.1 is from note 760 to 800. This represents the last eight (minus the last two) bars of the piece where a melodic line is played in which a lot of chromatic notes are involved. One can not become aware of one specific key until the last two bars where the piece again resolves in $C\sharp$ major. In the regions in between the marked parts (white space in fig 6.1) no modulations are present. In those regions the music is in a certain key, which can vary over time, i.e. there can be (sudden) key changes from one bar to another. This method of looking at non-convex sets is therefore only usable for longer modulation processes. We have seen that some melodic parts with no modulations involved can be wrongly marked as modulation parts. This would suggest that pieces that are harmonically more dense are easier to analyze in this way. As we have seen in tables 6.1, 6.3 and 6.4, from the small and the large sets we can learn the most about the music. We observe that the regions 1 (notes 250 to 365) and 4 (notes 760 to 800) have relatively more 3 note- and 7 note non-convex sets than the 2nd (notes 520 to 590) and 3rd (notes 690 to 760) regions. Therefore regions 1 and 4 have strong(-er) indications of a modulation. Thus, this third prelude serves as an indication that the modulation finding program works well.

Unfortunately, the method did not work well on all the pieces. In figure 6.2 the analysis of the tenth prelude of the Well-tempered Clavier is shown. The key

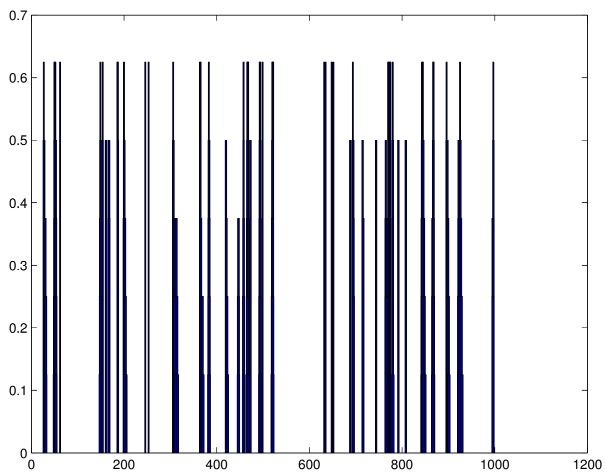


Figure 6.2: Histogram of non convex sets in the tenth prelude from the Well-tempered clavier. On the x-axis is the number of notes from the piece, the piece ends at note 1139.

of the prelude is E minor. From table 6.4 we understand that a piece in a minor key contains far more non-convex sets than a piece in a major key. This makes it difficult to localize modulation on the basis of non-convex sets since there is more 'back-ground noise', that is, non-convex sets which do not indicate a modulation but just refer to a non-convex set as part of the key.

According to the analysis by Bruhn (1993), the tenth prelude can be analyzed as given in table 6.5. The first three modulations (to G major, A minor and E minor) can be recognized in fig. 6.2, although not very clearly. G major is the parallel key of E minor (both have the same number of accidentals at the clef), so a modulation from one to the other does not bring out a lot of extra notes. The

I	1	bars 1-4m	notes 1-98	complete cadence in E minor
	2	bars 4m-9m	notes 98-235	modulation to G major
	3	bars 9m-15d	notes 235-386	modulation to A minor
	4	bars 15m-21d	notes 401-543	modulation back to E minor
	5	bars 21m-23d	notes 557-584	modulation to A minor
II	6	bars 23-26m	notes 584-695	complete cadence in A minor
	7	bars 26m-28d	notes 695-741	modulation back to E minor
	8	bars 28-41	notes 741-1139	confirmation of E minor

Table 6.5: Harmonic analysis of tenth prelude of Well-tempered Clavier by Bruhn (1993). The 'm' refers to middle, the 'd' refers to downbeat.

modulation to A minor in bars 21 to 23 is not visible at all. Although the analysis in table 6.5 states that bars 28 to the end of the piece represent a confirmation of E minor, this does not mean that E minor is the only key that one becomes aware of. Indeed, Bruhn (1993) writes that bars $27-30$ (notes $709-860$) correspond to bars $14-17$ (notes $358-463$), and bars $31-33$ (notes $803-916$) correspond to bars $19-21$ (notes $485-563$). This indicates that some modulation is going on, which is translated into vertical bars in figure 6.2. However, the figure does not give a very good overall indication of the modulations present in the prelude. As may be clear from table 6.4, from the small and the large sets we can learn the most about the music. However, the large sets (of $n = 10, 11$) are in the case of a minor key too large to compute, it requires too much computational time (again, since the number of possible compositions is a number between 2^n and 3^n). For the small two note sets, almost any combination of notes can be represented in a convex set[2] which means that almost no non-convex sets will arise in a piece of music, even if a modulation occurs. Of the three note sets, only 80.91% of the sets within a minor key represents a convex set. This means that quite a substantial part (19.09%) of the non-convex three note sets do not indicate modulations.

Overall, we see that studying non-convex sets can give a rough analysis of the modulations in a piece. The pieces in a major key are easier to analyze than pieces in a minor key, since in the latter some 'background noise' of non-convex sets is present. This analysis method uses only little information of the music (only the note names under octave equivalence) which indicates that the method can still be improved. Furthermore, it can perhaps be integrated in other modulation finding theories to optimize the results.

[2]When the tone space is infinitely big, *all* combinations of two notes can be represented in a convex set. This is because the line of fifths which is represented as a line in the tone space, represents all note names. The nearest neighbor parallel line of the line of fifths is again a line of fifths, only the note names are shifted along the line with the length of a syntonic comma. Therefore there can always be a line drawn from the C on the line of fifths to any other note on the other line of fifths without passing another note, which makes the two-note set convex.

6.2 Pitch spelling

The process of pitch spelling addresses the question which note names should be given to specific pitches. In most computer applications tones are encoded as MIDI pitch numbers which represent the different semi tones. For example, middle C is represented by pitch number 60, the $C\sharp/D\flat$ immediately following middle C is represented by pitch number 61, and so on. This MIDI notation is similar to pitch number notation where the 12 semitones within the octave are indicated by the numbers 0 to 11. Both systems do not distinguish between enharmonically equivalent notes like $C\sharp$ and $D\flat$. However, in tonal music, there is a lot of information in the note names about harmony, melody, scales, and intonation. Therefore, it is very useful to be able to disambiguate the music encoded as MIDI pitch numbers and transcribe it into note names. Pitch spelling is the process that deals with this problem. There has been an increasing interest in pitch spelling algorithms over the last decades, and various algorithms have been proposed (Longuet-Higgins 1987a; Temperley 2001; Meredith 2003, 2006; Cambouropoulos 2003; Chew and Chen 2003, 2005).

As an example of the pitch spelling problem, consider figure 6.3, taken from Piston and DeVoto (1989). The first dyad in the upper staff in every bar is the same, if represented in MIDI numbers. In their harmonic context however, the difference becomes clear due to the function of the chords.

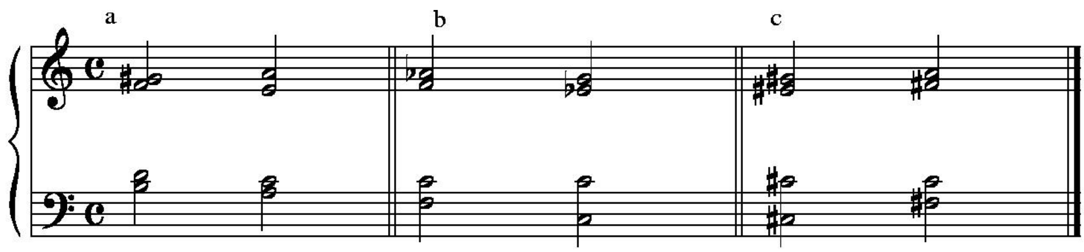


Figure 6.3: Example from Piston and DeVoto (1989).

Most experts seem to agree that the pitch name of a note in a passage of tonal music is primarily a function of 1) the key at the point where the note occurs, and 2) the voice-leading structure of the music in the note's immediate context (Meredith 2006). To understand that the local key plays an important role, imagine a passage in C major. Using the pitch number system in which $C = 0, C\sharp/D\flat = 1$, etc., a pitch number 4 is most likely spelled as an E since this note is part of the scale of C major. Any other possibility to spell the pitch number 4 ($D\sharp\sharp, F\flat$) would not be part of the scale of C major and therefore unlikely to be spelled as such. To understand that voice leading also has an influence on the pitch spelling process, consider figure 6.4 in which a one bar melody is displayed. In case of a semi-tone distance between consecutive notes, in ascending direction the preceding notes are notated with sharps and in descending direction the preceding notes are notated with flats.

Figure 6.4: Voice leading example: ascending and descending chromatic scale.

6.2.1 Review of other models

Longuet-Higgins and Steedman (1971) already touched on the work of pitch spelling in their paper on key finding algorithm. Thereafter Longuet-Higgins (1987a) developed a pitch spelling program which was part of a program that takes a melody as input and generates the phrase structure, articulation, note names and metrical structure. Longuet-Higgins states that only monophonic melodies can serve as input for the program. The spelling algorithm defines the value of sharpness of each note, which is the integer indicating the position on the line of fifths ($C = 0, G = 1$, etc.). The local key is derived by assuming that the first note of a melody is either the tonic or the dominant. Then, appropriate spellings are selected in the way that the distance on the line of fifth between the note to be spelled and the tonic is less than six steps. In this way, diatonic spellings are favored over chromatic spellings. Later, the choice of the tonic is checked again by evaluating a local set of notes to see if an alternative interpretation of the tonic would result in more diatonic intervals among the notes.

Cambouropoulos (2001, 2003) developed a pitch spelling algorithm which uses a shifting overlapping windowing technique. All the pitches in each window are spelled, but only the ones in the middle one-third section of the window are retained. The suggested size of the window is 9 or 12 pitches. The spelling process is based on two principles 1) Notational parsimony (i.e., spell notes making minimum use of accidentals), 2) Interval optimization (i.e., avoid augmented and diminished intervals). Penalty scores are introduced for these two principles, such that the spelling process results in searching through possible spellings for a window and selecting the spelling with the lowest penalty score. The specific penalty values have been selected after trial and error optimization of the test corpus. Cambouropoulos (2003) used a test corpus consisting of 10 sonatas by Mozart and 3 waltzes by Chopin. The best implementation of the algorithm gave 98.8% correctly spelled notes for Mozart (54418 notes) and 95.8% correctly spelled notes for Chopin (4876 notes).

Temperley (2001) proposed a preference rule system for pitch spelling, which he called "A preference rule system for tonal-pitch class labeling". The system consists of three tonal-pitch-class preference rules, or TPRs. The first rule is said to be the most important. TPR1: Prefer to label nearby events so that they are close together on the line of fifths. Temperley states that in most of the cases, this rule is sufficient to ensure the correct spelling of the passages. The second TPR is a voice leading rule, designed to account for the way chromatic notes are usually spelled when they are a half-step apart in pitch height. TPR 3 states to prefer tonal pitch class (TPC) representations which result in good harmonic representations. This harmonic feedback rule says that, when certain factors favor interpreting a group of events in a certain way (for example as a chord), there will be a preference to spell the notes accordingly. The definition of "good harmonic representation" is given in the chapter on preference rules for a theory of harmony. Temperley's tonal-pitch class labeling system is therefore dependent on his theory of harmony, which is in turn dependent on his theory of metrical structure. The implementations of the TPC-labeling system have been largely developed and written by Daniel Sleator and are described in chapters 5 and 6 of Temperley (2001). The system was tested on a 8747 note corpus from the Kostka and Payne (1995) theory workbook and had a correct spelling rate of 98.9%. Reviewing a number of pitch spelling models, Meredith (2006) has shown that a pitch spelling model based on only the first TPR gives equally good and even more robust results when it was tested on a large corpus containing music from 8 baroque and classical composers (Meredith 2006). A consequence of using only TPR1 is that the model does not require temporal information about the notes as input anymore.

