

ComputerMusic Notes

(Representations and Models)

Student: Angelo Antona.

With Contributions from: Francisco Messina, Filippo Marri, Enrico Torres.

Professor: Alberto Bernardini. Music and Acoustic Engineering - 2023/2024.

COMPUTERMUSIC

INDEX

1	Psychoacoustics introduction	5
2	An Overview of the Ear's Structure and Function	6
2.1	The ear	6
2.2	From The Cochlea to The Brain.....	8
2.3	Questions from previous exams	8
3	Perception Of Sound.....	9
3.1	Sounds Illusions	9
3.1.1	Superposition of Pure Tones - 1st Order Effect (<i>originated in the cochlea</i>)	9
3.1.2	Superposition of Pure Tones - 2 st Order Effect (<i>originated in the neural system</i>)	9
3.2	Timbral Consonance and Dissonance	10
3.3	Pitch: The Subjective Impression of Frequency	11
3.3.1	Just Noticeable Difference (JND)	11
3.3.2	Fundamental Note	11
3.3.3	Pitch and Intensity	11
3.3.4	Acoustic Intensity Measurement	12
3.3.5	Intensity Thresholds.....	12
3.4	Masking.....	12
3.4.1	Frequency Masking	12
3.4.2	Temporal Masking.....	13
3.5	Overall Effects of Subjective Sound Perception.....	13
3.6	Auditory Scene Analysis	14
3.6.1	Grouping Principles in Auditory Scene.....	14
3.6.2	Temporal Order Judgement.....	14
3.6.3	Timbre	15
3.7	Questions from previous exams	15
4	Perception of Music.....	17
4.1	Introduction to the Elements of Music.....	17
4.1.1	Fundamental Elements of Music	17
4.1.2	Complex Aspects of Music	17
4.1.3	Higher-Level Musical Concepts	17
4.1.4	The Gestalt of Music	18

4.2	The Mind Machine: Music and Brain Function	18
4.2.1	The Brain Structure	18
4.2.2	Functional Segregation in Music Processing.....	18
4.2.3	Specific Brain Areas Involved in Music.....	19
4.2.4	Neuroplasticity and Music	19
4.3	Anticipation and expectation.....	20
4.3.1	Neuron Anatomy.....	20
4.3.2	Neural basis for musical expectations	21
4.4	Studying Brain Response.....	23
4.4.1	Electroencephalography (EEG)	23
4.4.2	Magnetoencephalography (MEG).....	23
4.4.3	Functional Magnetic Resonance Imaging (Fmri).....	23
4.5	Brain Memory Management Theories.....	23
4.5.1	Constructivist theory (<i>relational</i>).....	23
4.5.2	Record-Keeping Theory.....	23
4.5.3	Multiple Trace Theory (MTT)	23
4.5.4	Mind vs. Brain – Common Illusions.....	24
4.6	Bottom-up vs. Top-down processing	24
4.6.1	Bottom-Up Processing	24
4.6.2	Top-Down Processing.....	24
4.7	Questions from previous exams	25
5	Building Blocks of Tonal Organisation	26
5.1	Keyboard Types in Western Music.....	26
5.1.1	The Stradella Button Keyboard	26
5.1.2	The Uniform/Janko Keyboard	26
5.1.3	The piano keyboard	27
5.2	Tonal Progression of Partial.....	27
5.3	Relations Between Tones.....	28
5.3.1	Intervals.....	28
5.3.2	Triads.....	29
5.3.3	Tetrads	29
5.3.4	Beyond Tetrads	30
5.3.5	Scales.....	30
5.3.6	Modal Scales and the Ionian Mode.....	31
5.3.7	Non-Seven-Note Scales.....	31

5.3.8	The Circle of Fifths.....	32
5.4	Questions from previous exams	33
6	Rhythmic Structures and Models	34
6.1	The Concept of Rhythm	34
6.2	Musical Meter	34
6.2.1	Difference between Rhythm and Metric	34
6.3	Metric Structure.....	34
6.4	Metric Modulation.....	35
6.4.1	Simplification of Periodic Metric Modulations	36
6.5	Polyrhythm and Polimeter	36
6.6	Polyrhythm.....	36
6.6.1	Polyrhythms	36
6.6.2	Hemiola	36
6.6.3	Cross-Rhythm.....	37
6.7	Polimeters	37
6.7.1	Signature Simplification for Polimeters (Polimeters in “Standard” Western Notation).....	37
6.7.2	Polimeters in Brian Ferneyhough's Notation (BFN)	37
6.8	Algebraic Structure of Rhythms	38
6.8.1	Substructures and Symmetry in Polyrhythms.....	38
6.9	The rhythmic wheels.....	39
6.10	Modern Western Notation & Brian Ferneyhough's Notation: Insights	39
6.11	Rhythm and Meter in Poetry	41
6.12	Questions from previous exams	41
7	Harmony	42
7.1	Relation between Melody and Harmonic/Rhythmic structures.....	42
7.1.1	Harmonic Interaction	42
7.1.2	Rhythmic Interaction.....	42
7.1.3	Structural Role.....	42
7.2	Four-Part Harmony	43
7.2.1	Four-Part-Harmony Overview.....	43
7.2.2	Rules and Techniques of Four-Part Harmony	43
7.3	Basic theory.....	44
7.3.1	The Diatonic Major Scale and Its Modes	44
7.4	Tonal Patterns	44
7.4.1	Tetrads and Common Chords Progressions.....	45

7.4.2	Chord Scale Theory	45
7.4.3	Major Modes Chords	45
7.4.4	Minor Modes Chords	46
7.5	Voicings	47
7.6	Voice Leading	47
7.6.1	Cadence.....	47
7.6.2	Clausula Vera (<i>hint</i>).....	48
7.7	Harmonic (Voicing) and Melodic (Voice Leading) Dissonances	48
7.8	Counterpoint.....	48
7.8.1	Basic information on Counterpoint.....	48
7.8.2	Clausula Vera.....	48
7.8.3	Counterpoint species	49
7.8.4	Additional Counterpoint Transformation Techniques	49
7.9	Questions from previous exams	50
8	Musical Structure.....	52
8.1	Most Common Musical Structures	52
8.2	Structural Analysis.....	53
8.2.1	Feature Vectors in Music Analysis	53
8.2.2	Chromagram	53
8.2.3	Mel-Frequency Cepstral Coefficients (MFCCs)	54
8.2.4	Self-Similarity Matrix (SSM)	54
8.2.5	Analytical Methods	55
8.2.6	Phase correlation	55
8.3	Questions from previous exams	56

1 PSYCHOACOUSTICS INTRODUCTION

Psychoacoustics is a field that extends beyond the basic medical understanding of sound. It incorporates not just the physical properties of sound but also its subjective impact, which cannot be fully analyzed through traditional scientific methods. This field focuses on understanding the complete process of how we hear sounds, which involves several key stages:

1. The Source of Sound:

- **Exciter:** This is the part that initiates the sound by providing energy to the system.
- **Resonator:** It shapes the unique quality or 'color' of the sound, often referred to as 'timbre'. The resonator does this by enhancing specific harmonics, which are the various tones that make up a sound.

2. Transmission Medium:

- **Environment:** This refers to how the sound travels or propagates through the air or other mediums.
- **Boundaries:** These include the surfaces and obstacles the sound encounters, leading to reflections, reverberation (echoing), and absorption (diminishing of sound).

3. The Receiver – Our Ear:

- The ear functions as a transducer, a device that converts one form of energy into another. It changes the acoustic wave (which is a pressure wave) into a mechanical wave. This mechanical wave is then transformed into electrical signals in the cochlea, a part of the inner ear.

Each of these stages plays a crucial role in how we perceive and understand sound. The process is complex, involving both the physical properties of sound and our subjective interpretation of it.

2 AN OVERVIEW OF THE EAR'S STRUCTURE AND FUNCTION

2.1 THE EAR

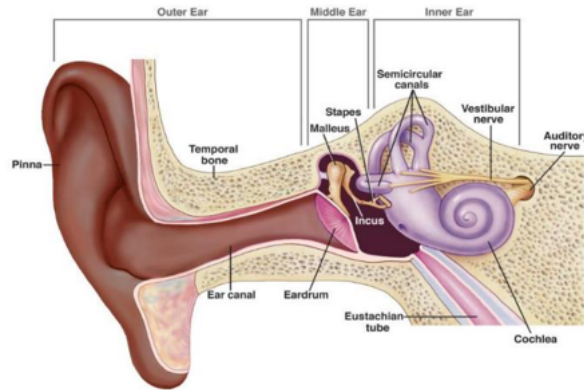


Figure 1

The human ear is a complex and highly specialized organ responsible for hearing. It consists of three main sections: the outer ear, middle ear, and inner ear, each playing a vital role in the process of hearing.

Let's analyze the composition of the human auditory system by following the path of sound from the environment to the brain:

1. Outer Ear:

- **Pinna:** Sound waves enter through the pinna (the visible part of the ear), which captures these waves and funnels (*incanala*) them into the ear canal. The pinna is unique to everyone, similar to a fingerprint. This uniqueness contributes to slight differences in how each person hears.
- **Ear Canal:** The sound waves travel down this canal, reaching the eardrum (tympanic membrane).

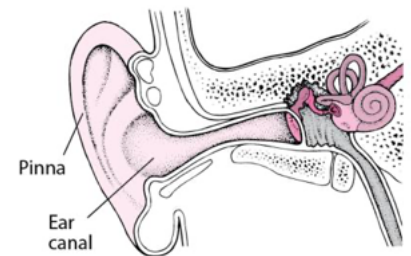


Figure 2

2. Middle Ear:

- **Eardrum:** This membrane vibrates in response to the sound waves.
- **Ossicles:** The middle ear contains three tiny bones named the malleus (hammer), incus (anvil), and stapes (stirrup).
 - **Malleus (Martello):** Attached to the eardrum, it vibrates with the eardrum.
 - **Incus (Incudine):** Receives vibrations from the malleus and passes them to the stapes.
 - **Stapes (Staffa):** The stapes vibrates against the oval window, a membrane-covered opening (*un'apertura coperta da membrana*) to the inner ear, transferring the sound vibrations.

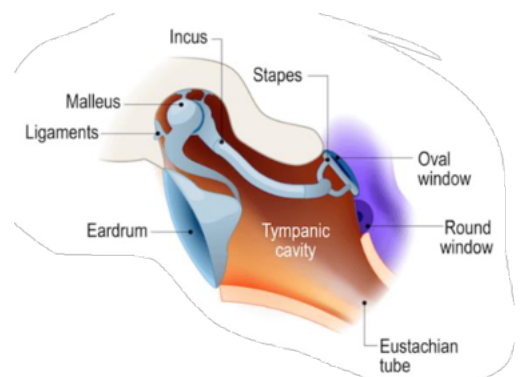


Figure 3

They function like an electrical transformer, amplifying pressure and reducing motion while preserving energy. This process adapts the sound wave from air to the liquid environment of the cochlea.

- **Eustachian Tube:** This tube links the ear canal to the throat (*gola*), helping to equalize pressure between the ear's internal and external environments. This equalization protects the eardrum from sudden changes in air pressure and loud noises.

3. Inner Ear:

- **Cochlear Windows (Oval and Round):** The oval window connects to the stirrup, while the round window acts as a damping mechanism. This setup compensates for pressure changes caused by the oval window's oscillations.
- **Cochlea:** The most crucial part of the inner ear, the cochlea is a labyrinthine structure where vibrations are converted into electrical signals. It is filled with perilymph, a fluid with an ionic composition similar to plasma and cerebrospinal fluid. The vibrations cause the fluid inside the cochlea to move.
- **Hair Cells:** The cochlea's basilar membrane is lined with over 30,000 hair cells. These cells convert the mechanical vibrations into electrical signals as they bend (*mentre si piegano*) in response to fluid movement. Different regions of the basilar membrane are sensitive to different frequencies, following the logarithmic function $\omega_0 = 2^{\frac{x}{3.5}} + a$ (The variable a represents the characteristic frequency location on the basilar membrane of the cochlea closest to the oval window. It is the frequency for which the distance x from the oval window is zero). High-frequency sounds primarily vibrate the base of the membrane (near the oval window), which is stiffer and narrower (*più rigido e più stretto*). Low-frequency sounds, conversely, vibrate the apex (*apice*) of the membrane, which is wider and more flexible.
- **Auditory Nerve:** These electrical signals are transmitted to the brain through this nerve.

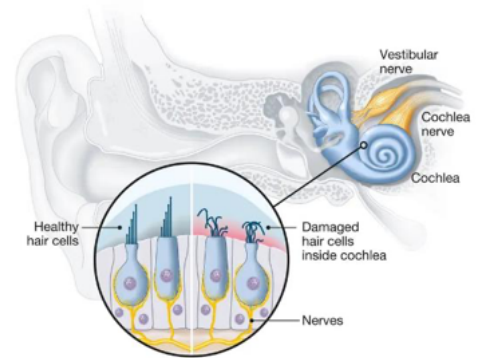


Figure 4

The middle and inner Ear are covered by the **Temporal Bone**, that is one of the hardest parts of the body.

2.2 FROM THE COCHLEA TO THE BRAIN

The process by which the cochlea turns acoustic vibrations into electrical signals and then conveys them to the brain involves several steps within the auditory pathway:

1. **Cochlea:** As we have seen, sound waves enter the ear and are translated into mechanical movements by the eardrum and the ossicles of the middle ear. The cochlea, a fluid-filled spiral structure in the inner ear, contains the organ of Corti, where hair cells are deflected by these mechanical movements, converting them into electrical signals.
2. **Cochlear/Auditory Nerve:** These electrical signals are then transmitted by the auditory nerve fibres, which make up the cochlear nerve. Each hair cell is connected to a neuron that carries the signal. The cochlear nerve carries the electrical impulses to the brainstem (*figure 5*).
3. **Cochlear Nucleus:** The signal first arrives at the cochlear nucleus in the brainstem (*tronco encefalico*), which is the first relay station (*stazione di trasmissione*) in the central auditory system.
4. **Superior Olivary Nucleus:** From the cochlear nucleus, the signal is sent to the superior olivary complex, where sounds from both ears are compared to help determine the direction they come from.
5. **Lateral Lemniscus:** Signals then travel to the lateral lemniscus, which is a tract of axons carrying the information regarding sound through the brainstem.
6. **Inferior Colliculus:** The signal reaches the inferior colliculus, an important midbrain nucleus in the auditory pathway, which acts as a major relay and processing station.
7. **Medial Geniculate Nucleus:** From the inferior colliculus, the signal is sent to the medial geniculate body of the thalamus, which is involved in the tonotopic organization and relay of auditory information (*Just like the cochlea, the MGB is organized tonotopically, which means that neurons in different parts of the MGB respond preferentially to different frequencies of sound. This organization is maintained from the cochlea through the auditory pathway and into the auditory cortex*).
8. **Auditory Cortex:** Finally, the information reaches the auditory cortex located in the temporal lobe of the brain, specifically the primary auditory cortex. Here, the processing of sound information such as pitch and volume occurs.

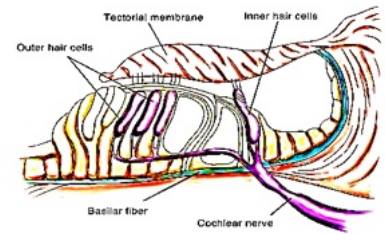


Figure 5

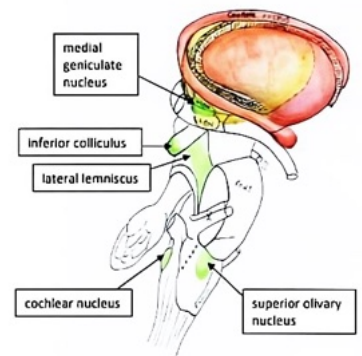


Figure 6

2.3 QUESTIONS FROM PREVIOUS EXAMS

Q2. Our auditory organ (cochlea) turns acoustic vibrations into electrical signals which are conveyed to the brain through the cochlear nerve. Can you describe the path followed by these electrical signals through the brain? (*start from cochlear nerve and continue to where this nerve connects to, and so on... End with the region of the cortex devoted to processing sounds*).

(More questions about the ear's structure that pertain to sound illusions and subjective perception will be covered in the next section of this text).

3 PERCEPTION OF SOUND

OUR AUDITORY SYSTEM, WHILE COMPLEX, DOES HAVE ITS LIMITATIONS.

3.1 Sounds Illusions

3.1.1 Superposition of Pure Tones - 1st Order Effect (originated in the cochlea) 🎧

When two sounds with closely related frequencies, denoted as f_1 and f_2 , are played together, we don't hear them as separate. Instead, we perceive a single tone that oscillates. This perceived tone has a frequency that is the average of f_1 and f_2 , and it also beats or pulsates at a frequency that is the difference between $\frac{f_1 - f_2}{2}$.

Mathematically, this is represented as follows:

$$\sin(A) + \sin(B) = 2\sin\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right)$$

$$\sin(2\pi f_1 t) + \sin(2\pi f_2 t) = 2\sin\left(2\pi \frac{(f_1 + f_2)}{2} t\right)\cos\left(2\pi \frac{(f_1 - f_2)}{2} t\right)$$

The sine component on the right-hand side of the equation corresponds to the average frequency of the two tones, which is the frequency we perceive. The cosine component represents the beat frequency, which is how fast the sound seems to pulse.

What's fascinating is that we are hearing a frequency that is not actually being played by any of the two sources. This is a testament to the constructive and destructive interference patterns that sound waves can form and how our auditory system interprets these patterns.

Risset's musical beat: In addition to this phenomenon, there is Risset's musical beat. This illusion occurs when the following conditions are simultaneously satisfied:

- a base frequency is played along with its harmonics (*Harmonics are frequencies that are integer multiples of the base frequency*).
- each harmonic's amplitude decreases inversely with its frequency—meaning the higher the frequency, the lower its amplitude.

In this way, the combination of these sounds can create the illusion of a rhythm.

If, for example, we consider a sound of $f = 100 \text{ Hz}$ accompanied by seven harmonics, the perceived rhythm can feel like a 7/8-time signature. 🎧

3.1.2 Superposition of Pure Tones - 2st Order Effect (originated in the neural system) 🎧

This effect is rooted in the neural system. Unlike physical sound wave interference, the 2nd-order effect involves the brain's interpretation of the sound waves: when two pure tones are played with frequencies f_1 and f_2 , where f_2 is slightly different from $2 * f_1$, the brain perceives a tone beat at the frequency $f_2 - 2 * f_1$.

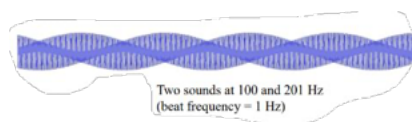
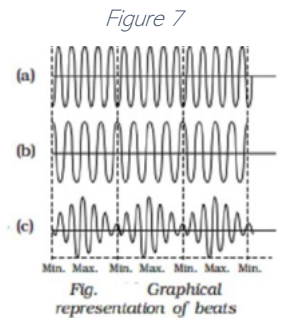


Figure 8



In contrast to the 1st-order effect, where beats are heard as changes in amplitude, the 2nd-order effect maintains a constant amplitude. Instead of amplitude modulation, listeners experience cyclic changes in the shape of the waveform. This can be perceived as a variation in the sound's timbre or quality, without fluctuations in loudness.

3.2 TIMBRAL CONSONANCE AND DISSONANCE

Understanding whether two sounds are consonant or dissonant involves analyzing their frequency separation and how this separation is perceived by the human auditory system. In particular, it is important the concept of critical Bandwidth, which refers to the frequency range within which two sounds affect the same group of hair cells in the cochlea. The sounds within a critical bandwidth are more likely to interact, affecting our perception of consonance and dissonance.

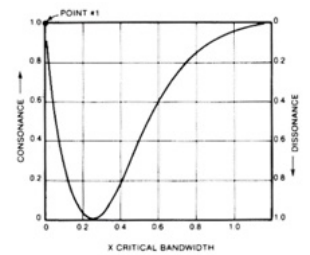


Figure 9

When two tones are close in frequency (within a small separation), we perceive a single beat

or a combined tone rather than two distinct sounds

because these sounds are in the same critical bandwidth and stimulate the same hair cells. As the frequency separation increases, we initially perceive a sensation of roughness (*ruvidosità*, *rozzezza*), which is associated with dissonance. When the frequency separation exceeds the critical bandwidth, we perceive two distinct tones, indicating a transition from dissonance to consonance.