Chew and Chen (2002, 2003, 2005) have created a real-time pitch spelling algorithm using the Spiral Array Model which was first proposed by Chew (2000). The spiral array is a spiral configuration of the line of fifths (see section 2.3). In the spiral array, note names are arranged on a helix, such that adjacent note names on the helix are a perfect fifth away from each other, and adjacent note names along the length of the cylinder where the helix is embedded in, are a major third away from each other. Chew and Chen define the center of effect (CE) of a set of notes as the point in the interior of the spiral array that is the convex combination of the pitch positions weighted by their respective durations. The basic principle in Chew and Chen's model is that a pitch should be spelled as the note which is closest (in the spiral array) to the center of effect of the notes in a window preceding the note to be spelled. Chew and Chen then describe a boot-strapping algorithm in which they take into account a local and a global key. In this algorithm the notes in a chunk are first spelled so that they are closest to a global center of effect calculated over a window containing a certain number of chunks preceding the chunk currently being spelled. Then the notes in the chunk are respelled so that they are closest to a weighted average of two centers of effect: a local center of effect calculated from a short window, and a cumulative center of effect constructed from all music preceding the chunk being spelled. The overall results of the spiral array pitch spelling model yields 99.37% (Chew and Chen 2005). However, the test set that was used included only two piano

Sonatas from Beethoven and a piece consisting of variations for violin and piano on a popular Taiwanese folksong. Meredith (2006) and Meredith and Wiggins (2005) ran Chew and Chen's algorithm on a large test corpus containing music of 8 baroque and classical composers, resulting in a percentage of 99.15% correctly spelled notes. Furthermore Meredith (2006) concluded that this algorithm was the least dependent on style of all algorithms tested.

Meredith (2003, 2006) has written a program called $ps13$ which works in two stages. In the first stage, the local sense of key at each point in a passage is represented by the number of times pitch numbers occur within a context surrounding that point. This frequency is used as a measure of likelihood of the pitch number being the tonic at that point. Then the note name implied by the pitch is the note name that lies closest to the tonic on the line of fifths. The strength with which a particular note name is implied is the sum of occurrences of the tonic pitch numbers that imply that note name, within the context surrounding the note. The size of the context is defined by two parameters, K_{pre} and K_{post}, which specify the number of notes preceding and following the note to be spelled. In Stage 2 of Meredith's algorithm, voice-leading is taken into account by correcting those instances in the output of Stage 1 where a neighbor note or passing note is predicted to have the same letter name (for example $B\flat$, B and $B\sharp$ have the same letter name) as either the note preceding it or the note following it. However, it was shown that in some cases Stage 2 failed to correct the neighbor and passing note errors, because of lacking information about voice leading. Meredith showed that removing Stage 2 from the algorithm has beneficial effects of improving time complexity, making the algorithm more robust to temporal variations in data derived from performances, and simplifying the algorithm. The algorithm only consisting of Stage 1, $ps13s1$, with the parameters K_{pre} and K_{post} adjusted to the values 10 and 42 respectively spelled 99.44% notes correctly. The test set Meredith used consisted of pieces by Bach, Beethoven, Corelli, Händel, Haydn, Mozart, Telemann and Vivaldi, in total consisting of 195972 notes. Comparing his algorithm to the models of Longuet-Higgins (1987a), Temperley (2001), Cambouropoulos (2003) and Chew and Chen (2005), his model turned out to perform best (Meredith 2005; Meredith 2006). With the fact that $ps13$ performed better without its Stage 2, Meredith found that the local sense of key is more important than voice-leading considerations when determining note names in tonal music.

6.2.2 Pitch spelling using compactness

We have seen that the major and minor diatonic scales can be found in convex regions in the tone space. Also, all diatonic chords turned out to be convex sets. Since tonal music is usually built from diatonic scales and chords, the convexity property may be used as a tool in a pitch spelling algorithm. For example, the set of pitch numbers $0, 4, 7$ can refer to a variety of possible sets, since the pitch number 0 could refer to $\ldots, B\sharp, C, D\flat\flat \ldots$, the pitch number 4 could refer to

$\ldots, D\sharp\sharp, E, F\flat \ldots$, and the pitch number 7 could refer to $\ldots, F\sharp\sharp, G, A\flat\flat \ldots$. From all these possible sets of note names, only the set C, E, G (and transformations of the set resulting from diminished second transpositions, like $D\flat\flat, F\flat, A\flat\flat$) refers to a diatonic chord in a certain scale *and* it is represented by a convex set. The set $0, 4, 7$ is therefore most likely spelled as C, E, G, and hence the convexity of the set might be a useful tool to find the right pitches. It is not enough to just consider the convexity of the set, since there may exist more convex configurations of a set. We have seen that compactness is closely related to consonance. Therefore, after extracting the convex sets, the most compact of the convex sets is said to represent the right spelling. In the process of searching for convex sets, it became clear that, especially for small sets, many possible convex configurations exist. For these sets, finding the right spelling was rather a process of finding the most compact one, than finding the convex one (because there were so many possibilities). Therefore, the idea arose to neglect the first stage of finding the convex sets, but directly search for the most compact sets. Test runs of two pitch spelling algorithms, one based on convexity-compactness and one based only on compactness, showed that the algorithm based only on compactness performed even better than the algorithm based on convexity-compactness at least up to runs with maximum 5 notes in a set.

For the simple example of spelling the set $0, 4, 7$ given above, compactness applies just as well as convexity. The set $0, 4, 7$ is most likely spelled as C, E, G since this is the most compact configuration of the set. As we already mentioned in section 5.2 compactness should be considered on the lattice constructed from the projection:

$$\phi : 2^p (\frac{3}{2})^q (\frac{5}{4})^r \rightarrow (q, r), \tag{6.1}$$

resulting in the tone space constructed from perfect fifths and major thirds as we have seen before. For convenience, figure 6.5 shows again the resulting tone spaces of frequency ratios, note names and pitch numbers.

	25/18	25/24	25/16	75/64	225/128		
40/27	10/9	5/3	5/4	15/8	45/32	135/128	
32/27	16/9	4/3	1	3/2	9/8	27/16	81/64
256/135	64/45	16/15	8/5	6/5	9/5	27/20	81/80
	256/225	128/75	32/25	48/25	36/25	27/25	

	F#	C#	G#	D#	A#		
G	D	A	E	B	F#	C#	
Eb	Bb	F	C	G	D	A	E
Cb	Gb	Db	Ab	Eb	Bb	F	C
	Ebb	Bbb	Fb	Cb	Gb	Db	

	6	1	8	3	10		
7	2	9	4	11	6	1	
3	10	5	0	7	2	9	4
11	6	1	8	3	10	5	0
	2	9	4	11	6	1	

Figure 6.5: Tone spaces constructed from projection $\phi : 2^p (\frac{3}{2})^q (\frac{5}{4})^r \rightarrow (q, r)$.

The model we will now describe, is based on two very simple rules. When the music is segmented into small sets of notes,

1. Choose the spelling that is represented by the most compact set.

2. Among the sets that are equally compact, the set that is closest in key to the previous set is chosen

As may be clear, in the pitch number tone space there is always more than one set with the same shape and therefore the same compactness. This is illustrated in figure 6.6. The two rules for pitch spelling given above can be summarized in *one* principle, that of compactness, since the second rule elects the set that forms together with the previous set the most compact structure.

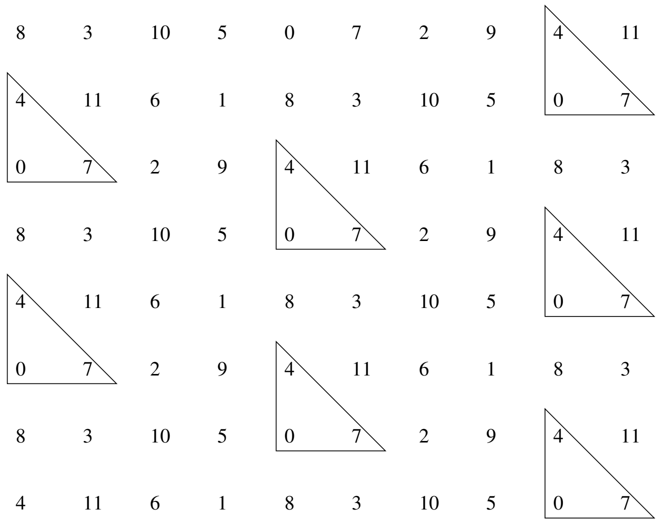


Figure 6.6: Piece of the pitch number tone space which illustrates that there exists more than one set $\{0, 4, 7\}$ with the same compactness.

For the first set of the piece, among the equally compact sets, the set that has the projection on the note name space with the least number of accidentals is chosen. For the sets thereafter, we would want to choose the set which is closest in key (number of accidentals) to the the preceding set. However, if the music is segmented in very small sets, and there is a sudden change of key, this may not work properly. Therefore, the average is calculated between the number of accidentals in the previous set and the sets before that set.

In figure 6.8 an example is given of the pitch spelling process of the first bar from Fugue II from Bach's Well-tempered Clavier book I. This bar is displayed in figure 6.7. The notes of this bar, given in pitch numbers are: $0, 11, 0, 7, 8, 0, 11, 0, 2$. From the most compact sets, the one with the least number of accidentals is chosen, as can be seen from the projection in figure 6.8. This set, C, B, C, G, $A\flat$, C, B, C, D, indeed represents the correct notes from the first bar of the piano sonata.

It can happen that a complete piece is spelled (according to the algorithm) in a different key than the original. For example, a piece written in $C\sharp$ major (with 7 sharps) will be spelled by the algorithm as a piece in $D\flat$ major (with 5 flats) because the latter key contains fewer accidentals. This does not mean that the algorithm wrongly spelled the piece, it may be notated correctly but in a different key. Therefore, we want our pitch spelling algorithm to allow for so-called enharmonic spellings. To this end, the definition we use here (which is also used by Meredith 2003, 2006, Temperley 2001 and others) for a correctly spelled piece of music is: A piece is spelled correctly if every note name assigned by the algorithm is the same interval away from the corresponding note name in the original score. The algorithm therefore generates three spellings. One spelling is directly generated by the algorithm, one spelling is generated from the first by transposing all notes a diminished second up, and one spelling is generated from the first by transposing all notes a diminished second down, such that three enharmonic spellings result. The spelling with the smallest number of errors is then considered to be the correct spelling for the piece of music. In solving this problem of enharmonic spellings, we followed Meredith (2006).

Each MIDI file is segmented in sets each consisting of n notes. If the number of notes the whole musical piece consists of, is not a multiple of n, the last pitches are undetermined. To overcome this problem, after the last set of n pitches, the remainder of pitches form a set (which contains less than n pitches) to be spelled using the same algorithm.

The input to our model are scores encoded in OPND (onset, pitch-name, duration) format (Meredith 2003). Each OPND representation is a set of triples (t,n,d) given the onset Time, the pitch Name and the Duration of a single note or sequence of tied notes in the score. The triples are ordered in a file according to onset time. For a chord, the triples are ordered from the lowest to the highest note in the chord. The note names from the OPND file are used at the end to check whether the pitches are correctly spelled. The actual input of the program are MIDI numbers that are obtained from the note names (Meredith 2003). The compactness algorithm uses only the pitch information of the OPND format; the onset time and the duration of the notes are neglected. Furthermore, the pitch information is transcribed to MIDI numbers which is in turn transcribed to pitch numbers modulo 12. The algorithm spells the notes only on the basis of these pitch numbers, ordered in onset time and in case of equal onset time, and ordered

Figure 6.7: First bar from Fugue II from Bach's Well-tempered Clavier book I.

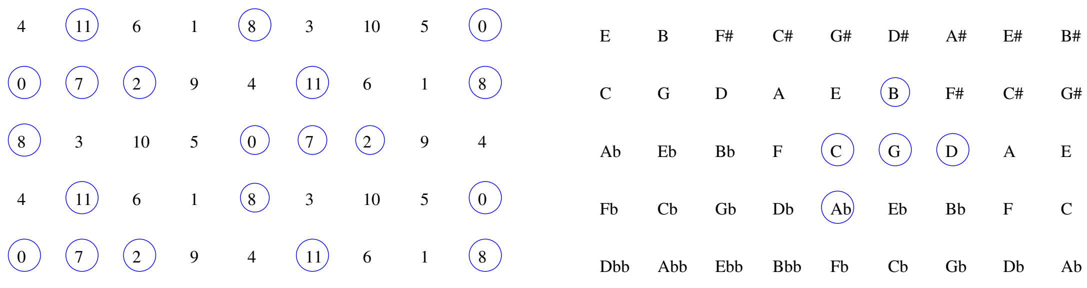


Figure 6.8: Encoding of first bar from Fugue II from Bach's Well-tempered Clavier.

from low to high frequency. As a consequence, the algorithm does not distinguish simultaneously played notes from consecutive notes.

6.2.3 The algorithm

In algorithm 1 the core of the compactness program is explained in pseudo code. In the actual algorithm, note names are encoded as numbers according to the line of fifths, $C = 0$, $G = 1$, etc., since numbers are easier to handle than letters. In line 1 of the code a number of notes is chosen. The music that is to be spelled is segmented into pieces containing this number of notes. In line 2 the key is assigned, the number 0 represents the key of C. With this 'initial key', we want to make sure that the spelling with the least number of sharps and flats will be chosen in the first set of notes. The coordinates of all possible configurations of a set on the pitch number lattice are stored in arrays. For each array representing the coordinates of a set the induced note names are determined. The note names corresponding to the configurations in the pitch number space are found by just projecting the pitch number space onto the note name space (see figure 6.5). In practice, when the coordinates of the configurations are determined, the note names could just be looked up in the note name space, using the same coordinates.

Of every configuration of a set, the notes in terms of fifth line values, are summed (line 16). When this sum is divided by the number of notes in the set (line 19), it gives an indication of the 'tonal center', which is important to be able to spell a set of notes with a tonal center close to the tonal center of the previous set of notes. For example, the notes $C, D, F, C\sharp$ are indicated by the numbers $0, 2, -1, 7$ and result in a value for the 'tonal center' as $(0+2-1+7)/4 = 2\,(= D)$. For every configuration the compactness is calculated as shown in algorithm 2. The compactness of a set of coordinates is calculated by summing the euclidean distances between all points (after removing the points that are equal to other points already in the set). The lower the resulting value, the more compact the configuration is. The most compact configurations are selected by sequentially calculating the compactness thereof and comparing whether the configuration is more compact than or equally compact as the - until then - most compact

```
n=number of notes in the set
Key=0
for all sets of n notes in the piece of music do
    for all possible configurations of a set in the pitch number space do
        Store the coordinates from pitch number space of the notes in array
        M
        for each array M do
            Calculate the value Compactness
            if the value Compactness is smaller than or equal to a value
            MostCompact then
                if the value of Compactness is smaller than a value
                MostCompact then
                    Store the Compactness value in the variable MostCompact
                    clear the array Best
                    Set PreferredKey to initial value
                end
                for each array M do
                    Store the induced note names in array NoteNames
                    Sum the note values and store this in Sum
                end
                Store the indexes of the sets that are most compact in Best
                Calculate TonalCenter as Sum/Number
                if the difference between TonalCenter and Key is smaller
                than value PreferredKey then
                    Store the difference between TonalCenter and Key in
                    PreferredKey
                    Store the index from the sets in array Best corresponding
                    to the set for which the tonal center is closest to Key, in
                    Closest
                end
            end
        end
    end
    Store the average of Key and TonalCenter(Closest) in Key
    for all notes in the set do
        Compare NoteNames(Best(Closest)) with the corresponding note
        name in the original score
    end
    Sum the number of correctly spelled notes and store in
    CorrectSpelling
end
The total percentage of correctly spelled notes equals CorrectSpelling
divided by the total number of notes in the piece of music
```

Algorithm 1: Simplified version of main code of pitch spelling algorithm.

```
input:set M consisting of n notes
Reduce the set M to the same set but with no repetitions of elements
Compute the distances between all (n choose 2) pairs of points in the set
Sum the distances and store in value compactness
output: compactness
```

Algorithm 2: Code for function compactness.

configuration (line 7 to 9). The variable 'MostCompact' is given a high value before going through the for-loop the first time. Then, this value is lowered to the compactness of the most compact set at that moment (line 10). The indices of the configurations having the same compactness (figure 6.6 reminds us that there are always configurations having the same compactness) which are also the most compact at that moment, are stored in the array 'Best' (line 18). If a configuration appears having a higher degree of compactness, the array 'Best' is cleared (line 11) and filled with indexes of the configurations all possessing the new highest degree of compactness. Of the most compact configurations, the 'TonalCenter' of each is compared to the value 'Key' by calculating the difference of these values (line 20). This difference has been given the name 'PreferredKey' and is changed every time a configuration is closer to the 'Key' and set to its initial value (line 12) when a new array 'Best' of most compact configurations is created. The value of 'Key' is changed after every set to provide a value representing a balanced combination of the key of the whole piece and the 'local key' of the previous set (line 27). Of the most compact configurations of a set, the one whose 'TonalCenter' is closest to the 'Key' is selected (line 22). The notes from this configuration are the notes spelled according to the algorithm for one set of notes in the music. To check if this represents the correct notes, each note in this set is compared with the corresponding note in the original score. If a note is correctly spelled, a goodness counter is incremented so that the total of this divided by the total number of notes in the piece of music represents the percentage of correctly spelled notes in this piece (line 28 to 33).

The compactness program has been tested on the preludes and fugues of Bach's Well-tempered Clavier. The fact that more authors have used this test corpus allows us to compare our algorithm with other models. Results are given in table 6.6 for n ranging from 1 to 7. For $n = 1$, the algorithm reduces to rule no. 2 described in the previous section, since the compactness of one single point always equals zero. It is therefore interesting to see that, considering the compactness of only two notes, increases the result already with around 30%. For the best result at $n = 7$, the goodness rates for all preludes and fugues are given in table 6.7.

While algorithm 2 shows that the computational time of compactness is quadratic in the number of notes n (i.e. $\frac{1}{2}n(n-1)$), the number of *possible*

n	percentage correctly spelled notes
1	65.76 %
2	96.57 %
3	96.42 %
4	98.80 %
5	98.58 %
6	98.98 %
7	**99.21** %

Table 6.6: Results for the pitch spelling algorithm based on compactness, as a function of the number of notes n used in the segmentation.

configurations in the Euler lattice for which compactness has to be computed is exponential in n. When using a 9×9 lattice, pitch numbers appear between 6 and 9 times (for example there are 9 locations where the pitch number 0 is situated - when choosing this number in the origin). Therefore, a set of n notes has a minimum of 6^n and a maximum of 9^n configurations. For increasing n, this number becomes high very fast, and slows down the pitch spelling process. Numbers like $n = 20$ are not uncommon to represent one or two bars in music. A first improvement to diminish the computational time of the algorithm has been made by rejecting certain configurations - the ones that are definitely not the most compact - in an early stage. In the algorithm, the compactness of an n-note set is compared with the most compact set up to that moment. If a subset of this n-note set is less compact than the most compact set, then the particular n-note set and also all other sets containing this subset are by definition (see algorithm 2) less compact than the most compact set. Thus, their compactness does not need to be evaluated. This reasoning has been incorporated in the algorithm and increased the speed of the spelling process considerably. However, still, the algorithm requires time exponential in n, therefore $n = 7$ is the practical limit here. From table 6.6 it can be seen that the best performance occurs at $n = 7$, the worst performance is for $n = 3$.

6.2.4 Error analysis

Studying the errors, i.e. the pitches that were not correctly spelled, we can observe a number of problems with the compactness algorithm. A voice leading problem exists due to the fact that the compactness of a set is independent of the order of the notes in the set. For example, the model would always prefer the spelling $C - D\flat$ over the spelling $C - C\sharp$ (in a 2-note set), independent of the order of those two pitches, while this is important for their spelling (see again figure 6.4).

However, the problem that causes most errors has to do with the local spelling character of the model. Contrary to Chew and Chen (2005), who report spelling

no.	prelude		fugue	
	no. of notes	correctness	no. of notes	correctness
1	549	99.45 %	729	99.86 %
2	1091	99.08 %	751	99.47 %
3	810	99.26 %	1408	99.43 %
4	658	99.09 %	1311	99.39 %
5	718	98.61 %	772	100.00 %
6	784	96.94 %	715	98.46 %
7	1411	99.57 %	886	99.44 %
8	681	98.24 %	1378	98.84 %
9	421	98.57 %	732	99.86 %
10	1148	99.39 %	810	99.26 %
11	572	99.48 %	667	99.55 %
12	504	99.01 %	1309	98.24 %
13	402	99.75 %	853	99.88 %
14	604	99.01 %	807	98.88 %
15	607	99.01 %	1690	99.53 %
16	534	99.25 %	747	98.26 %
17	661	100.00 %	883	99.66 %
18	553	99.46 %	798	99.75 %
19	603	98.84 %	1172	99.74 %
20	608	97.86 %	2372	99.20 %
21	632	99.68 %	946	99.47 %
22	775	99.35 %	732	99.18 %
23	417	99.76 %	821	99.76 %
24	720	99.17 %	1792	98.88 %

Table 6.7: Results of pitch spelling algorithm for $n = 7$, for all preludes and fugues from the first book of Bach's Well-tempered Clavier.

errors due to unexpected local key changes, we obtain errors because our algorithm does not take into account enough context. The compactness model adapts quickly to local key changes since the most important part of the algorithm deals with the compactness of the spelled set rather than the context. Examples of wrongly spelled pitches are given in figure 6.9. The figure represents one measure from the sixth prelude of Bach's Well-tempered Clavier which was spelled using the compactness algorithm with $n = 4$. The circled 4-note sets in the figure indicate the three sets that have misspelled notes. The first set $G, G, C\sharp, B\flat$ was spelled incorrectly as $G, G, D\flat, B\flat$ since the latter forms a more compact set in the tone space. In the other two circled sets all $C\sharp$'s are incorrectly spelled as $D\flat$'s as well. However, not all $C\sharp$'s are incorrectly spelled like for example the first $C\sharp$ in the measure. Therefore, if a larger set would have been spelled by taking into account more context, these errors could probably have been avoided.

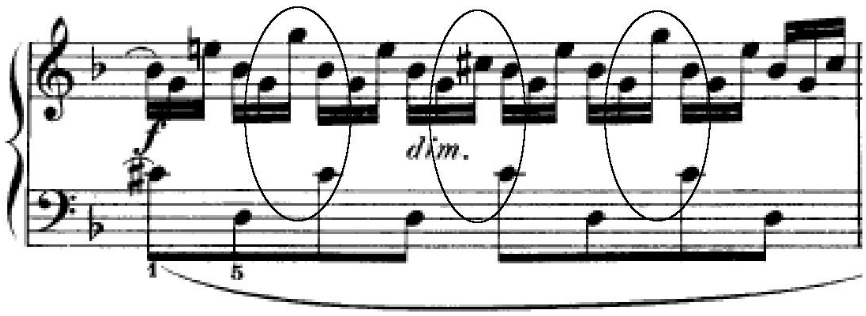


Figure 6.9: Measure no. 16 from prelude VI of Bach's Well-tempered Clavier, showing three 4-note sets from which notes were misspelled.

Analyzing all errors from the sixth prelude (which has been chosen for analysis because of its many errors: only 96.05% was spelled correctly), it turns out that the majority of the errors are due to the misspelling of three semitones as a minor third instead of an augmented second. In the example of figure 6.9, the $D\flat$ was spelled as a minor third above the $B\flat$ instead of an augmented second $C\sharp$ above the $B\flat$.

6.2.5 Evaluation and comparison to other models

With our pitch spelling model we obtained a percentage of 99.21% correctly spelled pitches on the first book of Bach's Well-tempered Clavier. Since Meredith (2003) did a comparative study on pitch spelling algorithms[3] in which he used this corpus, our results can be exactly compared (table 6.8). It may be clear

Algorithm	percentage correct
Cambouropoulos	93.74 %
Longuet-Higgins	99.36 %
Temperley	99.71 %
Meredith	99.74 %
Honingh	99.21 %

Table 6.8: Comparison of pitch spelling models all tested on the 41544 notes of the first book of Bach's Well-tempered Clavier.

[3]Longuet-Higgins's algorithm (1987a) was not designed to be used on polyphonic music. However, with each piece presented as a sequence of MIDI note numbers in appearing order and chords represented from bottom to top (same input representation as we used for our model) his algorithm could still be used. However, results from Meredith (2006) suggest that the algorithm works much better when the music is processed a voice at a time than when it is processed a chord at a time.

that our compactness algorithm does not perform outstandingly compared to the other algorithms. However, the differences are small, and we think that this performance is promising, given that the algorithm is based on only *one* simple principle. We will discuss possible improvements of the algorithm.

Although the algorithm performs on average better using a higher value for n (for example $n = 5$ gives an overall better result than $n = 2$), this does not mean that for every single prelude and fugue this is the case. For example, for prelude number 1 the number of correctly spelled notes using $n = 2$ is 99.82%, in contrast to the number of correctly spelled using $n = 5$ which equals 99.27% If we could find out to what feature this is related, we could design the algorithm such that n is variable, to obtain improved results. This investigation belongs to our plan for future research.

In a way, our model is congenial to Chew and Chen's model (2005). The algorithms are both based on geometrical models of tonal pitch relations. Another similarity with Chew and Chen's model (2005), is that our model uses only present and past information. An advantage hereof is that these models can be integrated into real-time systems of pitch perception. Different from our algorithm, Chew and Chen use metrical information by dividing the music in equal time slices and compute the center of effect with use of the durations of the notes. Our model works without this information. Our algorithm does not take into account onset time so that there is no difference between a (harmonic) chord and an arpeggio of the same chord. We believe the model could be improved by taking this into account. However, neither Cambouropoulos (2003) or Meredith (2006) have used such information. Furthermore, in our model the piece is divided into chunks. Since Meredith and Wiggins (2005) found that the windowing scheme of Chew and Chen's pitch spelling algorithm was critical for high note accuracy, changing our windowing scheme could possibly result in an improved performance.