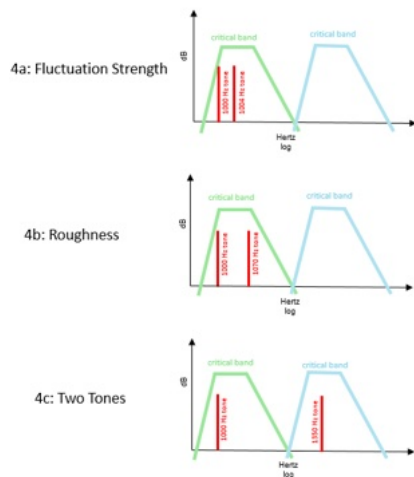


Figure 10

The width of the critical bandwidth changes in function of the frequency: at lower frequencies, the critical bandwidth is larger, which can make intervals sound muddy (*torbido*) or less distinct; at higher frequencies, the critical bandwidth becomes narrower, and intervals sound more consonant and clearer.

This phenomenon explains why certain intervals on a piano sound consonant in higher registers but dissonant in lower registers.

For the same reason, the spacing of chords in Western music tends to become wider in lower registers, which helps avoid sensory dissonance. The left hand of a pianist often plays voicings that span a wider range when the lowest note is around the octave below middle C. In contrast, the right hand can play more closely spaced intervals without creating dissonance because the critical bandwidth in higher registers is narrower.



3.3 PITCH: THE SUBJECTIVE IMPRESSION OF FREQUENCY

Pitch is a subjective impression of frequency. It doesn't always correspond to the real frequency of the sound, a phenomenon due to the way our brain processes sound information.

3.3.1 Just Noticeable Difference (JND)

It is the smallest variation of a frequency we can perceive. Our hearing is most sensitive in the mid-frequency range of 20Hz to around 20kHz, with the most sensitivity lying in the logarithmic scale from 1-4 kHz, crucial for the frequencies characterizing speech. *Figure 11* presents a graph that illustrates this concept. The difference limen is the smallest change in a stimulus that can be detected. On the y-axis, we have the difference limen for frequency expressed as a percentage, and on the x-axis, we have the frequency of the tones in hertz (Hz). We can observe from the graph that at lower frequencies (below 500 Hz), the percentage change needed for a difference to be noticed is larger. This implies that at lower frequencies, our ears are less sensitive to changes in frequency. As the frequency increases, particularly in the mid-range (around 1000 to 5000 Hz), our sensitivity to frequency changes increases, meaning that smaller percentage changes are detected. However, at very high frequencies (above 5000 Hz), the curves level out, indicating that our ability to discern changes does not improve significantly in this range.

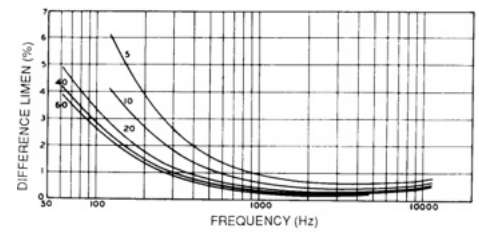


Figure 11

3.3.2 Fundamental Note

All sounds that create a sensation of pitch are periodic. Their spectrum consists of harmonics that are integer multiples of the fundamental frequency. The pitch of a complex periodic tone is closely related to the pitch of a sine wave at the fundamental. Helmholtz postulated that pitch is heard at the fundamental frequency because this frequency creates the lowest frequency peak on the basilar membrane in the ear.

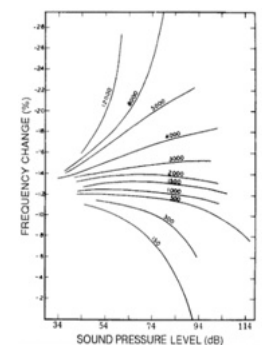
Seebeck and later Schouten demonstrated that even complex periodic sounds with no energy at the fundamental frequency can still produce a clear pitch sensation at the fundamental. This phenomenon is called **Missing fundamental**.



Figure 12

3.3.3 Pitch and Intensity

Pitch not only depends on the frequency but also on the intensity or amplitude of sound. Even when playing the same pitch at different volume levels, our pitch perception may change, leading us to perceive different notes. The *figure 12* shows how pitch changes in function of the intensity (*when intensity increases: low frequencies are perceived as lower pitch, high frequencies are perceived as higher pitch*).



3.3.3.1 Mel Scale

Pitch is measured using the Mel scale (*figure 13*). It is a scale that relates the perceived frequency of a tone to the actual measured frequency. It's based on subjective tests to quantify how listeners perceive pitches: the idea is that equal distances on the Mel scale sound equally distant to the human ear. The scale is non-linear: it is more linear at lower frequencies and more logarithmic at higher frequencies. This reflects how the human ear perceives sound. The unit of measurement on this scale is "Mels." For reference, 1000 Hz is usually set to 1000 Mels, and other frequencies are scaled relative to this.

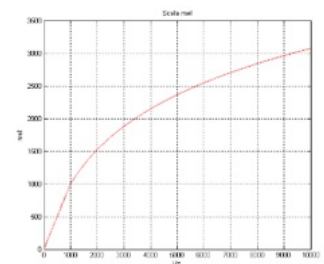


Figure 13

Figure 13 shows what we have just said: the horizontal axis represents the frequency of the sound in Hertz (Hz), while the vertical axis represents the perceived pitch of the sound measured in Mels. As usually, 1000 Hz is set as a reference point where the frequency in Hertz and the perceived pitch in Mels is the same.

3.3.4 Acoustic Intensity Measurement

Acoustic intensity, often described as loudness, is how we subjectively experience the power of a sound. It's not just about how strong the sound is; it's about how we perceive its strength. For example, two sounds might have the same Sound Pressure Level (SPL), which is an objective measurement, but we might hear them as being of different loudness if they are at different frequencies.

3.3.4.1 Phon scale and Fletcher and Munson's iso-phonic curves

The unit of measurement of perceived intensity is the phon. The phon scale is based on the dB scale (decibels) but adjusted for the ear's varying sensitivity at different frequencies. For instance, our ears are less sensitive to very low/very high frequencies. The relation between perceived and real intensity is:

$$0 \text{ phons} = \text{perceived intensity of a 1000hz tone at 0 Db (SPL)}$$

So, if a sound is said to have a level of x phons, it is as loud to us as a 1000 Hz tone would be at x decibels SPL.

To better understand this concept, we could use the Fletcher and Munson's iso-phonic curves (figure 14). This plot shows several curves, each representing a constant loudness level in phons. The curves swoop upwards at the lower and higher ends of the frequency spectrum, indicating that at these frequencies, a sound must have a higher SPL for us to perceive it as being equally loud to a sound at 1000 Hz and 0dB. The lowest curve marks the threshold of audibility—below this line, sounds are too faint to hear. The highest curve denotes the threshold of pain—above this, sounds are uncomfortably loud. This plot illustrates that our perception of loudness changes with frequency and is not a straight linear relationship with the SPL.

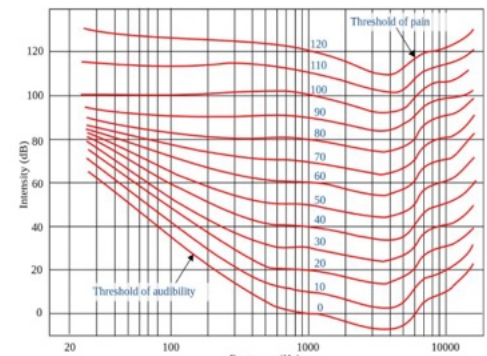


Figure 14

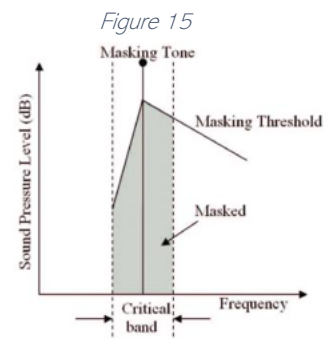
3.3.5 Intensity Thresholds

The audibility of sound also depends on intensity thresholds, which are the minimum and maximum SPLs at which sound is perceivable without causing pain. The threshold of pain indicates the SPL beyond which sound becomes painfully loud, while the hearing threshold indicates the lowest SPL at which sound is just audible.

3.4 MASKING

3.4.1 Frequency Masking

Frequency masking occurs when the presence of a stronger sound, or "masker," makes it difficult to hear a second sound at a similar or neighbouring frequency. This happens because the masker and the masked sound compete at the level of the auditory system. When two sounds are played together, if one is significantly louder (the masker), it can obscure our ability to hear the other sound (the test tone). The point at which the *less loud* tone becomes just audible is known as the threshold of audibility.



3.4.2 Temporal Masking

Temporal masking describes the auditory system's delayed response to sounds following a loud noise. After exposure to a loud sound (the masker), the ear requires time to recover and become sensitive again to softer sounds at similar frequencies. This recovery time can involve both a pre-mask period, occurring immediately before the masker is introduced, and a post-mask period, following the cessation of the masker. During these times, our perception of other sounds is 'masked,' and our ability to hear them is diminished.

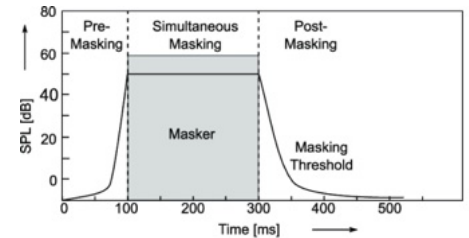


Figure 16

The concept of auditory masking is used in the development of audio compression algorithms, such as those used in MP3 and AAC encoding, which remove sounds that are likely to be masked and therefore inaudible, allowing for more efficient compression.

3.5 OVERALL EFFECTS OF SUBJECTIVE SOUND PERCEPTION

When you record sound at a certain level of loudness and then listen to the recording at a much lower sound intensity, several things can happen related to the perception and quality of the sound:

- **Perceived Loudness Changes** (*ovviamente*).
- **Frequency Response Changes**: our perception of the frequency balance of sounds changes with loudness (Fletcher-Munson curves).
- **Psychoacoustic Effects**: There are psychoacoustic effects where the brain fills in some missing frequencies and harmonics (frequency and temporal masking, missing fundamental note).
- **Speaker or Headphone Limitations**: The output device (speakers or headphones) may have its own frequency response that changes with volume, further altering the sound.

In order to partially compensate to these perceptual changes, sound engineers often check the mix at various volume levels to ensure that it sounds good both loud and soft. Let's break down how they approach the task with respect to volume levels:

- **Mixing**: during the mixing process, engineers adjust the levels of individual tracks (instruments, vocals, etc.) to achieve a harmonious balance. Since our hearing sensitivity changes at different volume levels, particularly for low and high frequencies, engineers tweak equalization (EQ) settings to ensure the mix retains a pleasing balance and that all elements are audible, even at low volumes.
- **Mastering**: Mastering involves final adjustments to the mixed track, ensuring uniformity and consistency of sound across all playback systems. This can include applying compression and limiting to reduce the dynamic range, which helps make quiet sounds more audible at lower volumes without making loud sounds overwhelming (*opprimente*).

The goal is to create a final product that maintains its intended emotional impact, clarity, and balance whether it's played softly in the background or blasted at maximum volume.

3.6 AUDITORY SCENE ANALYSIS

Auditory scene analysis is the process by which the brain organizes sound into perceptually meaningful elements. The brain groups sounds based on several principles to make sense of the complex acoustic environments we experience daily.

3.6.1 Grouping Principles in Auditory Scene

The grouping Principles in Auditory Scene Analysis are:

- **Proximity:** The principle of proximity in auditory grouping is based on the concept that sounds which are close together in frequency are often perceived as being related or produced by the same source. This grouping is not just about physical proximity in space, but more about the spectral or tonal proximity. For instance, two notes played on a piano that are close in pitch will likely be heard as related or part of the same melodic line. Similarly, sounds that share similar qualities, such as timbre, are also grouped together. This is because our auditory system uses timbre to differentiate between different instruments or voices, even if they're playing or singing the same note. So, if two sounds have a similar 'color' or textural quality, we tend to perceive them as coming from the same source.
- **Common Fate:** Sounds that start and stop together or share frequency and amplitude modulations, such as vibrato, are grouped together, suggesting a common source. A classic example is a group of instruments in an orchestra playing the same musical phrase; the synchronous onset and offset of their notes make us hear them as a cohesive unit. In a more dynamic sense, sounds that modulate in frequency or amplitude in the same way, such as the vibrato in a singer's voice or the consistent swell of volume in a piece of music (the 'envelope'), are often perceived as being linked.
- **Good Continuation:** The principle of Good Continuation in auditory scene analysis refers to our perceptual tendency to group together sound elements that follow a predictable pattern over time. When a sound displays consistent properties, such as a steady frequency or pitch, our auditory system interprets this consistency as a cue that the sound is emanating from a single, continuous source. For example, if you're listening to a siren that momentarily gets drowned out (*sopraffatto*) by the noise of a passing vehicle, you still perceive the siren as ongoing. Despite the interruption, the brain 'connects the dots' and continues to perceive the siren as a constant sound because it has established a pattern of what the siren sounds like and expects this pattern to continue.
- **Continuity Effect:** The Continuity Effect builds on the concept of Good Continuation. It happens when a continuous sound is momentarily obscured by a louder noise, known as a 'masker.' The auditory system, having already recognized the sound as continuous, will fill in the gaps caused by the masking noise. This effect shows the brain's ability to restore interrupted sounds based on the expectation that they continue unchanged during the period of masking. An everyday example of the Continuity Effect is when you're listening to music through headphones and someone speaks to you briefly; you don't perceive the music as having stopped. Instead, your brain fills in the sound of the music during the moment of speech, smoothing over the interruption.

3.6.2 Temporal Order Judgement

This refers to the brain's ability to identify the sequence of sounds. If you hear a series of notes played on a piano or a sequence of spoken digits, your brain can easily determine the order in which they were played or spoken. This capability is fundamental to activities such as following a conversation, enjoying music, or understanding the temporal pattern of sounds in nature, like the sequence of bird calls.

3.6.3 Timbre

Timbre is the quality of a sound that makes it distinct from other sounds with the same pitch and loudness. It depends on several factors: Spectral Content (The distribution of energy across various frequencies), Transients (Brief, high-energy bursts at the onset of sounds), Harmonic Structure (The ratio and series of frequencies that make up the sound; harmonic sounds have frequencies that are integer multiples of a fundamental frequency, while anharmonic sounds do not). Timbre allows us to distinguish between different instruments playing the same note, for instance.

3.7 QUESTIONS FROM PREVIOUS EXAMS

- Q1.** Briefly describe the physiology of the hearing system, with particular reference to the organ of Corti (cochlea). Can you explain why the structure of the cochlea introduces limits in your ability to resolve sinusoids whose frequencies are very close to each other?
- Q1.** How do you graphically describe the subjective impression of sound intensity (loudness) as a function of frequency? Please specify the involved measurement units. What happens when you record sound at a certain level of loudness and then you listen to the recording at a much lower sound intensity? And how do you compensate for it?
- Q1.** A sound generator produces two simultaneous sinusoidal tones. The former is a loud tone at 1046 Hz (corresponding to the note C6). Describe what the average person hears in the following situations, and explain why:
- a) The magnitude of the second tone is the same as that of the first tone, and its frequency shifts continuously from 1036Hz to 1056Hz.
 - b) The second tone is a sinusoid whose SPL is 40dB below that of the first one, and its frequency shifts progressively from 1319Hz (corresponding to about E₆) to 1046Hz (C6).
- Offer an interpretation of such phenomena in relation to the physiology of human hearing.
- Q1.** Briefly describe (with the help of a diagram) how pitch (i.e., perceived tonal frequency) is affected by the loudness of the tone itself. Also discuss the possible problems that might arise from this non-ideal aspect of our perception.
- Q1.** Describe the perception of tonal (timbral) dissonance and explain how it can be explained in relation of hearing physiology. Does this phenomenon depend on the frequency of the fundamental of the involved tones? If so, in what way? and what impact does this have on a piano performance?
- Q1.** Briefly describe the subjective impression of sound intensity as a function of frequency. How do you define its measurement unit? What happens when you record sound at a certain level of loudness and then you listen to the recording at a much lower sound intensity?

- Q1.** A sound generator produces two simultaneous sinusoids. The former is a loud tone at 880Hz (corresponding to A5), and the corresponding critical band has an approximate range of 700-1100Hz. Considering all that you know of sound perception, make an educated guess of what the average person would hear in the following situations:
- a) The second tone is an equally loud sinusoid whose frequency shifts continuously from 870Hz to 890Hz
 - b) The second tone is an equally loud sinusoid of frequency 932Hz (corresponding to A5#).
 - c) The second tone is a sinusoid 40dB below the first one, and its frequency shifts progressively from 1109Hz (corresponding to C#6) to 932Hz (corresponding to A5#)

Please be specific in your answer and offer an interpretation of such phenomena in relation to the physiology of human hearing.

- Q1.** Describe the perception of tonal consonance and dissonance and explain how this phenomenon can be explained in physiological terms (refer to the role of hair cells as resonating filters with overlapping bandwidths and describe how this translates in terms of spatial firing density of the neural receptors of the hair cells). How does the perception of dissonance change as the tonal frequency grows (please explain using a graphical representation)? And how does this translate in practical terms, for example, for a piano player?
- Q1.** Describe the subjective perception of sound intensity using equal-loudness contour diagrams, and describe the related unit of measurement. Discuss the meaning of the audibility threshold in relation to such diagrams. Also discuss what happens to this threshold in the presence of a loud sinusoidal tone and elaborate on the practical implications that this might have.

4 PERCEPTION OF MUSIC

Understanding how our brain processes musical information involves several key concepts rooted in the elements of music. The term music itself derives from the Greek word "mousike," which denotes the art of the Muses and encompasses not just sound but also rhythm, melody, and the expression of emotion.



4.1 INTRODUCTION TO THE ELEMENTS OF MUSIC

4.1.1 Fundamental Elements of Music

- **Tone:** This refers to a discrete musical sound. The terms 'tone' and 'note' are often used interchangeably. A tone is the actual sound produced, while a note is the written symbol representing that sound in musical notation.
- **Pitch:** It is the perceptual interpretation of the frequency of a tone. Pitch enables us to categorize tones as high or low on a musical scale.
- **Rhythm:** This aspect concerns the duration of notes and their grouping into patterns over time, creating the temporal structure of music.
- **Tempo:** Tempo denotes the overall speed of a musical piece.
- **Contour** (*Linea melodica*): It is the perceived shape of a melody, characterized by the directional movement of the notes.
- **Timbre:** Often described as the "color" of music, timbre is the characteristic that distinguishes different sounds of the same pitch and loudness, often influenced by the spectral content and transient characteristics of the sound.
- **Loudness:** A subjective attribute relating to the perceived intensity or volume of a sound.
- **Spatial Location:** The perceived origin of a sound source in an environment.
- **Reverberation:** It refers to the acoustic reflections in a space that contribute to the sense of environment around the sound.

4.1.2 Complex Aspects of Music

- **Meter:** It is how we perceive and organize time in music and involves grouping tones in a recurring pattern, such as measures in 3/8 or 4/4 time.
- **Key:** This establishes the tonal center of a composition, often characterized by a scale and the tonic note or chord that provides a sense of resolution.

4.1.3 Higher-Level Musical Concepts

The previous concepts combine to create:

- **Harmony:** It pertains to the vertical arrangement of notes and includes the progression of chords and the tensions and resolutions created within a piece. Counterpoint is a related concept involving the interplay of independent melodic lines to produce harmonious effects.
- **Melody:** The memorable theme or main line of a musical piece, consisting of a sequence of tones arranged in a recognizable pattern. Melody is akin (simile) to the narrative thread of a piece, often remembered for its distinct contour and rhythm.

4.1.4 The Gestalt of Music

The brain's ability to integrate the principles listed above allows us to experience music not just as a series of isolated notes, but as a rich, unified expression of human creativity. This synthesis is not merely a passive process; it involves active interpretation and the dynamic interplay of sensory input and cognitive functions. This capability is called Gestalt of Music.