In our algorithm, the 'key' is represented as a point on the line of fifths instead of a point in the Euler-lattice. Although the key condition (choose the set that is closest to the previous key) does not have that much influence on the algorithm as the compactness condition (choose the most compact set) because the latter condition is applied first and is the most restrictive, it may still give an improvement in the algorithm if the key is represented as a point on the Euler lattice.

As already mentioned by Meredith (2003), the test-corpus of Bach's Well-tempered Clavier cannot be considered to represent a wide variety of musical styles and genres. To be able to show that our compactness algorithm presents a robust pitch spelling algorithm over a variety of musical pieces, the test corpus needs to be enlarged with other music to represent a balanced corpus of a variety of composers within the baroque and classical music.

It was already noted by Longuet-Higgins and Steedman (1971) that the placing of the notes on the Euler lattices "indicates how the subject would have to be played in just intonation". It may be clear that this is an advantage of the compactness model over all the other pitch spelling models described above.

Another question is about the statistical significance of the results. From table 6.7, the standard deviation can be calculated which can tell something about the dispersion of the results. If every prelude and fugue is used as one data point, the mean correctness is 99.20% (which is indeed close to the overall correctness 99.21%), and the standard deviation $\sigma = 0.60$. Of course, the preludes and fugues are not at all of equal length so this standard deviation can only serve as a rough indication. Meredith (2005, 2006) discussed the problem of calcultating the statistical significance of the difference between the results of pitch spelling algorithms. Meredith (2006) explains that the matched-sampled t-test (Howell 1982) is an appropriate test for measuring the significance of the difference between the spelling accuracies achieved by two algorithms over the same corpus. This test involves calculating the mean difference between pairs of results using two spelling algorithms. This mean value should then be divided by the standard error in the mean, resulting in the value for t. However, Meredith (2006) has concluded that "in general, the p value returned by the t-test depends quite heavily on the way that the test corpus is partitioned and that there is no strong a priori reason for choosing one partition over any other". For this reason this statistical analysis has not been incorporated here.

Meredith (2006, 2005) and Meredith and Wiggins (2005) gave an overview of the best performing pitch spelling algorithms of the last decades and concluded that most algorithms use the line of fifths to find the correct spelling of the pitches. One of the few algorithms that do not use this is Chew and Chen's (2002, 2003, 2005) spiral array model. However, Meredith investigated the spiral array model and implemented it with the line of fifths instead of the spiral array and claims that it made no difference in performance. He raises the question: "Would it be possible to improve on existing algorithms by using some pitch space other than the line of fifths?". In our compactness model the two-dimensional Euler-lattice is used to find the correct spelling of pitches. To be sure this model performs better than the line of fifths, we implemented the line of fifths instead of the lattice in a second algorithm, leaving the rest of the code the same. Searching for compactness in a two dimensional space translated naturally to compactness on the one dimensional line of fifths. For n is 2 to 7, the pitch spelling algorithm using the line of fifths was used to calculate the percentages of correctly spelled notes. The resulting percentages were approximately 1% below the percentages from table 6.6 for each n. Although this does not seem to be a huge difference at first sight, an increased performance of 10 correctly spelled notes in a piece of 1000 notes makes a good improvement. A matched-paired t-test was used to calculated the significance of this difference. To perform the test, the corpus, consisting of 41544 notes, was divided into 8 approximately equal sets. The correctness for the algorithm using the tone-space applied to set 1 was compared with the correctness for the algorithm using the line of fifths, and the difference between the percentages was notated. This was done for all sets, and the mean of the differences could be calcultated. With this, the value for t was calcultated as the quotient of the mean difference and the standard error in the mean, and resulted in a value of $t = 3.04$. The degree of freedom was 8, since the corpus was divided in 8 sets. These values gave rise to a chance level of $p < 0.05$. This means that the difference of 1% is not likely to be a chance finding; we would expect such a result to occur by chance only 5% of the time. Thus we say that the obtained difference of 1% is significantly different from zero. This statistical analysis has only been carried out for $n = 4$, but since the difference between the results was approximately 1% for every n, we expect the significance finding to hold for every n. Therefore, we want to make the conjecture that using the two-dimensional Euler lattice instead of the line of fifths is preferred in the process of pitch spelling. However, as mentioned before, the p-value returned by the t-test depends on the partitioning of the corpus. As a consequence, more investigation on the difference between the tone-space and the line of fifths in pitch spelling algorithms, is desirable.

It would be very interesting to compare the tone space we have used to Chew's spiral array in the case of pitch spelling. The spiral array is calibrated so that spatial proximity corresponds to perceived relations among the represented entities (Chew and Chen 2005). This is comparable with compactness representing a measure of consonance in the Euler-lattice. However, one of the biggest differences between those tone spaces in the light of pitch spelling, is that the Euler-lattice can represent frequency ratios as well. The spiral array is a spiral configuration of the line of fifths. As a consequence there is only one representative of every note name contrary to the Euler-lattice, in which there exist more than one of every note name, corresponding to frequency ratios that differ a number of syntonic commas. It would be interesting to examine whether the extra information of the implied frequency ratios yield improved results in the case of pitch spelling.

Chapter 7

Concluding remarks

In this thesis, we have discussed several criteria to derive a scale. These criteria can operate as models explaining the existence of some scales, or they may serve as an evaluation of certain existing scales. Furthermore, scales resulting from these criteria can be interpreted as suggestions for new scales that have not been explored until now. The criteria we have focused on in this thesis have been concerned with equal temperament and well-formedness.

We have considered equal temperament systems, where also microtonal systems with a division higher than 12 tones to the octave were addressed. There are several ways to evaluate the 'goodness' of an n-tone equal tempered system, and we have addressed two of them. The first condition was that an equal temperament should approximate a number of ratios from just intonation as well as possible. The second requirement had to do with the application of equal temperament to a notational system where Western harmony is incorporated. Together, these conditions have resulted in n-tone equal tempered systems where n equals $12, 19, 31, 41$ or 53. Equal tempered systems of these sizes have indeed been found in music theory or practice.

The well-formedness approach to scales addresses the question of the quality and origin of scales from another viewpoint. We have focused on the notion of mathematical beauty or geometrical well-formedness applied to tonal pitch structures, that can possibly serve as the principled basis for tonal music. It has turned out that there is a highly persistent principle holding for pitch structures like scales (also non Western scales), diatonic chords and harmonic reductions: if presented in the tone space described in chapter 2 they form compact and convex or star-convex shapes. We have explained convexity and star-convexity in terms of consonance, such that in a convex musical item the consonance is optimized. For the star-convex scales it had turned out that consonance is optimized according to the tonic of that scale. The property of compactness was not quantified and used as such until chapter 5.

In answer to the questions above, the convexity model can help to explain the existence of certain scales, and furthermore it can serve as a (boolean) evaluation function. Contrary to other theories of well-formedness, the convexity model does not immediately predict tone systems of a specific size. Although we found that most tone systems form a convex set in the tone space, not every convex set in the tone space presents a possible tone system. Still, convexity may help to make predictions about suitable n-tone systems when this condition is combined with others.

Equal tempered scales can also be studied with respect to the property of convexity. However, the property of convexity in the pitch number space is somewhat less interesting than in the frequency ratio space, since the finite number of pitch numbers in the space give rise to a high chance to obtain a convex set. Moreover, scales that are constructed from all the notes of the equal tempered system in which they are embedded (like the chromatic 12-tone scale in ET, and the above mentioned $19, 31, 41$ and 53 tone scales), form necessarily a convex region since they contain the whole (toroidal) pitch number space.

The tone lattices as displayed in figure 7.1 have been central to this thesis, and the projections from one space to the other have formed the basis of two problems addressed in this thesis. Pitch spelling is the problem that is concerned

			9	1	5	9
		2	6	10	2	6
	7	11	3	7	11	3
0	4	8	0	4	8	0
9	1	5	9	1	5	
6	10	2	6	10		
3	7	11	3			

			Bbb	Db	F	A
		Ebb	Gb	Bb	D	F#
	Abb	Cb	Eb	G	B	D#
Dbb	Fb	Ab	C	E	G#	B#
Bbb	Db	F	A	C#	E#	
Gb	Bb	D	F#	A#		
Eb	G	B	D#			

			216/125	27/25	27/20	27/16
		144/125	36/25	9/5	9/8	45/32
	192/125	48/25	6/5	3/2	15/8	75/64
128/125	32/25	8/5	1	5/4	25/16	125/64
128/75	16/15	4/3	5/3	25/24	125/96	
64/45	16/9	10/9	25/18	125/72		
32/27	40/27	50/27	125/108			

MIDI $\Longrightarrow$ note names $\Longrightarrow$ preferred intonation

Figure 7.1: Two projections that are concerned with the problems: pitch spelling and intonation finding.

with attaching the right note name to a MIDI number. In other words, pitch spelling deals with the projection from the tone space of pitch numbers to the tone space of note names. The problem with intonation, and in this case intonation of chords in isolation, is how to interpret (as which frequency ratio) a note name: this problem is concerned with the projection from the tone space of note names to the tone space of frequency ratios. Both problems were dealt with by lifting the convex or the most compact set to the richer space along the specific projection. It has been shown that using compactness instead of convexity gave better results in both cases. For the preferred intonation of chords, we have seen that the most compact set represents the most consonant set in around 90 percent of the cases, depending on the number of notes in the chord. In the case of pitch spelling, a percentage of 99.21% correctly spelled notes was obtained by testing our algoritm based on the notion of compactness, on the preludes and fugues of the first book of Bach's Well-tempered Clavier. Furthermore, we have seen that the algorithm performs worse when the line of fifths is used instead of the 2-dimensional tone space. This suggests that other pitch spelling algorithms, most of which using the line of fifths, can possibly be improved by using the 2-dimensional tone space instead.

We conclude by summarizing that the concepts of convexity and compactness have arisen as important principles reflecting a notion of consonance in scales and chords, and have been successfully applied to well-known problems from music research.

Appendix A

Notes on lattices and temperaments

A.1 Isomorphism between P_3 and $\mathbb{Z}^3$

Since $\mathbb{Z}$ is the set forming a group under addition, the set $\mathbb{Z}^3$ can be represented as a three dimensional space of all points (a, b, c), where $a, b, c \in \mathbb{Z}$, and is a group under vector addition with unit element $(0, 0, 0)$. As mentioned the elements of the group $\mathbb{Z}^3$ are 3-tuples (p, q, r). The group operation is vector addition which means

$$(p, q, r) \circ (p', q', r') = (p + p', q + q', r + r') \tag{A.1}$$

The group P_3 is isomorphic with $\mathbb{Z}^3$, this means that there is a one to one correspondence between the elements of the groups. The isomorphism is given by the map ϕ:

$$\phi : (p, q, r) \in \mathbb{Z}^3 \longleftrightarrow (2^p(\frac{3}{2})^q(\frac{5}{4})^r) \in P_3 \tag{A.2}$$

To prove that the two groups are isomorphic to each other we have to show that the map ϕ is a group homomorphism:

$$\phi((p + p', q + q', r + r')) = \phi((p, q, r)) \bullet \phi((p', q', r')) \tag{A.3}$$

(where $\bullet$ is the group operation in P_3), and prove that ϕ is injective:

$$(a, b, c), (d, e, f) \in \mathbb{Z}^3 : \phi((a, b, c)) = \phi((d, e, f)) \Rightarrow (a, b, c) = (d, e, f) \tag{A.4}$$

and surjective:

$$\forall y \in P_3, \exists (a, b, c) \in \mathbb{Z}^3 : \phi((a, b, c)) = y \tag{A.5}$$

First we prove that ϕ is a homomorphism:

$$\begin{aligned} \phi((p + p', q + q', r + r')) &= 2^{p+p'}(\frac{3}{2})^{q+q'}(\frac{5}{4})^{r+r'} = 2^p(\frac{3}{2})^q(\frac{5}{4})^r \bullet 2^{p'}(\frac{3}{2})^{q'}(\frac{5}{4})^{r'} \\ &= \phi((p, q, r)) \bullet \phi((p', q', r')) \end{aligned} \tag{A.6}$$

We prove injectivity of the map ϕ by the knowledge that every element from P_3 can be written as a unique product of the first three primes.

$$\phi((a,b,c)) = \phi((d,e,f)) \Rightarrow 2^a(\frac{3}{2})^b(\frac{5}{4})^c = 2^d(\frac{3}{2})^e(\frac{5}{4})^f \Rightarrow$$
$$2^{a-d}(\frac{3}{2})^{b-e}(\frac{5}{4})^{c-f} = 1 \Rightarrow (a,b,c) = (d,e,f) \tag{A.7}$$

We prove surjectivity of the map ϕ by the definition of an element from P_3. The elements in P_3 are defined as $\{2^p(\frac{3}{2})^q(\frac{5}{4})^r | p,q,r \in \mathbb{Z}\}$ so there is always an element $(p,q,r) \in \mathbb{Z}$ such that $\phi((p,q,r)) = 2^p(\frac{3}{2})^q(\frac{5}{4})^r$.