4.2 THE MIND MACHINE: MUSIC AND BRAIN FUNCTION

4.2.1 The Brain Structure

Music, a complex auditory experience, is intricately processed by nearly every region of the brain and various neural subsystems. To understand music's impact, we look at the brain's structure, acknowledging its division into four lobes plus the cerebellum:

1. **Frontal Lobe:** associated with planning, self-control. It is called the "perceptual organisation" connected to the Gestalt principle.
2. **Parietal Lobe:** associated with motor movements and spatial skill.
3. **Temporal Lobe:** associated with hearing and memory.
4. **Occipital Lobe:** associated with vision.
5. **Cerebellum:** manages the emotional aspect and coordination of music-related movements.

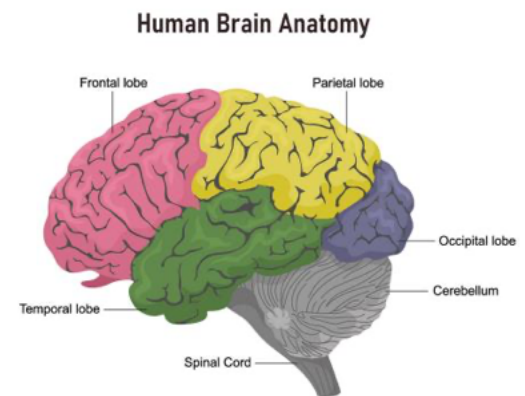


Figure 17

4.2.2 Functional Segregation in Music Processing

Functional segregation in the brain refers to the way that different cognitive processes are handled by different regions of the brain. This concept is based on the observation that certain functions are localized to specific areas, a principle known as localization of function. The brain is partitioned into areas that specialize in processing different types of information, such as vision, hearing, language, and executive functions.

The level of abstraction at which these functions operate can be both low and high. Lower-level functions tend to be more strictly localized. For example, primary sensory and motor areas have very specific locations and functions. Higher-level functions, such as problem-solving, decision-making, and abstract thinking, tend to involve more complex networks that span multiple brain regions.

Areas involved in music processing often overlap with those used in language processing, suggesting a shared mechanism for organizing sounds into structured patterns.

4.2.3 Specific Brain Areas Involved in Music

When we try to follow along (*seguire*) with music that is familiar, there are some parts of the brain responsible for this:

- The **cerebellar vermis** is a narrow, worm-like structure that runs along the midline of the cerebellum, essentially dividing the cerebellum into two hemispheres. It plays a significant role in the non-cognitive aspects of music perception, particularly in the processing of rhythm and tempo. It contributes to motor control and coordination, which are essential when you're tapping your foot to a beat or dancing to music. Musicians who have undergone extensive training often show changes in the size and activity of the cerebellar vermis, reflecting its role in the fine motor skills required for playing instruments.
- The **hippocampus** (temporal lobe) plays a pivotal (*fondamentale*) role in recalling musical experiences and anticipating music based on past encounters.
- The **cerebellum** is involved in beat following.
- The **frontal lobes** are heavily involved when performing music.
- The **motor cortex** (parietal lobe) is activated during musical performances and when moving to the beat.
- The **sensory cortex** (parietal lobe) provides tactile feedback when interacting with instruments or feeling music vibrations.
- The **auditory cortex** (temporal lobe) is critical for the early stages of hearing and processing sounds.
- The **language centres**, including Broca's (frontal lobe) and Wernicke's (temporal lobe) areas, are utilized when deciphering lyrics.

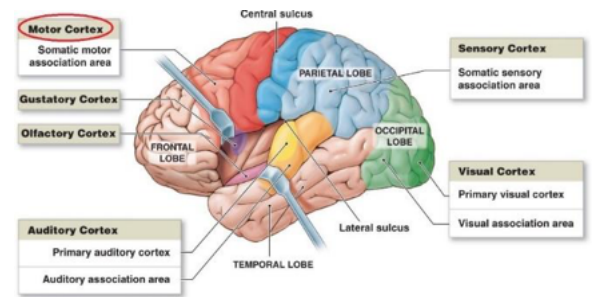


Figure 18

4.2.4 Neuroplasticity and Music

The brain exhibits neuroplasticity, the ability to reorganize functions if certain areas are damaged. This adaptability is evident in musicians, who may develop specific brain regions more extensively due to their practice and experience.

4.3 ANTICIPATION AND EXPECTATION

Music is of course organized, but we need a form of variation, or "surprise", otherwise there would be no emotion. The appreciation we have for music is intimately related to our ability to learn the underlying structure of the music we like and to be able to make predictions about what will come next. Composers will fill music with emotion by knowing what our expectations are, and then deliberately control when those expectations will be met, and when they won't. Basically, there should be a sweet spot between a piece being boring and incoherent or unstructured (but this also depends on the listener).

4.3.1 Neuron Anatomy

To understand the neural basis for musical expectations, it is important to have clear the following concepts:

Neurons: neurons, or nerve cells, are the building blocks of the nervous system. They are specialized cells designed to transmit information throughout the body. This information is carried in the form of electrical signals. Each neuron has a cell body (soma), which contains the nucleus and the machinery for the cell's functions. Extending from the cell body are dendrites, which receive signals from other neurons, and a long axon, which sends signals to other neurons.

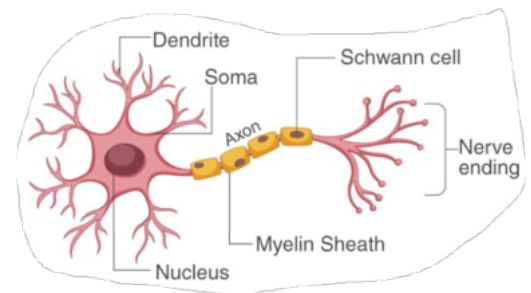


Figure 19

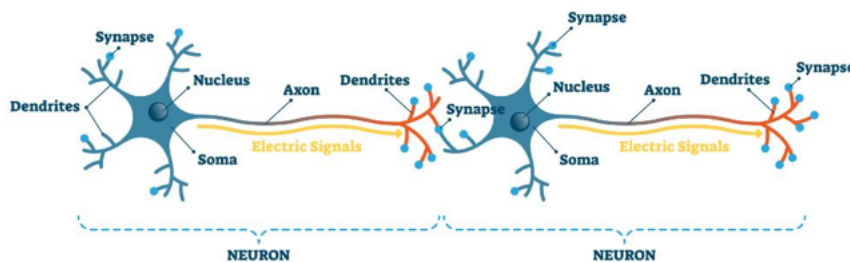


Figure 20

Synapses: a synapse is a junction between two neurons. It is not a physical connection; there is a small gap called the synaptic gap that separates the neurons. When an electrical signal reaches the end of an axon (the axon terminal), it cannot jump the synaptic cleft. Instead, it triggers the release of chemicals called neurotransmitters.

Neurotransmitters: neurotransmitters are the brain's chemical messengers. They are released from the axon terminal of one neuron and cross the synaptic cleft to the next neuron where they bind to receptors on the neuron's surface. The binding of neurotransmitters to receptors is often described by the lock and key model. Each neurotransmitter (the key) has a specific shape that fits into a specific receptor (the lock). When the right neurotransmitter binds to its receptor, it can trigger or inhibit an electrical impulse in the receiving neuron.

The neurotransmitters most involved in listening to music are:

- **Dopamine:** This neurotransmitter is often associated with the pleasure and reward system of the brain. It is released in response to pleasurable activities, including listening to music. Dopamine is mainly produced in the substantia nigra and the ventral tegmental area (VTA) of the brain.
- **Serotonin:** This neurotransmitter is involved in the regulation of mood and emotion. It may contribute to the mood-stabilizing effects of music.
- **Oxytocin:** Sometimes released in response to music, oxytocin is associated with bonding and trust and may contribute to the social and communal experience of music. It is produced in the hypothalamus.

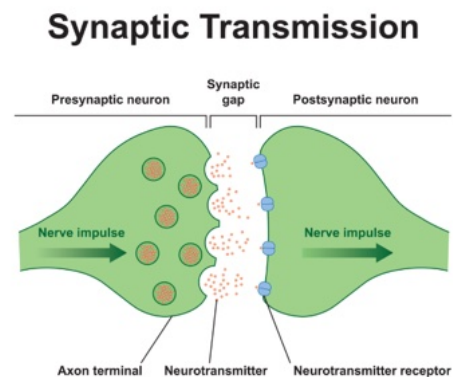


Figure 21

- **Endorphins:** These are natural opioids that can induce feelings of euphoria and pain relief. They may be part of the reason why music can be emotionally moving and sometimes even provide a sense of relief or healing. Endorphins are produced in the pituitary gland and other parts of the brain.

Listening to music can lead to the release of these neurotransmitters, which can elevate our mood and provide pleasure. This is part of why we find music rewarding or why certain songs can change our mood. This can enhance cognitive performance and memory: the release of neurotransmitters during music listening can modulate these cognitive processes, which are also mediated by neuronal activity.

Reuptake: reuptake is a process by which neurotransmitters that have been released into the synaptic cleft are transported back into the presynaptic neuron after they have performed their function of transmitting a neural impulse. The neurotransmitters are taken back up into the presynaptic neuron via specialized proteins called transporters. This process serves as a way of terminating the signal between neurons and also conserves neurotransmitters by recycling them for future use. The most commonly known reuptake process is for the neurotransmitter serotonin, facilitated by the serotonin transporter (SERT).

4.3.2 Neural basis for musical expectations

The brain's music processing system operates independently from its language processing system. Initially, both systems start in similar regions of the frontal and parietal lobes, but as we age, they become distinct. We've discovered that memories associated with music are stored differently, relying heavily on tune recognition. This process involves complex neural computations that interact with our memory. Our brains selectively focus on consistent features across multiple listenings, ignoring the variable ones.

The categorization of music in our brains involves several key aspects:

- We form categories based on prototypes, or standard examples. These prototypes are rooted in our physiology (meaning that these prototypes and our ability to categorize music are not just psychological or cultural constructs, but are also deeply integrated into, and affected by, our biological and neurological structures).
- We evaluate new musical pieces based on their resemblance to these prototypes.

The boundaries defining these musical categories are not rigid.