A.2 Alternative bases of $\mathbb{Z}^2$

The lattice $\mathbb{Z}^2$ is a subgroup and discrete subspace of the vector space $\mathbb{R}^2$. They share the same basis: $e_1 = (1,0)$, $e_2 = (0,1)$. We can choose another basis for $\mathbb{Z}^2$:

$$\{\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} | a,b,c,d \in \mathbb{Z}, \mathrm{Det}\begin{pmatrix} a & c \\ b & d \end{pmatrix} \neq 0\} \tag{A.8}$$

For simplicity we will use

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \tag{A.9}$$

and its inverse

$$A^{-1} = \begin{pmatrix} e & g \\ f & h \end{pmatrix}. \tag{A.10}$$

We want to prove the statement we made in the text: $\{(a,b),(c,d)\}$ is a basis of $\mathbb{Z}^2 \Leftrightarrow \mathrm{Det}(A) = 1$.

First assuming that $\mathrm{Det}(A) = 1$, we know that A^{-1} consists of integer elements[1]: $e,f,g,h \in \mathbb{Z}$. Then,

$$AA^{-1} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} e & g \\ f & h \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{A.11}$$

implies that

$$e\begin{pmatrix} a \\ b \end{pmatrix} + f\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e_1, \tag{A.12}$$
$$g\begin{pmatrix} a \\ b \end{pmatrix} + h\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = e_2. \tag{A.13}$$

[1] Because for an 2×2 matrix $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$: $A^{-1} = \frac{1}{Det(A)}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$.

And if e_1 and e_2 are elements of the space spanned up by (a, b) and (c, d), any element of $\mathbb{Z}^2$ is in that space, so

$$\{\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\} \text{ is a basis of } \mathbb{Z}^2 \tag{A.14}$$

and the first part of the proof is done.

To go the other way around, we assume $\{(a, b), (c, d)|a, b, c, d \in \mathbb{Z}\}$ is a basis of $\mathbb{Z}^2$, therefore:

$$e_1 = e\begin{pmatrix} a \\ b \end{pmatrix} + f\begin{pmatrix} c \\ d \end{pmatrix}, \tag{A.15}$$

$$e_2 = g\begin{pmatrix} a \\ b \end{pmatrix} + h\begin{pmatrix} c \\ d \end{pmatrix}, \tag{A.16}$$

with $e, f, g, h \in \mathbb{Z}$, which is equivalent to

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e & g \\ f & h \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} \tag{A.17}$$

and therefore

$$\mathrm{Det}(A^{-1})\mathrm{Det}(A) = 1. \tag{A.18}$$

Since all elements of A^{-1} and A are elements of $\mathbb{Z}$, and therefore $\mathrm{Det}(A^{-1})$ and $\mathrm{Det}(A)$ should both be elements of $\mathbb{Z}$ there is no other possibility for $\mathrm{Det}(A)$ then to be equal to 1 or -1.

$$\mathrm{Det}(A) = \pm 1 \tag{A.19}$$

See Regener (1973, ch. 8), for an alternative proof. For more details on linear algebra, see for example (Lang 2002).

A.3 Generating fifth condition

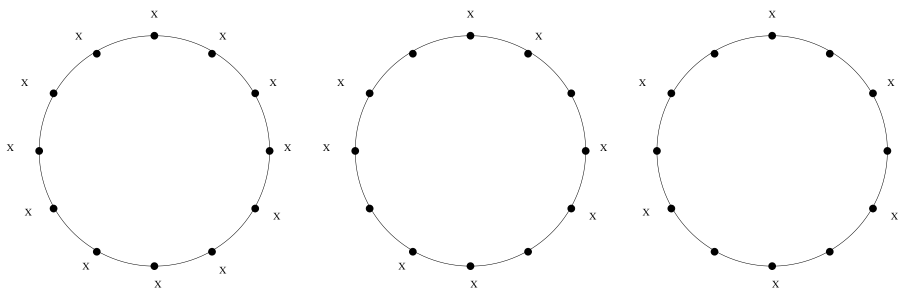
When the n-tone equal tempered system does not have generators among $m_{\frac{3}{2}}, m_{\frac{5}{4}}$ or $m_{\frac{6}{5}}$, the n-tone temperament can be simplified to an n'-tone temperament (section 3.3), such that

$$n = k \cdot n' \qquad k \in \mathbb{N}. \tag{A.20}$$

This can be understood as follows. If the note names are distributed by m where m satisfies the condition

$$\mathrm{GCD}[m, n] = 1, \tag{A.21}$$

a distribution like figure A.1a arises in which every unit can be identified with a note name. If equation A.21 is not true, one may wonder if a distribution like figure A.1b could arise, such that the note names are unequally distributed? Looking at figure A.1b, apparently one could go from unit number t to unit



(a) (b) (c)

Figure A.1: Several distributions of note names (here in a division of $n = 12$), whereby the X's represent possible note names. Distribution b is not possible, see text for explanation.

number $t + 1$ ($t \in \mathbb{Z}$), by adding $x \cdot m \bmod n$, $x \in \mathbb{Z}$. In the same way, one should be able to go from $t + 1$ to $t + 2$. Therefore, an unequal distribution such as A.1b is never possible. If equation A.21 does not hold, (equal) distributions like for example figure A.1c result, which is reducible to an n'-tone temperament (eq. A.20). A temperament system that is generated by the fifth, such that all their notes can be arranged in a continuous series of equal fifths is defined by Bosanquet (1874a, 1874b) to be a regular system.

Samenvatting

Het onderzoek beschreven in dit proefschrift concentreert zich rond de oorsprong van toonstructuren zoals toonladders of akkoorden in muziek. Het is vaak onduidelijk hoe dit soort toonstructuren zich hebben ontwikkeld en waar ze vandaan komen. We kunnen denken aan de volgende vragen. Waarom heeft de Westerse majeur toonladder (do, re, mi, fa, sol, la, si) 7 tonen, en zijn het er niet 6 of 8 of een ander aantal? En waarom bestaat de Japanse pentatonische toonladder uit 5 noten? Met andere woorden, zijn deze getallen willekeurig ontstaan uit verschillende culturen, of zijn deze getallen gerelateerd en wellicht ontstaan uit één en dezelfde oorsprong? Er is veel onderzoek dat de laatste visie ondersteunt. Dit onderzoek kan verdeeld worden in verschillende gebieden. Zo is er bijvoorbeeld het onderzoeksgebied evolutionaire musicologie waarin onderzoek wordt gedaan naar de evolutie van muziek vanuit biologisch en cultureel oogpunt. Verder zijn er studies die suggereren dat de toonladders uit verschillende culturen samenhangen met de instrumenten waarop gespeeld wordt. Ook zou gelijkzwevende stemming (een term die hieronder zal worden uitgelegd) een rol kunnen spelen. Bepaalde gelijkzwevende toonladders zijn gewild om verschillende redenen, bijvoorbeeld omdat ze de reine stemming goed benaderen en omdat ze tegelijkertijd de mogelijkheid hebben om te moduleren. Als laatste noem ik hier dan nog het onderzoek op het gebied van de zogenaamde "wel-gevormde" toonladders. Een toonladder kan wel-gevormd worden genoemd om verschillende redenen bijvoorbeeld omdat hij een symmetrische vorm heeft wanneer hij wordt weergegeven op een tonenrooster of kwintencirkel.

Dit proefschrift focust zich op de laatste twee onderzoeksgebieden (gelijkzwevende stemming en wel-gevormdheid) om een gezamelijke oorsprong van toonladders mogelijk te kunnen verklaren. Naast een gezamelijke oorsprong kunnen deze studies ook dienen als evaluatie van bestaande toonladders (zijn sommige wellicht beter dan andere, of meer geschikt voor een bepaald doel?). Tenslotte kunnen toonladders die voortgebracht worden door de gevormde theorieën, dienen als suggesties voor nieuwe toonladders, die voor muziektheoretici, componisten en wetenschappers interessant zijn om te bestuderen. In dit proefschrift is een gedeelte gewijd aan de studie en evaluatie van gelijkzwevende toonladders. Verder is er een nieuwe notie van wel-gevormdheid geïntroduceerd, waarvan in de laatste twee hoofdstukken toepassingen worden besproken.

Gelijkzwevende stemming

Sinds Pythagoras is reeds bekend dat een interval waarbij de verhouding van de frequenties is gegeven door 2:1 een rein (zuiver) interval oplevert: het octaaf. De kwint met de verhouding 3:2 is eveneens een rein interval. Deze (en meer) reine intervallen blijken onverenigbaar te zijn in een muziekinstrument. Als je bijvoorbeeld boven iedere toon op een piano een reine kwint of octaaf wil kunnen spelen, zouden er een oneindig aantal toetsen nodig zijn. Als oplossing van dit probleem is in de 19^e eeuw de gelijkzwevende stemming ingevoerd, waarbij het octaaf verdeeld wordt in 12 gelijke delen. In deze stemming worden bepaalde intervallen uit de reine stemming goed benaderd. Veel onderzoekers, muziektheoretici en componisten hebben zich daarna afgevraagd of het ook mogelijk is het octaaf in een ander aantal dan 12 gelijke stukken te verdelen, waardoor mogelijk meer intervallen uit de reine stemming benaderd worden, of sommige intervallen wellicht beter benaderd worden. Er zijn veel studies gedaan naar een n-toons gelijkzwevende stemming of toonladder, waarbij geprobeerd werd n zo optimaal mogelijk te kiezen. De vraag is nu, wat is optimaal? Reine stemming beschrijft een oneindig aantal intervallen. Welke van deze intervallen moeten benaderd worden in een (eindige) gelijkzwevende stemming? Om tot een optimale keuze van n te komen moet dus een set van intervallen uit de reine stemming gekozen worden die benaderd dient te worden. De volgende vraag is: binnen deze set van intervallen, welk interval moet het beste benaderd worden, en welke daarna, en daarna? Of zijn alle intervallen even belangrijk? In deze studie heb ik een poging gedaan bovenstaande vragen te formaliseren om zo tot een model te komen dat de optimale waarden voor n voorspelt. De gevonden waarden voor n zijn: 12, 15, 19, 27, 31, 34, 41, 46, 53. Inderdaad blijkt de 12 toons gelijkzwevende stemming, degene die tegenwoordig gebruikt wordt, een goede stemming te zijn volgens dit model. Een aantal van de andere stemmingen is ook (in mindere mate dan 12) gebruikt en onderzocht.

Als de resulterende gelijkzwevende stemmingen gebruikt worden om Westerse muziek mee te spelen, dient deze stemming wel consistent zijn met betrekking tot het Westerse notatiesysteem. Hiermee wordt bedoeld dat een element uit de gelijkzwevende toonladder wel naar meerdere nootnamen (zoals $A, C\sharp$) mag verwijzen, maar dat een nootnaam niet naar meerdere elementen in de toonladders mag verwijzen. Als dit laatste wel het geval zou zijn, zou het bijvoorbeeld niet duidelijk zijn welke toets op een piano in te drukken wanneer iemand je vraagt om een A te spelen. Het gegeven dat een element naar meerdere noten kan verwijzen wordt enharmonische equivalentie genoemd. Bijvoorbeeld, op een (12-toons gelijkzwevende) piano verwijst de toets die naar de $C\sharp$ verwijst, ook naar de $D\flat$. Deze voorwaarden vormen restricties op het aantal mogelijkheden voor n, in een n-toons gelijkzwevende stemming. Dit betekent dat in sommige n-toons stemmingen niet gespeeld kan worden binnen het Westerse muziek-notatiesysteem. Om toch in deze 'verboden' n-toons gelijkzwevende stemmingen te kunnen spelen gegeven bovenstaande regels, zou een ander notatiesysteem gebruikt moeten worden. Hoofdstuk 3 van dit proefschrift gaat in op deze vragen en maakt een voorspelling van de mogelijkheden van waarden voor n aan de hand van de opgelegde restricties door het notatiesysteem. Gecombineerd met de bovenstaande voorspelde waarden voor n (die verkregen waren door goede benadering van reine intervallen), voorspelt het Westerse notatiesysteem, dat systemen met n gelijk aan 12, 19 of 31 goede keuzes zouden zijn. Instrumenten in deze stemming zijn inderdaad vervaardigd.

Wel-gevormdheid

In dit onderdeel van deze studie is gefocust op toonladders en akkoorden in reine stemming. De centrale vraag is hier: wanneer noem je een set tonen een toonladder of akkoord en wat maakt een *goede* toonladder of akkoord? Er bestaat (tot op heden) geen eenduidig antwoord op deze vraag, en daarom beschouwen we een groot aantal toonladders in een toonruimte (figuur 7.1a) en kijken we naar de overeenkomsten. Het blijkt dat vrijwel alle toonladders een convexe vorm beschrijven in deze ruimte. Een convexe vorm is een vorm zonder inhammen of gaten (bijvoorbeeld een cirkel, vierkant of ovaal hebben een convexe vorm, maar een ster of donut hebben geen convexe vorm). Voor Westerse akkoorden geldt hetzelfde, alle laddereigen akkoorden (akkoorden uit de toonladder) hebben een convexe vorm. Dit proefschrift betoogt dat de convexiteit van toonladders en akkoorden te maken heeft met consonantie. Hoe meer de tonen met elkaar in verbinding staan (dus zonder inhammen of gaten tussen twee tonen), hoe makkelijker je van de ene toon naar de andere kunt gaan via consonante intervallen. Hiermee is nu ook een evaluatiemodel voor toonladders gemaakt: is de toonladder convex, dan noemen we hem wel-gevormd. Convexiteit blijkt onafhankelijk te zijn van de gekozen basis van de toonruimte, wat deze eigenschap nog specialer maakt: het is geen artefact van de ruimte. Doordat convexiteit is aangetoond voor een groot aantal toonladders waaronder ook niet Westerse toonladders, is dit een goede aanwijzing dat convexiteit unificerende eigenschappen van toonstructuren weer kan geven. Een eigenschap die verwant is aan convexiteit is compactheid: de mate waarin de elementen van een toonstructuur dicht bij elkaar zitten in de toonruimte. In tegenstelling tot convexiteit is compactheid wel afhankelijk van de gekozen basis van de toonruimte. Echter, het blijkt dat als de basis gekozen wordt die de meest consonante intervallen projecteert op de kleinste afstanden in de ruimte, de compactheid eveneens geïnterpreteerd kan worden als een maat van consonantie: hoe compacter de set noten, hoe consonanter. We kunnen nu kijken naar toepassingen van de eigenschappen convexiteit en compactheid.

Toepassing 1: juiste intonatie van akkoorden

Als we praten over akkoorden, is dat meestal in termen van nootnamen of een aanduiding van waar een akkoord in een toonladder zit. We kunnen bijvoorbeeld spreken over het tonica akkoord, het akkoord dat op de grondtoon van de toonladder staat. Of we kunnen het hebben over het dominant septiem akkoord, het akkoord dat in de toonladder van C, de noten G, B, D, F beschrijft. Zelden echter, hebben we het over akkoorden in termen van frequentie-verhoudingen (zoals het akkoord $1, \frac{5}{4}, \frac{3}{2}$) wanneer we een akkoord in een stuk muziek aanduiden. Dit komt omdat voor de meeste akkoorden het niet volledig duidelijk is hoe ze geïntoneerd moeten worden. Natuurlijk bestaan hier veel meningen over, maar er is geen eenduidige theorie die iedereen volgt. Er vanuit gaande dat een akkoord zo consonant mogelijk moet klinken, kunnen we de mate van convexiteit en compactheid gebruiken om te ontdekken welke intonatie (welke frequentie-verhoudingen) geprefereerd wordt voor een aantal akkoorden. Als evaluatiemethode gebruiken we een andere bestaande maat voor consonantie, de functie voorgesteld door Euler. Het blijkt dat compactheid een betere indicator is voor consonantie van akkoorden dan convexiteit.

Toepassing 2: juiste notatie van tonen

In veel computer toepassingen worden tonen gecodeerd als MIDI getal. In het MIDI systeem is de centrale C gecodeerd als het getal 60; de toon die een halve toon hoger is ($C\sharp/D\flat$) als 61 en zo verder. Deze MIDI notatie is analoog aan de 12-toons gelijkzwevende stemming. Beide maken geen onderscheid tussen enharmonisch equivalente noten zoals de $C\sharp$ en de $D\flat$. Echter, juist deze nootnamen bevatten veel informatie over bijvoorbeeld de toonsoort van een stuk, de harmonie, melodie en intonatie, en zijn dus heel belangrijk voor een muzikant om te weten. Om deze reden is het nuttig als er een model zou bestaan, die MIDI getallen in nootnamen zou omzetten. Dit is lastig, want de ene keer representeert een bepaald MIDI getal bijvoorbeeld een $A\sharp$, maar een andere keer representeert datzelfde MIDI getal een $B\flat$, afhankelijk van de muzikale context. In de literatuur zijn reeds een aantal modellen voorgesteld die proberen de noten juist te 'spellen', gegeven een muziekstuk in MIDI notatie. Geen van de voorgestelde modellen werkt voor 100 procent goed, wat wil zeggen dat geen van deze modellen alle noten van alle ingegeven muziekstukken goed codeert. De 'goede' codering wordt bij dit probleem gegeven door de notatie van de componist van het stuk. In dit proefschrift presenteer ik een nieuw model voor het juist 'spellen' van noten, gebaseerd op de notie van compactheid. De toonsoort van een stuk draagt voor een groot gedeelte bij aan de muzikale context die ervoor zorgt dat een noot op een bepaalde manier gespeld wordt. Het blijkt dat, door de meest compacte vorm van een set noten te kiezen, deze noten zich vaak binnen één toonsoort bevinden wat er meestal voor zorgt dat dit de juist spelling van deze set noten weergeeft. In het resulterende compactheids-model kan het aantal noten in zo'n set gevarieerd worden, en het blijkt dat hoe meer noten in de set zitten, hoe beter het model werkt. Het compactheids-model is getest op alle preludes en fuga's uit het Wohltemperierte Klavier van J.S. Bach dat in totaal 41544 noten bevat. Een score van $99,21$ procent wordt bereikt als de stukken verdeeld worden in sets van 7 noten. Dit betekent dat $99,21$ procent van alle noten goed gespeld wordt met dit model. Hoewel deze score vergelijkbaar is met die van andere modellen bekend uit de literatuur, is het bijzonder dat een model dat slechts gebaseerd is op één principe zulke goede resultaten kan geven.

Tenslotte

De bestudeerde problemen samenvattend, kunnen deze gezien worden als projecties tussen verschillende aanduidingen voor tonen. In de studie over de gelijkzwevende stemming bijvoorbeeld, hebben we een projectie gemaakt van de frequentie-verhoudingen naar de elementen van de gelijkzwevende stemming en van de nootnamen naar de elementen van de gelijkzwevende stemming. Daarna, bij de problematiek rond de juiste stemming van akkoorden, hebben we ons bezig gehouden met hoe de nootnamen juist te projecteren op de frequentie-verhoudingen. Tenslotte draait het probleem van de juiste notatie van noten om een geschikte projectie van de elementen van gelijkzwevende stemming (of MIDI) naar de nootnamen. We hebben een aantal algemene regelmatigheden gevonden in toonstructuren, op basis waarvan de twee laatst genoemde projecties tot stand zijn gekomen.

Terugkomend op de vragen die in het begin gesteld zijn, kan gezegd worden dat een aantal aspecten mogelijk heeft bijgedragen aan het ontstaan van verschillende toonladders. De n-toons gelijkzwevende toonladders die theoretisch gevonden werden door het zoeken naar een goede benadering van de reine stemming en een geschikte notatie, zijn tevens gevonden in de praktijk. Dit ondersteunt de aanname dat 'reine stemming' en 'geschikte notatie' onderliggende eisen zijn geweest voor het ontstaan van deze toonladders. Verder is convexiteit gevonden als overkoepelende eigenschap van een groot aantal reine toonladders. Enerzijds suggereert dit dat het principe van convexiteit een onderliggend principe geweest kan zijn dat een rol heeft gespeeld bij het ontstaan van toonladders. Anderzijds kan convexiteit gebruikt worden als evaluatiemodel zoals hierboven geschreven. Tenslotte kan de convexiteits-eigenschap gebruikt worden voor het verder exploreren en ontwikkelen van nieuwe toonladders.

References

Agmon, E. (1989). A mathematical model of the diatonic system. *Journal of Music Theory 33*(1), 1–25.

Alperin, J. and R. B. Bell (1995). *Groups and Representations.* Springer-Verlag New York Inc.

Bachem, A. (1950). Tone height and tone chroma as two different pitch qualities. *Acta Pshychologica 7*, 80–88.

Balzano, G. J. (1980). The group theoretical description of 12-fold and microtonal pitch systems. *Computer Music Journal 4*(4), 66–84.

Barbour, J. M. (2004/1951). *Tuning and Temperament, A Historical Survey.* Dover.

Benson, D. (2006). Mathematics and music. book in progress, available at `http://www.maths.abdn.ac.uk/~bensondj/html/maths-music.html`.

Bod, R. (2002). Memory-based models of melodic analysis: Challenging the Gestalt principles. *Journal of New Music Research 31*(1), 27–37.

Boomsliter, P. C. and W. Creel (1961). The long pattern hypothesis in harmony and hearing. *Journal of Music theory 5*(2), 2–31.

Boomsliter, P. C. and W. Creel (1963). Extended reference: An unrecognized dynamic in melody. *Journal of Music Theory 2*(1), 2–22.

Bosanquet, R. (1874a). Temperament; or, the division of the octave[part i]. *Proceedings of the Musical Association, 1st Sess. 1*(1), 4–17.

Bosanquet, R. (1874b). Temperament; or, the division of the octave[part ii]. *Proceedings of the Musical Association, 1st Sess. 1*(1), 112–158.

Bosanquet, R. (1876). *Musical intervals and temperaments.* London: Macmillan and Co. Reprinted with commentary by Rudolph Rasch, Diapason Press, Utrecht, 1986.

Bruhn, S. (1993). *J.S. Bach's Well-Tempered Clavier: In-depth Analysis and Interpretation*. Hong Kong: Mainer International Ltd. Transcription for the Web published 2002-2003.

Burns, E. (1999). Intervals, scales and tuning. In D. Deutsch (Ed.), *Psychology of Music* (second ed.)., Chapter 7, pp. 215–264. Academic Press.

Burns, E. and S. Campbell (1994). Frequency and frequency ratio resolution by possessors of relative and absolute pitch: Example of categorical perception? *Journal of the Acoustical Society of America 96*, 2704–2719.

Burns, E. and W. Ward (1978). Categorical perception-phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America 63*, 456–468.

Cafagna, V. and T. Noll (2003). Algebraic investigations into enharmonic identification and temperament. In G. Di Maio and C. di Lorenzo (Eds.), *Proceedings of the 3rd International Conference Understanding and Creating Music*, Caserta, Italy.

Cambouropoulos, E. (2001). Automatic pitch spelling: From numbers to sharps and flats. In *Proceedings of the VIII Brazilian Symposium on Computer Music*, Fortaleza, Brasil.

Cambouropoulos, E. (2003). Pitch spelling, a computational model. *Music Perception 20*(4), 411–429.

Carey, N. (1998). *Distribution Modulo 1 and Musical Scales*. Ph. D. thesis, University of Rochester.

Carey, N. and D. Clampitt (1989). Aspects of well-formed scales. *Music Theory Spectrum 11*(2), 187–206.

Cariani, P. (2004). A temporal model for pitch multiplicity and tonal consonance. In *Proceedings of the Eighth International Conference on Music Perception and Cognition(ICMPC)*.

Cazden, N. (1980). The definition of consonance and dissonance. *International Review of the Aesthetics and Sociology of Music 11*(2), 123–168.

Chalmers, J. (1989). The nineteen-tone instruments of W. A. Piehl and Tillman Schafer. *Experimental Musical Instruments 4*(6), 5–7.

Chew, E. (2000). *Towards a mathematical model of tonality*. Ph. D. thesis, Operations Research Center, massachusetts Institute of Technology, Cambridge.