This categorization process explains why children often prefer simpler songs, as their frontal lobes and other brain parts are not yet fully developed, limiting their capacity to process complex music. As adults, we can appreciate more complex compositions. However, there's a threshold to this complexity, described by the inverted-U hypothesis, suggesting that beyond a certain point, additional complexity lessens our enjoyment of music.

4.3.2.1 Brain's reward system

The feeling of pleasure we experience when listening to music we enjoy can be attributed to several factors related to the brain's processing of music and its effect on our reward and pleasure systems. Here are some key aspects:

- **Dopamine Release:** a neurotransmitter associated with the brain's reward and pleasure centers.
- **Emotional Response:** These emotional responses are closely linked to the brain's reward system. When music triggers positive emotions, it can reinforce our desire to listen to it again.
- **Pattern Recognition and Prediction:** The human brain is wired to recognize patterns and make predictions. Music often involves patterns and predictability, which can be satisfying when our expectations are met or pleasantly surprised. This aspect of music can engage the brain's reward system by providing a sense of achievement or surprise.
- **Social Connection:** Music often plays a role in social bonding and shared experiences, which can be rewarding.
- **Memory and Nostalgia:** Music can trigger memories and nostalgic feelings, which are closely linked to emotional responses. The ability of music to evoke vivid memories can be rewarding and emotionally fulfilling.

- **Physical Response:** Music can also induce physical responses, such as dancing or tapping to the beat, which can be enjoyable and contribute to the overall rewarding experience of music listening.

The brain's reward system is a key point in each of the aforementioned reactions. It is an intricate and essential network that plays a pivotal role in how we experience pleasure, motivation, and reinforcement learning. This system, crucial for survival and overall well-being, drives us towards behaviors beneficial for survival, such as eating, socializing, and reproducing. Delving deeper into the key components of the brain's reward system, particularly in the context of music listening, we find:

- **Ventral Tegmental Area (VTA):** Located in the midbrain, the VTA is a neural hub, often considered the starting point of the dopamine-mediated reward pathway. Its role in the reward system is paramount, primarily through the release of dopamine, a neurotransmitter intricately linked to feelings of pleasure and reward. During music listening, the VTA's activation leads to the release of dopamine, especially when the music is perceived as emotionally rewarding or pleasurable. This neurochemical release is a critical component of the enjoyable and sometimes euphoric sensations associated with music.
- **Nucleus Accumbens (NAc):** Situated in the basal forebrain, the NAc is a primary recipient of dopamine signals from the VTA. This area is integral to processing the rewarding aspects of stimuli, including music. The NAc acts as a sort of neural 'crossroads' where pleasure, motivation, and reinforcement converge. Its role in music appreciation is significant, as it processes the rewarding aspects of music, contributing to the desire to continue listening or to experience similar types of music again.
- **Hippocampus and Amygdala:** These structures play a crucial role in forming the emotional and memory aspects of rewarding experiences, such as listening to music.
 - The **hippocampus** (temporal lobe) is key in forming new memories, including those associated with music, such as remembering a melody or lyrics.
 - The **amygdala**, situated within the temporal lobes of the brain, is pivotal in emotional processing. It contributes to the emotional resonance of music, helping to explain why certain pieces of music can evoke strong emotional reactions. This emotional response can be both innate and learned, depending on cultural context and personal experiences. The amygdala's activity influence our music preferences and tastes.

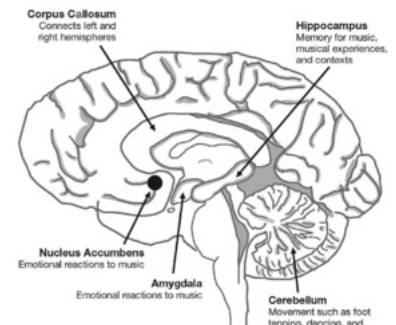


Figure 22

The interplay between these two structures is essential in creating lasting musical memories and the emotional ties that come with them, and in recalling musical experiences and anticipating music based on past encounters.

- **Dopamine Pathways:** The mesolimbic pathway, running from the VTA to the NAc, is critical in the reward system, particularly in the sensation of reward mediated by dopamine. This pathway is highly active during pleasurable experiences, including music listening, facilitating the feeling of pleasure and motivation. The mesocortical pathway, extending from the VTA to the prefrontal cortex, is involved in more cognitive aspects, such as motivation, emotional response, and decision-making related to music, such as choosing what music to listen to or making judgments about its quality or emotional content.

The brain's reward system also plays a significant role in learning and predicting rewards. When listening to music, especially familiar pieces, the brain often anticipates upcoming sections, crescendos, or choruses. This anticipation of pleasurable parts of the music can activate the brain's reward system, releasing dopamine in a way that not only rewards the current listening experience but also reinforces the memory and enjoyment of the music. This is part of the reason why we might feel compelled to listen to certain songs repeatedly.

4.4 STUDYING BRAIN RESPONSE

4.4.1 Electroencephalography (EEG)

EEG allows the observation of brain activity with high temporal precision, akin to a high-speed camera capturing the rapid fluctuations of neural responses to stimuli. While EEG excels in determining the "when" of brain activity, its resolution is akin to listening to the buzz (*brusio*) in a crowded room; it captures the general hum (*ronzio*) but cannot isolate individual conversations. Thus, it detects the collective activity of neurons rather than the specifics of single neuron firings. Despite these limitations, EEG's swift tracking of temporal patterns makes it invaluable for studying anticipatory processes in music perception, capturing how the brain reacts to and expects harmonic progressions.

4.4.2 Magnetoencephalography (MEG)

MEG measures the brain's magnetic fields resulting from electrical currents of neural activity. This technique is comparable to EEG in its temporal acuity but uses a different modality to sense the brain's subtle magnetic signals. MEG is particularly useful for its unique sensitivity and contributes to the temporal mapping of brain activity.

4.4.3 Functional Magnetic Resonance Imaging (Fmri)

fMRI provides insights into brain activity from another angle, detecting changes related to blood flow. This technique operates on the principle that cerebral blood flow is coupled with neuronal activation, increasing in regions of the brain that are actively engaged. By tracking these changes, fMRI effectively highlights the "where" of brain activity, offering a spatial map of neural engagement akin to a traffic report that indicates the busiest routes within a city.

Each method offers distinct advantages: EEG and MEG provide rapid updates on brain activity timing, while fMRI locates the activity. By integrating these techniques, researchers can overcome the limitations inherent in each.

4.5 BRAIN MEMORY MANAGEMENT THEORIES

Our brain has about 100 billion of neurons, but this does not explain its complexity, because it has no problem of "capacity", but rather to access information. Each neuron is usually connected to one thousand to ten thousand other neurons. The number of possible connections between n neurons grows exponentially with n^2 .

The fact that the brain's capacity is virtually unlimited makes it necessary to organize information so that it will be easily accessible and retrievable during its operational life. The information that we store is organised according to some models and it allows us to retrieve it in an efficient way. There are 3 main theories:

4.5.1 Constructivist theory (*relational*)

The constructivist theory posits (*postula*) that memories are constructed by the mind rather than recorded verbatim (*parola per parola*). According to this theory, each time we recall a memory, we reconstruct it using both the original information and additional input from our current state, which can lead to changes or distortions in the memory over time. Constructivist theory views memory as an interpretive process, where the brain actively works to make sense of information by fitting it into existing schemas or creating new ones.

4.5.2 Record-Keeping Theory

Contrary to constructivist theory, the record-keeping theory posits that the brain maintains detailed records of experiences. However, research suggests that while some memories can be vivid and detailed, they are not always accurate recordings of events.

4.5.3 Multiple Trace Theory (MTT)

Both the previously cited hypothesis have pros and cons, so scholars attempted to create a theory that brought together the most convincing sides of the two theories: the Multiple Trace Theory.

MTT posits that memories are dynamic rather than static. When we form a memory, it doesn't reside in a single, unchangeable form. Instead, it is subject to alteration every time it's retrieved. This process creates multiple, slightly different neural traces of the same memory. Each memory trace corresponds to a pattern of neural activity. These traces are distributed across various regions of the brain, especially within the hippocampus and related structures, which are critically involved in the consolidation of new memories.

When a memory is recalled, it becomes malleable and it can be updated, strengthened (*rafforzato*), or even weakened before being 'saved' again. This is a continuous process that occurs each time we remember something. When a memory is recalled and reconsolidated, the neural connections that constitute that memory can be strengthened. This is often why frequently recalled memories can become more vivid or durable over time.

MTT also incorporates the concept of prototype theory, which suggests that rather than having strict categories in our mind, our categorization is 'fuzzy'—meaning that some members of a category are more central (prototypical), while others are more peripheral. For example, when we think of a bird, a robin (*pettirosso*) might come to mind more readily than a penguin.

4.5.4 Mind vs. Brain – Common Illusions

Illusions confirm that our perception is biased by these models that we create from reality. When functioning properly, in fact, our perceptual system tends to “distort” the world that we perceive in order to make sense of it. We interact with the world around us through our senses, what we see must come to an agreement with our understanding of the world around us. Visual illusions are perhaps the most compelling proof of sensory distortion. Most of the time the information we receive at our sensory receptors is incomplete or ambiguous, and our brain's ability to make these identifications is nothing short of remarkable, considering that it starts out with whatever sensory receptors pass up to it.

4.6 BOTTOM-UP VS. TOP-DOWN PROCESSING

The concepts of bottom-up and top-down processing are fundamental in understanding how the brain interprets sensory information.

4.6.1 Bottom-Up Processing

This is a data-driven process that begins with sensory input. Imagine your brain as a complex, high-tech computer. When you experience the world, raw data comes in through your senses like microphones and cameras hooked up to the computer. This is the "bottom" where the processing starts. After your “sensors” kept the data, the brain's specialized systems start by picking apart the basic elements of this data. For music, this would be akin to an audio program identifying the different instruments, notes, and rhythms in a song. After the extraction, the brain combines these basic elements into a cohesive whole. Continuing with the music analogy, it's like taking those identified instruments, notes, and rhythms, and understanding them as a song with a particular genre or emotion.

4.6.2 Top-Down Processing

This is a concept-driven process that is guided by higher-level mental processes, like your brain's own expectations and prior knowledge. Imagine you're listening to music in a language you don't understand. Top-down processing uses your previous experiences and context to fill in gaps and anticipate what you're hearing. If you've heard similar music before, your brain predicts the rhythm or the flow of the music. It's like having a predictive text feature that guesses the next word you'll type based on what you usually say.