Chew, E. (2002). The spiral array: An algorithm for determining key boundaries. In C. Anagnostopoulou, M. Ferrand, and A. Smaill (Eds.), *Music and Artificial Intelligence - Second International Conference, ICMAI 2002*, Edinburgh, Scotland, UK, pp. 18–31.

Chew, E. (2003). Thinking out of the grid and inside the spiral. Technical Report 03-002, IMSC.

Chew, E. (2006). Slicing it all ways: Mathematical models for tonal induction, approximation and segmentation using the spiral array. *INFORMS Journal on Computing 18*(3).

Chew, E. and Y.-C. Chen (2002). Mapping midi to the spiral array: Disambiguating pitch spellings. In H. K. Bhargava and N. Ye (Eds.), *Computational Modeling and Problem Solving in the Networked World, Proceedings of the 8th INFORMS Computer Society Conference, ICS2003*, Volume 21 of *OR/CS Interfaces Series*, pp. 259–275. Kluwer Academic Publishers.

Chew, E. and Y.-C. Chen (2003). Determining context-defining windows: Pitch spelling using the spiral array. In *Proceedings of the 4th International Conference for Music Information Retrieval, ISMIR*.

Chew, E. and Y.-C. Chen (2005). Pitch spelling using the spiral array. *Computer Music Journal 29*(2), 61–76.

Clampitt, D. (1997). *Pairwise Well-Formed Scales: Structural and Transformational Properties*. Ph. D. thesis, State University of New York at Buffalo.

Clough, J., N. Engebretsen, and J. Kochavi (1999). Scales, sets, and interval cycles: A taxonomy. *Music Theory Spectrum 21*(1), 74–104.

Clough, J. and G. Myerson (1985). Variety and multiplicity in diatonic systems. *Journal of Music Theory 29*(2), 249–270.

Cohn, R. (1998). Introduction to neo-Riemannian theory: A survey and a historical perspective. *Journal of Music Theory 42*(2), 167–180.

Dahlhaus, C. and R. O. Gjerdingen (1990). *Studies in the Origin of Harmonic Tonality*. Princeton University Press.

De Klerk, D. (1979). Equal temperament. *Acta Musicologica 51*, 140–150.

Desain, P. and H. Honing (2003). The formation of rhythmic categories and metric priming. *Perception 32*(3), 341–365.

Douthett, J., R. Entringer, and A. Mullhaupt (1992). Musical scale construction: The continued fraction compromise. *Utilitas Mathematica 42*, 97–113.

Dowling, W. J. (1982). Musical scales and phsychophysical scales: their psychological reality. In R. Falk and T. Rice (Eds.), *Cross Cultural Perspectives on Music*. University of Toronto Press.

Drabkin, W. (2005). Grove music online. http://www.grovemusic.com/. Accessed 23 May 2005.

Drobisch, M. W. (1855). Über musikalische Tonbestimmung und Temperatur. *Abhandlungen der Königlich sächsischen Gesellschaften der Wissenschaften zu Leipzig 4*, 3–121.

Eitz, C. A. (1891). *Das mathematisch-reine Tonsystem.* Leipzig.

Erlich, P. (1998). Tuning, tonality, and twenty-two-tone temperament. *Xenharmonikon 17.* revised version 4/2/02.

Erlich, P. (2005). A middle path: Between just intonation and the equal temperaments, part i. *Xenharmonikon 18.*

Euler, L. (1926/1739). Tentamen novae theoriae musicae. In E. Bernoulli et al. (Ed.), *Opera Omnia*, Volume 1 of *III*, Stuttgart. Teubner.

Fleischer (Volk), A. (1998). A lifting criterion for affine endomorphisms of finitely generated Z-modules and its music-theoretic meaning. Master's thesis, Humboldt University in Berlin.

Fokker, A. D. (1945). *Rekenkundige bespiegelingen der muziek.* J. Noorduijn en Zoon N.V.

Fokker, A. D. (1949). *Just Intonation.* The Hague: Martinus Nijhoff.

Fokker, A. D. (1955). Equal temperament and the thirty-one-keyed organ. manuscript. http://www.xs4all.nl/ huygensf/doc/fokkerorg.html.

Fokker, A. D. (1969). Unison vectors and periodicity blocks in the three-dimensional (3-5-7)harmonic lattice of notes. In *Proceedings of Koninklijke Nederlandsche Akademie van Wetenschappen*, Number 3 in B 72.

Fuller, R. (1991). A study of microtonal equal temperaments. *Journal of Music Theory 35*(1/2), 211–237.

Fyk, J. (1995). *Melodic Intonation, Psychoacoustics, and the Violin.* Zielona Góra, Poland: Organon.

Galilei, G. (1638). *Discorsi e dimostrazioni matematiche interno à due nuove scienze attenenti alla mecanica ed i movimenti locali.* Leiden: Elsevier, 1638. Translated by H. Crew and A. de Salvio as *Dialogues concerning Two New Sciences.* New York: McGraw-Hill Book Co., Inc, 1963.

Hall, D. E. (1973). The objective measurement of goodness-of-fit for tunings and temperaments. *Journal of Music Theory 17*(2), 274–290.

Hall, D. E. (1981). The difference between difference tones and rapid beats. *American Journal of Physics 49*(7), 632–636.

Hall, D. E. (1988). Acoustical numerology and lucky equal temperaments. *American Journal of Physics 56*(4), 329–333.

Hartmann, G. C. (1987). A numerical exercise in musical scales. *American Journal of Physics 55*(3), 223–226.

Helmholtz, H. (1954/1863). *On the Sensations of Tone* (second English ed.). Dover.

Honingh, A. K. (2003a). Group theoretic description of just intonation. In G. Buzzanca and et al. (Eds.), *Proceedings of the 3rd International Conference Understanding and Creating Music*, Volume 3, Caserta, Italy.

Honingh, A. K. (2003b). Measures of consonances in a goodness-of-fit model for equal-tempered scales. In *Proceedings ICMC*, Singapore, pp. 435–438.

Honingh, A. K. (2004). Limitations on fixed n-tone equal tempered divisions. In G. Buzzanca and et al. (Eds.), *Proceedings of the 4th International Conference Understanding and Creating Music*, Volume 4, Caserta, Italy.

Honingh, A. K. (2006a). Convexity and compactness as models for the preferred intonation of chords. In *Proceedings of the Ninth International Conference on Music Perception and Cognition (ICMPC 9)*, Bologna, 22-26 August.

Honingh, A. K. (2006b). Pitch spelling using compactness. In *Proceedings of the Ninth International Conference on Music Perception and Cognition (ICMPC 9)*, Bologna, 22-26 August.

Honingh, A. K. (2006c). Scales and chords in the euler-lattice: Convexity as a preferred intonation model. In *Mathematical and Computational Musicology*, Berlin. State Institute of Music Research. in press.

Honingh, A. K. and R. Bod (2004). The notion of convexity in music. In G. Buzzanca and et al. (Eds.), *Proceedings of the 4th International Conference Understanding and Creating Music*, Volume 4, Caserta, Italy.

Honingh, A. K. and R. Bod (2005). Convexity and the well-formedness of musical objects. *Journal of New Music Research 34*(3), 293–303.

Houtgast, T. (1976). Subharmonic pitches of a pure tone at low S/N ratio. *Journal of the Acoustical Society of America 60*, 405–409.

Howell, D. C. (1982). *Statistical Methods for Psychology.* Boston: Duxbury Press.

Husmann, H. (1953). *Vom Wesen der Konsonanz.* Heidelberg.

Huygens, C. (1691). Lettre touchant le cycle harmonique. Letter to the editor of the journal 'Histoire des Ouvrage de Scavans'. Reprinted with English and Dutch translation (ed. Rudolph Rasch), Diapason Press, Utrecht, 1986.

Hyer, B. (2006). 'tonality', grove music online. http://www.grovemusic.com/. Accessed 27 April 2006.

Jeans, J. (1968). *Science and Music.* Dover Publications Inc. New York.

Jones, H. (1990). *Groups, Representations and Physics.* IOP Publishing Ltd.

Kameoka, A. and M. Kuriyagawa (1969a). Consonance theory part I: Consonance of dyads. *Journal of the Acoustical Society of America 45*, 1451–1459.

Kameoka, A. and M. Kuriyagawa (1969b). Consonance theory part II: Consonance of complex tones and its calculation method. *Journal of the Acoustical Society of America 45*, 1460–1469.

Kanno, M. (2003). Thoughts on how to play in tune: Pitch and intonation. *Contemporary Music Review 22*(1/2), 35–52.

Keller, H. (1976). *The Well-Tempered Clavier by Johann Sebastian Bach*. London: Norton. translated by L. Gerdine.

Knopoff, L. and W. Hutchinson (1983). Entropy as a measure of style: The influence of sample length. *Journal of Music Theory 27*, 75–97.

Kopiez, R. (2003). Intonation of harmonic intervals: Adaptability of expert musicians to equal temperament and just intonation. *Music Perception 20*(4), 383–410.

Kostka, S. and D. Payne (1995). *Workbook for Tonal Harmony*. New York: McGraw-Hill.

Krantz, R. J. and J. Douthett (1994). A measure of the reasonableness of equal-tempered musical scales. *Journal of the Acoustical Society of America 95*(6), 3642–3650.

Krantz, R. J. and J. Douthett (2000). Construction and interpretation of equal-tempered scales using frequency ratios, maximally even sets, and P-cycles. *Journal of the Acoustical Society of America 107*(5), 2725–2734.

Krueger, F. (1904). Differenztone und Konsonanz. *Archiv für die gesamte Psychologie 2*.

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. Oxford Psychology Series No. 17. Oxford University Press.

Krumhansl, C. L. and E. Kessler (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review 89*(334-386).

Lang, S. (2002). *Algebra* (Revised third ed.). Springer-Verlag.

Latham, A. (Ed.) (2002). *The Oxford Companion to Music*. Oxford University Press.

Leman, M. (Ed.) (1997). *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology*. Springer-Verlag.

Lerdahl, F. and R. Jackendoff (1983). *A Generative Theory of Tonal Music*. MIT Press.

Lindley, M. (2005). Grove music online. http://www.grovemusic.com/. Accessed 28 June 2005.

Lindley, M. and R. Turner-Smith (1993). *Mathematical Models of Musical Scales: A New Approach*. Verlag für systematische Musikwissenschaft, Bonn.

Locke, S. and I. Kellar (1973). Categorical perception in a non-linguistic mode. *Cortex 9*, 355–368.

Longuet-Higgins, H. C. (1962a). Letter to a musical friend. *Music review 23*, 244–248.

Longuet-Higgins, H. C. (1962b). Second letter to a musical friend. *Music review 23*, 271–280.

Longuet-Higgins, H. C. (1976). The perception of melodies. *Nature 263*, 646–653. Reprinted as Longuet-Higgins 1987.

Longuet-Higgins, H. C. (1987a). The perception of melodies. In H. C. Longuet-Higgins (Ed.), *Mental Processes: Studies in Cognitive Science*, pp. 105–129. London: British Psychological Society/MIT Press. Published earlier as Longuet-Higgins 1976.

Longuet-Higgins, H. C. (1987b). The three dimensions of harmony. In H. C. Longuet-Higgins (Ed.), *Mental Processes: Studies in Cognitive Science*, pp. 59–63. London, England and Cambridge, Mass.: British Psychological Society/MIT Press.

Longuet-Higgins, H. C. and M. Steedman (1987/1971). On interpreting Bach. In H. C. Longuet-Higgins (Ed.), *Mental Processes: Studies in Cognitive Science*, pp. 82–104. British Psychological Society/MIT Press. Published earlier as Longuet-Higgins and Steedman (1971).

Loosen, F. (1993). Intonation of solo violin performance with reference to equally tempered, pythagorean, and just intonations. *Journal of the Acoustical Society of America 93*(1), 525–539.

Mazzola, G. (1990). *Geometrie der Töne - Elemente der Mathematischen Musiktheorie*. Birkhäuser Verlag.

Mazzola, G. (2002). *The Topos of Music: Geometric Logic of Concepts, Theory, and Performance*. Birkhäuser Verlag.

Meredith, D. (2003). Pitch spelling algorithms. In *Proceedings of the Fifth Triennial ESCOM Conference*, Hanover University of Music and Drama, Germany, pp. 204–207.

Meredith, D. (2005). Comparing pitch spelling algorithms on a large corpus of tonal music. In U. K. Wiil (Ed.), *Computer Music Modeling and Retrieval: Second International Symposium, CMMR*, pp. 173–192. Springer, Berlin.

Meredith, D. (2006). *Computing Pitch Names in Tonal Music: A Comparative Analysis of Pitch Spelling Algorithms*. D.Phil. dissertation. Faculty of Music, University of Oxford.