4.7 QUESTIONS FROM PREVIOUS EXAMS

- Q2.** Describe the concept of functional segregation in the brain. How is brain cortex partitioned based on specific functions? Are such functions at low or high level of abstraction? Are music listening and music cognition functionally relegated to specific areas of the brain cortex? If so, which and how?
- Q2.** Briefly explain the biochemical process behind neurotransmission. What is the role of neurotransmitters? What neurotransmitters are mainly involved in music perception and enjoyment and where are they mainly produced?
- Q2.** Describe the working principle of Functional Magnetic Resonance Imaging (fMRI) and explain what it shows of the brain's activity/response. What are its strengths and weaknesses? Where does it outperform an EEG (Electro-EncephaloGram)? Where does the EEG outperform the fMRI?
- Q2.** Briefly describe the role of the Ventral Tegmental Area and that of the Nucleus Accumbens in music enjoyment.
- Q2.** Briefly explain the biochemical process of neurotransmission. What is the role of neurotransmitters? What is "reuptake"? What are the most important neurotransmitters that are involved in music perception and enjoyment?
- Q2.** What regions of the brain are recruited when following along with music? Why?
- Q2.** Briefly explain Multiple Trace Theory as a memory consolidation model and comment on its relation with Constructivist (Relational) Theory and Record-Keeping Theory.

5 BUILDING BLOCKS OF TONAL ORGANISATION

5.1 KEYBOARD TYPES IN WESTERN MUSIC

While the piano keyboard (which we will talk about in a later paragraph) is the standard in western music, alternative structures like the Stradella or Uniform keyboards offer different approaches to tonal organization.

5.1.1 The Stradella Button Keyboard

The Stradella button keyboard is a type of layout used mainly on the left side of the accordion. It's designed to make it easier to play bass lines and chords to accompany the melody that is typically played with the right hand.

The Stradella layout is organized in rows. Each row corresponds to a specific chord or note. The buttons are arranged in a pattern known as the "circle of fifths," which is a musical concept that places notes and chords in a specific sequence that is commonly found in Western music.

Here's a simplified explanation using the *figure 23 (standard bass system)* as a reference:

- **Root Notes:** These are the base buttons that play the fundamental bass notes. You can recognize these buttons by the circle symbol, and they are typically used as starting points for finding your way around the keyboard.
- **Major Third Notes:** Located directly above the root note buttons, these play the major third note in relation to the bass note below them. These buttons are indicated by a triangle symbol.
- **Major Chords:** Above the major third notes, these buttons play a major chord based on the corresponding root note. The major chord consists of the root note, a major third, and a perfect fifth. These are indicated by the plus sign symbol.
- **Minor Chords:** These buttons, indicated by the minus sign, play a minor chord when pressed. A minor chord has a root note, a minor third, and a perfect fifth. These are indicated by the minus sign symbol.
- **Dominant 7th Chords:** The buttons with the number 7 symbol produce a dominant 7th chord, which adds a minor seventh note to the major chord.
- **Diminished 7th Chords:** Found at the top row, these buttons play a diminished 7th chord when pressed. This chord is built from a root note, a minor third, a diminished fifth, and a diminished seventh.

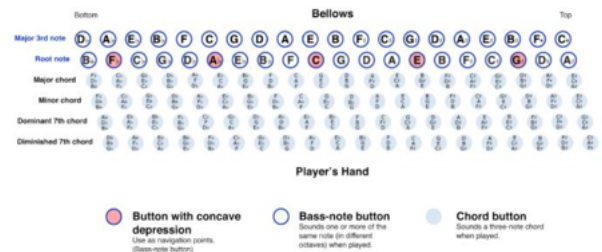


Figure 23

The Stradella button keyboard is common in Western music, particularly in genres that feature the accordion, such as polka, traditional folk music, and some forms of jazz and pop. It's a system that's been around for a long time and is widely used because it's relatively easy to learn and very versatile for accompaniment patterns.

5.1.2 The Uniform/Janko Keyboard

The uniform or Janko keyboard (*figure 24*) is an alternative design to the traditional piano keyboard layout: it is designed such that each row of keys is a whole tone apart from the next. This means that if you move from one key to the key directly next to it on the same row, you're moving by a whole tone interval. Moreover, if you move diagonally between rows, you're moving by a half tone, which is the smallest common interval in Western music.

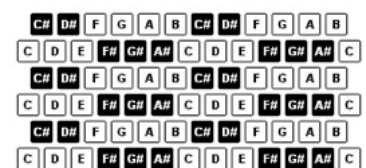


Figure 24

The advantage of this layout is that it allows for the same fingering patterns to be used for chords and scales regardless of the key you're playing in. This is different from the traditional piano keyboard, referred to here as the "classical Ionian keyboard," where the fingering for scales and chords changes depending on the key due to the irregular pattern of black and white keys.

However, despite its logical approach to keyboard design, the Janko layout has not been widely adopted in comparison to the standard piano keyboard.

5.1.3 The piano keyboard

In Western music, the piano keyboard (*figure 25*) serves as a fundamental framework for tonal organization. On a piano Ionian keyboard, the pitch of a note (P_{note}) is derived from a reference notes, for example C1, through the formula $P_{note} = P_{C_1} * 2^{\frac{n}{12}}$, where n is the distance from in term of semitones from the reference note (we use a base 12 because we are dividing an octave into 12 steps). The white keys can be interpreted by a combination of fifths, starting from F and going up a fifth until B.

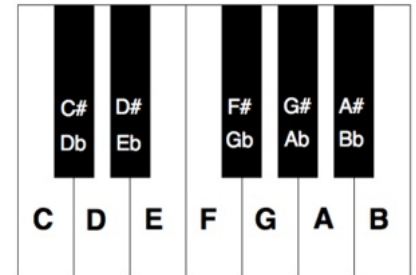


Figure 25

5.2 TONAL PROGRESSION OF PARTIALS

When a string is plucked or bowed, it doesn't just vibrate at a single frequency. Instead, it vibrates in multiple ways at once, producing what we call 'modes' or 'partials.' Each of these modes corresponds to a different frequency:

- the lowest frequency is known as the fundamental frequency, and it is usually the pitch that we perceive as the note being played.
- the other frequencies are higher and are multiples of this fundamental frequency, creating a series of harmonics, also known as overtones. These harmonics are integral multiples of the fundamental frequency, meaning if the fundamental frequency is 'f,' the second harmonic will be '2f,' the third will be '3f,' and so on.

The harmonic series is important in music because it explains why certain notes sound consonant or dissonant together; it's related to the physics of sound and how we perceive the combination of frequencies.

We can analyse this phenomenon looking to the figure 26: this image represent a diagram of the harmonic series as produced by a vibrating string, like that of a guitar or violin. The numbers on the left side (1 to 16) correspond to the modes or harmonics of the string vibration, starting from the fundamental frequency and going up to the 16th harmonic. Here's the detail of the image from the bottom up:

- **1st Harmonic (Fundamental):** The string vibrates in its entire length, producing the lowest pitch, noted as "C".
- **2nd Harmonic:** The string vibrates in two equal parts, producing a pitch one octave higher than the fundamental, also a "C".
- **3rd Harmonic:** The string vibrates in three equal parts, yielding (*producendo*) a pitch that is an octave plus a perfect fifth above the fundamental ("G").
- **4th Harmonic:** Produces another "C" two octaves above the fundamental.
- **5th Harmonic:** Results in an "E" that is two octaves plus a major third above the fundamental, with a -14 cents deviation.
- **6th Harmonic:** Yields a "G" again, now two octaves plus a perfect fifth above the fundamental, with a +2 cents deviation.

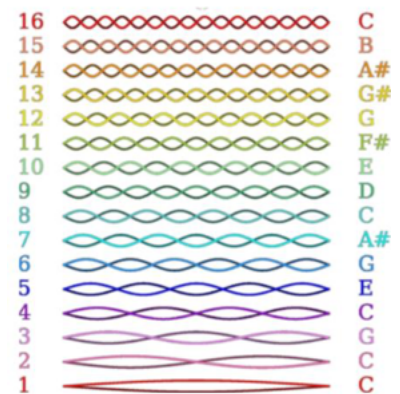


Figure 26

- **7th Harmonic:** Gives an "A#" (Bb), which is flat in equal temperament tuning, noted as -31 cents.
- **8th Harmonic:** Another "C", this time three octaves above the fundamental.

As we move up the series, the harmonics produce pitches that correspond to the notes in the series, but they may not perfectly align with the pitches of the equal-tempered scale. The waveforms shown next to the numbers represent the string's physical vibration mode that corresponds to each harmonic. The complexity of the waveform increases with each successive harmonic.

5.3 RELATIONS BETWEEN TONES

5.3.1 Intervals

Understanding the relationships between tones in music involves the study of intervals. An interval is the pitch difference between two tones. Intervals are foundational to the construction of scales and chords and can be classified as follows:

- The most fundamental interval is the perfect unison (0 semitones), which is the sound of the same pitch played together.
- **Perfect Intervals** include the octave, which spans 12 semitones and doubles the frequency of the starting pitch. Other perfect intervals are the perfect fifth (7 semitones) and the perfect fourth (5 semitones). These intervals are so named because of their consonant sound and stability in harmony.
- **Major Intervals** are considered to give a bright and stable sound but are less consonant than perfect intervals. These include Major second (2 semitones); Major third (4 semitones); Major sixth (9 semitones); Major seventh (11 semitones).
- **Minor Intervals** are one semitone smaller than their major counterparts, giving them a darker or more melancholic quality. They include Minor second (1 semitone); Minor third (3 semitones); Minor sixth (8 semitones); Minor seventh (10 semitones).

Number of semitones apart	Music interval	Notes
1	Minor 2nd	C - C#
2	Major 2nd	C - D
3	Minor 3rd	C - D#
4	Major 3rd	C - E
5	Perfect 4th	C - F
6	Augmented 4th / Diminished 5th	C - F#
7	Perfect 5th	C - G
8	Minor 6th	C - G#
9	Major 6th	C - A
10	Minor 7th	C - A#
11	Major 7th	C - B
12	Octave	C - C

Figure 27

Between the cited interval, the one that creates the bigger amount of tension and instability is the **tritone** (figure 28), which is an interval spanning 6 semitones (augmented fourth or diminished fifth). One of the most interesting curiosities about the tritone is its historical nickname "diabolus in musica," which translates to "the devil in music." During the Middle Ages and into the Renaissance, the tritone was avoided in musical composition due to its dissonant sound, which was unsettling in the context of the harmonic conventions of the time. Its use was discouraged or even banned by the church in liturgical music.