Meredith, D. and G. Wiggins (2005). Comparing pitch spelling algorithms. In *Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005)*, London.

Mithen, S. (2005). *The Singing Neanderthals: The Origins of Music, Language, Mind and Body*. London: Weidenfeld and Nicholson.

Montagu, J. (2002). temperament. In A. Latham (Ed.), *The Oxford Companion to Music*, pp. 1262–1263. Oxford University Press.

Nettl, B. (1956). *Music in Primitive Culture*, Chapter 4: Scale and Melody. Harvard University Press Cambridge.

Noll, T. (1995). *Morphologische Grundlagen der abendländischen Harmonik*. Ph. D. thesis, Technishe Universität Berlin. Published in: M. Boroda (ed.), *Musikometrika* 7, 1997, Bochum: Brockmeyer.

Noll, T. (2001). Geometry of chords. In E. L. Puebla (Ed.), *Electronic Bulletin of the Sociedad Matematica Mexicana*, Volume 1.

Noll, T. (2005). Facts and counterfacts: Mathematical contributions to music-theoretical knowledge. In S. Bab and et. al (Eds.), *Models and Human Reasoning - Bernd Mahr zum 60. Geburtstag*. W&T Verlag, Berlin.

Ono, T. (1990/1987). *An Introduction to Algebraic Number Theory* (Second ed.). Plenum Press. English translation.

Palisca, C. V. and B. C. J. Moore (2006). 'consonance', grove music online. http://www.grovemusic.com/. Accessed 26 April 2006.

Parncutt, R. (1989). *Harmony: A psychoacoustical approach*. Berlin: Springer-Verlag.

Parncutt, R. and H. Strasburger (1994). Applying psychoacoustics in composition: "harmonic"progressions of "nonharmonic"sonorities. *Perspectives of new music 32*(2), 88–129.

Partch, H. (1949/1974). *Genesis of a Music* (second ed.). Da Capo Press.

Piston, W. and M. DeVoto (1989). *Harmony* (revised and expanded ed.). Victor Gollancz Ltd.

Plomp, R. (1964). The ear as frequency analyzer. *Journal of the Acoustical Society of America 36*(9), 1628–1636.

Plomp, R. (1965). Detectibility threshold for combination tones. *Journal of the Acoustical Society of America 37*, 1110–1123.

Plomp, R. (1967). Pitch of complex tones. *Journal of the Acoustical Society of America 41*, 1526–1533.

Plomp, R. and W. J. Levelt (1965). Tonal consonance and critical bandwith. *Journal of the Acoustical Society of America 38*, 548–560.

Plomp, R. and A. M. Mimpen (1968). The ear as frequency analyzer II. *Journal of the Acoustical Society of America 43*(4), 764–767.

Povel, D.-J. (2002). A model for the perception of tonal melodies. In C. Anagnostopoulou, M. Ferrand, and A. Smaill (Eds.), *Music and Artificial Intelligence*. Heidelberg: Springer Verlag.

Rakowski, A. (1990). Intonation variants of musical intervals in isolation and in musical contexts. *Psychology of Music 18*, 60–72.

Rameau, J.-P. (1722). *Traité de l'harmonie réduite à ses principes naturels*. Paris.

Rapoport, P. (1993). The structural relationship of fifths and thirds in equal temperament. *Journal of Music Theory 37*(2), 351–389.

Rasch, R. A. and R. Plomp (1999). The perception of musical tones. In D. Deutsch (Ed.), *Psychology of Music* (second ed.)., Chapter 4, pp. 89–109. Academic Press.

Regener, E. (1973). *Pitch Notation and Equal Temperament: A Formal Study*. Berkeley: University of California Press.

Reti, R. (1958). *Tonality, Atonality, Pantonality: A study of some trends in twentienth century music*. Westport, Connecticut: Greenwood Press.

Riemann, H. (1914). Ideen zu einer Lehre von den Tonvorstellungen. *Jahrbuch der Musikbibliothek Peters 21/22*, 1–26. Leipzig.

Ritsma, R. (1967). Frequencies dominant in the perception of the pitch of complex sounds. *Journal of the Acoustical Society of America 42*, 191–198.

Salzer, F. (1962). *Structural Hearing, tonal coherence in music*. Dover, New York.

Schechter, M. (1980). Tempered scales and continued fractions. *American Mathematical Monthly 87*, 40–42.

Schenker, H. (1906). *Harmonielehre*. Wien, Universal edition.

Scholtz, K. P. (1998). Algorithms for mapping diatonic keyboard tunings and temperaments. *Music Theory Online 4*(4).

Schouten, J. (1938). The perception of subjective tones. *Proceedings of Koninklijke Nederlandsche Akademie van Wetenschappen 41*(10), 1086–1093.

Schulte, M., A. Knief, A. Seither-Preisler, and C. Pantev (2001). Gestalt recognition in a virtual melody experiment. In J. Nenonen, R. Ilmoniemi, and T. Katila (Eds.), *Biomag2000, Proceedings of the 12th International Conference on Biomagnetism*, Helsinki University of Technology, Espoo, Finland, pp. 107–110.

Seebeck, A. (1841). Beobachtungen über einige Bedingungen der Entstehung von Tönen. *Annals of Physical Chemistry 53*, 417–436.

Sethares, W. (1993). Local consonance and the relationship between timbre and scale. *Journal of the Acoustical Society of America 94*(3), 1218–1228.

Sethares, W. A. (1999). *Tuning, Timbre, Spectrum, Scale* (second ed.). Springer.

Shepard, R. N. (1982). Geometrical approximations to the structure of musical pitch. *Psychological Review 89*, 305–333.

Smoorenburg, G. F. (1970). Pitch perception of two-frequency stimuli. *Journal of the Acoustical Society of America 48*, 924–942.

Smoorenburg, G. F. (1972). Audibility region of combination tones. *Journal of the Acoustical Society of America 52*, 603–614.

Stumpf, K. (1898). Konsonanz und dissonanz. *Beitrage zur Akustik und Musikwissenschaft 1*, 1–108.

Temperley, D. (2001). *The Cognition of Basic Musical Structures*. The MIT Press.

Tenney, J. (1988). *A History of Consonance and Dissonance*. Excelsior, New York.

Tenney, J. and L. Polansky (1980). Temporal Gestalt perception in music. *Journal of Music Theory 24*(2).

Terhardt, E. (1974). Pitch, consonance and harmony. *Journal of the Acoustical Society of America 55*, 1061–1069.

Terhardt, E. (1977). The two-component theory of musical consonance. In E. Evans and J. Wilson (Eds.), *Psychophysics and Physiology of Hearing*, pp. 381–390. Academic, London.

Terhardt, E. (1984). The concept of musical consonance: A link between music and psychoacoustics. *Music Perception I*(3), 276–295.

Terhardt, E. (1987). Gestalt principles and music perception. In W. Yost and C. Watson (Eds.), *Perception of Complex Auditory Stimuli*, pp. 157–166. Erlbaum, Hillsdale, NJ.

Terhardt, E., G. Stoll, and M. Seewann (1982). Algorithm for extraction of pitch and pitch salience from complex tone signals. *Journal of the Acoustical Society of America 71*(3), 679–688.

Tramo, M. J., P. A. Cariani, B. Delgutte, and L. D. Braida (2001). Neurobiological foundations for the theory of harmony in western tonal music. *Annals of the New York Academy of Science 930*, 92–116.

Van de Craats, J. (1989). *De fis van Euler: Een nieuwe visie op de muziek van Schubert, Beethoven, Mozart en Bach*. Aramith Uitgevers, Bloemendaal.

Vogel, M. (1975). *Die Lehre von den Tonbeziehungen*. Verlag für systematische Musikwissenschaft, Bonn-Bad Godesberg.

Vos, P. G. and E. W. van Geenen (1996). A parallel-processing key-finding model. *Music Perception 14*(2), 185–224.

Wallin, N. L., B. Merker, and S. Brown (Eds.) (2000). *The Origins of Music.* The MIT Press.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. *Psychologische Forschung 4*, 301–350.

Yardi, S. and E. Chew (2004). Giving ragas the time of day: Linking structure, emotion and performance time in North Indian classical music using the harmonic network. In *Proceedings of the eighth International Conference on Music Perception and Cognition, ICMPC.*

Yasser, Y. (1975). *A theory of evolving tonality.* New York: Da Capo Press.

Youngblood, J. (1985). Style as information. *Journal of Music Theory 2*, 24–35.

Yunik, M. and G. Swift (1980). Tempered music scales for sound synthesis. *Computer Music Journal 4*, 60–65.

Zonis, E. (1973). *Classical Persian music: An introduction.* Cambridge, MA: Harvard University Press.

Zweifel, P. F. (1996). Generalized diatonic and pentatonic scales: A group-theoretic approach. *Perspectives of new music 34*(1), 140–161.

Zwicker, E. and H. Fastl (1990). *Psychoacoustics.* Springer-Verlag, Berlin.

Index

Titles in the ILLC Dissertation Series:

ILLC DS-2001-01: **Maria Aloni**
Quantification under Conceptual Covers

ILLC DS-2001-02: **Alexander van den Bosch**
Rationality in Discovery - a study of Logic, Cognition, Computation and Neuropharmacology

ILLC DS-2001-03: **Erik de Haas**
Logics For OO Information Systems: a Semantic Study of Object Orientation from a Categorial Substructural Perspective

ILLC DS-2001-04: **Rosalie Iemhoff**
Provability Logic and Admissible Rules

ILLC DS-2001-05: **Eva Hoogland**
Definability and Interpolation: Model-theoretic investigations

ILLC DS-2001-06: **Ronald de Wolf**
Quantum Computing and Communication Complexity

ILLC DS-2001-07: **Katsumi Sasaki**
Logics and Provability

ILLC DS-2001-08: **Allard Tamminga**
Belief Dynamics. (Epistemo)logical Investigations

ILLC DS-2001-09: **Gwen Kerdiles**
Saying It with Pictures: a Logical Landscape of Conceptual Graphs

ILLC DS-2001-10: **Marc Pauly**
Logic for Social Software

ILLC DS-2002-01: **Nikos Massios**
Decision-Theoretic Robotic Surveillance

ILLC DS-2002-02: **Marco Aiello**
Spatial Reasoning: Theory and Practice

ILLC DS-2002-03: **Yuri Engelhardt**
The Language of Graphics

ILLC DS-2002-04: **Willem Klaas van Dam**
On Quantum Computation Theory

ILLC DS-2002-05: **Rosella Gennari**
Mapping Inferences: Constraint Propagation and Diamond Satisfaction

ILLC DS-2002-06: **Ivar Vermeulen**
A Logical Approach to Competition in Industries

ILLC DS-2003-01: **Barteld Kooi**
Knowledge, chance, and change

ILLC DS-2003-02: **Elisabeth Catherine Brouwer**
Imagining Metaphors: Cognitive Representation in Interpretation and Understanding

ILLC DS-2003-03: **Juan Heguiabehere**
Building Logic Toolboxes

ILLC DS-2003-04: **Christof Monz**
From Document Retrieval to Question Answering

ILLC DS-2004-01: **Hein Philipp Röhrig**
Quantum Query Complexity and Distributed Computing

ILLC DS-2004-02: **Sebastian Brand**
Rule-based Constraint Propagation: Theory and Applications

ILLC DS-2004-03: **Boudewijn de Bruin**
Explaining Games. On the Logic of Game Theoretic Explanations

ILLC DS-2005-01: **Balder David ten Cate**
Model theory for extended modal languages

ILLC DS-2005-02: **Willem-Jan van Hoeve**
Operations Research Techniques in Constraint Programming

ILLC DS-2005-03: **Rosja Mastop**
What can you do? Imperative mood in Semantic Theory

ILLC DS-2005-04: **Anna Pilatova**
A User's Guide to Proper names: Their Pragmatics and Semanics

ILLC DS-2005-05: **Sieuwert van Otterloo**
A Strategic Analysis of Multi-agent Protocols

ILLC DS-2006-01: **Troy Lee**
Kolmogorov complexity and formula size lower bounds

ILLC DS-2006-02: **Nick Bezhanishvili**
Lattices of intermediate and cylindric modal logics

ILLC DS-2006-03: **Clemens Kupke**
Finitary coalgebraic logics

ILLC DS-2006-04: **Robert Špalek**
Quantum Algorithms, Lower Bounds, and Time-Space Tradeoffs

ILLC DS-2006-05: **Aline Honingh**
The Origin and Well-Formedness of Tonal Pitch Structures

ILLC DS-2006-06: **Merlijn Sevenster**
Branches of imperfect information: logic, games, and computation