Figure 28

Another interesting concept to grasp regarding intervals between notes is that of **inversion**: when you invert an interval, the note that was lower becomes higher, and the note that was higher becomes lower. This process of inversion changes the type and size of the interval, but there are consistent patterns to these changes.

One key rule of interval inversions is that the sizes of the original interval and its inversion will always add up to 9 when you count the number of steps (scale degrees, not semitones!) between the two notes. For example, if you have a fifth (which spans 5 steps), its inversion would be a fourth (which spans 4 steps), and together they add up to 9 (5 + 4 = 9).

Here's what happens to different types of intervals when they are inverted:

- **Major intervals become minor:** When you invert a major interval, it becomes a minor interval. For example, a major third (which spans 4 half-steps or semitones) becomes a minor sixth (which spans 8 half-steps) when inverted.
- **Minor intervals become major:** Conversely, inverting a minor interval turns it into a major interval. A minor third (3 half-steps) becomes a major sixth (9 half-steps) upon inversion.
- **Perfect intervals remain perfect:** Inversions of perfect intervals (like the perfect fourth or perfect fifth) also result in perfect intervals. For instance, a perfect fifth inverts to a perfect fourth.
- **Augmented intervals become diminished:** If you invert an augmented interval (one that is one half-step larger than a perfect or a major interval), it becomes a diminished interval (one that is one half-step smaller than a perfect or a minor interval), and vice versa.

5.3.2 Triads

Triads are the simplest form of chords in music and consist of three notes. These three notes are created by stacking two intervals of thirds on top of one another from a root note. The types of thirds used (major or minor) define the type of triad:

- **Major Triad:** Formed by a major third and then a minor third. The combination gives a sound that is often described as happy, strong, or triumphant.
- **Minor Triad:** Composed of a minor third followed by a major third, producing a sound that can be perceived as sad, pensive, or tragic.
- **Diminished Triad:** Two minor thirds stacked together, creating an unstable or tense sound.
- **Augmented Triad:** Consists of two major thirds, resulting in a unique, floating, or misty quality.



Figure 29

Each triad gives a distinct emotional character to the music and is fundamental to harmony.

We can also Inverting triads, changing the order of the notes so that a different note is the lowest pitch. Triads can be inverted in two ways:

- **First Inversion:** The third of the triad is now the lowest note.
- **Second Inversion:** The fifth of the triad is now the lowest note.

Inversions change the character of the triad without altering its fundamental quality.

5.3.3 Tetrads

Tetrads, commonly known as seventh chords, are a type of chord in music that consist of three specific intervals stacked on top of each other. These intervals are built on the root note of the chord and typically include:

- A third interval (either major or minor)
- A fifth interval (either perfect, diminished, or augmented)
- A seventh interval (either major or minor)

By stacking these intervals, we create various types of tetrads:

- **Major 7th:** Composed of a major third, perfect fifth, and major seventh. Also known as a 'major-major seventh' or 'delta chord', it is characterized by a rich, full sound.
- **Minor 7th:** Formed by a minor third, perfect fifth, and minor seventh. This chord has a more subdued, mellow quality.
- **Dominant 7th:** Has a major third, perfect fifth, and minor seventh. It's known for its strong resolution tendency, often leading to the tonic chord in a progression.
- **Half-Diminished:** Built with a minor third, diminished fifth, and minor seventh, creating a tense, unresolved sound.
- **Diminished 7th:** Consists of a minor third, diminished fifth, and diminished seventh (equivalent to a major sixth). This chord is very tense and unstable.
- **Augmented 7th:** With a major third, augmented fifth, and major seventh, it's also referred to as a 'major-minor seventh' chord.

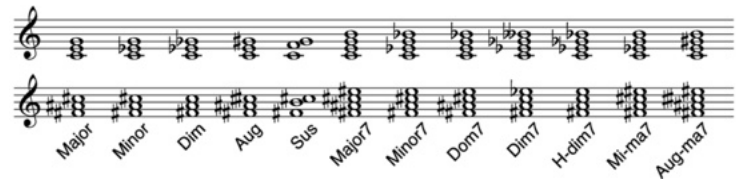


Figure 30

In the same way that triads can be inverted by changing the bass note, tetrads can also be inverted to create different sounds and harmonic functions. Additionally, voicings of tetrads, which involve spreading out the chord notes in various octaves, can create more interesting textures and lead to smoother transitions between chords in a progression.

5.3.4 Beyond Tetrads

While tetrads are a fundamental aspect of harmony, musicians can stack even more intervals to create extended chords, such as pentads (ninth chords), hexads (eleventh chords), heptads (thirteenth chords), and beyond. These extended chords are used to add tension and colour to the music, providing a rich palette for composers and performers.

5.3.5 Scales

A scale in music is an ordered sequence of notes that provides a framework that dictates how the music is structured harmonically and melodically. While there are 12 notes in the chromatic scale, most musical pieces do not utilize all 12. Instead, they employ a subset of these notes to create a particular scale, which imposes a structure on the music and defines its character. The key elements of a scale are:

- The **Tonic:** This is the 'home' note of the scale, where music in the key of the scale feels resolved or at rest.
- The **Scale Choice:** The selection of a scale determines the set of pitches used in the music and is chosen in relation to the tonic. For instance, in the key of C major, the C major scale is used, and it consists of the notes C, D, E, F, G, A, and B.

Accidentals (*alterazioni*), such as sharps and flats, are notated in music to indicate any alterations from the natural notes of the scale. These are strategically placed within the musical staff to ensure that only the pitches within the chosen scale are played. The number of accidentals is determined by the key signature, which harmonizes the composition.

5.3.6 Modal Scales and the Ionian Mode

Modal scales, also known as modes, are variations of the major scale (Ionian mode) that are created by starting the scale on a different note of the Ionian mode. Each mode has a unique sequence of whole and half steps, which gives it a distinct sound and emotional character:

1. Ionian Mode (major scale): The same as the major scale, with a bright and stable sound ([ex](#)).
2. Dorian Mode: Begins on the second note of the Ionian mode, offering a slightly darker, more minor sound ([ex](#)).
3. Phrygian Mode: Starts on the third note, known for its Spanish flamenco-like quality ([ex](#)).
4. Lydian Mode: Initiated from the fourth note, it has a brighter, major sound due to the raised fourth ([ex](#)).
5. Mixolydian Mode: Commences on the fifth note, similar to the major scale but with a lowered seventh, giving it a bluesy feel ([ex](#)).
6. Aeolian Mode (natural minor scale): Also known as the natural minor scale, begins on the sixth note and has a minor, sad quality ([ex](#)).
7. Locrian Mode: Starts on the seventh note and is characterized by a diminished fifth, creating a very dark and unstable sound ([ex](#)).



Figure 31

By shifting the starting point of the Ionian mode to different notes within the scale, we essentially rotate the pattern of intervals, creating the different modes. The characteristic quality of a mode, whether it sounds major or minor, is determined by the third interval from the root note.

We can see the translational transformation also as a bit rotation, as showed in *figure 32*:

	0	1	2	3	4	5	6	7	8	9	10	11	
Ionian	1	0	1	0	1	1	0	1	0	1	0	1	Major
Dorian	1	0	1	1	0	1	0	1	0	1	1	0	Minor
Phrygian	1	1	0	1	0	1	0	1	1	0	1	0	Minor
Lydian	1	0	1	0	1	0	1	1	0	1	0	1	Major
Mixolydian	1	0	1	0	1	1	0	1	0	1	1	0	Major
Aeolian	1	0	1	1	0	1	0	1	1	0	1	0	Minor
Locrian	1	1	0	1	0	1	1	0	1	0	1	0	Minor
Ionian	1	0	1	0	1	1	0	1	0	1	0	1	Major

Figure 32

5.3.7 Non-Seven-Note Scales

Music across different cultures uses various scales that do not conform to the seven-note structure. We can classify scales by the number of notes they have:

- Ditonic: A two-note scale, which can be viewed through the lens of intervals rather than a scale with a full range.
- Tritonic: A three-note scale, often seen in primitive musical forms.

- Tetratonic: A four-note scale, less common but found in some cultural music traditions.
- Pentatonic Scale: Consists of five notes and is prominent in folk music, especially in Asia. It's anhemitonic, meaning it lacks semitones, and corresponds to the black keys on a piano ([ex](#)).
- Blues Scale: A variation of the pentatonic scale that includes additional notes to create the signature blues sound ([ex](#)).
- Hexatonic Scale: Contains six notes and is often found in Western folk music ([ex](#)).
- Octatonic Scale: This eight-note scale is used in jazz and modern classical music and is characterized by alternating whole and half steps, creating a symmetric pattern ([ex](#)).

5.3.8 The Circle of Fifths

The [Circle of Fifths](#) ([figure 33](#)) is a visual representation that musicians use to understand and remember the relationships between different keys. It shows all 12 notes of the chromatic scale and arranges them in a circle where each note is a perfect fifth above the previous one. The Functions of the Circle of Fifths are:

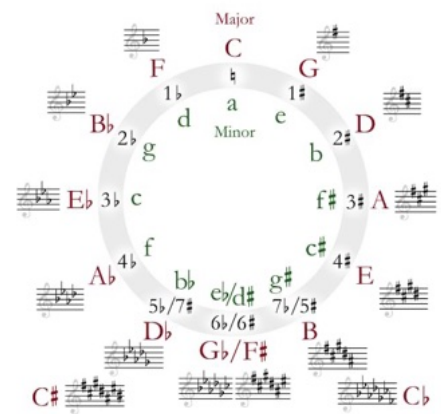


Figure 33

- **Key Identification:**
 - **moving clockwise** (by fifth steps) adds sharps to the key signature. The position of the “new sharp” is the seventh, starting from the current tonic note (es. 7th of Re = Do → the Re tonality has a Do# + the sharps inherited (ereditati) from the previous tonalities in the cycle).
 - **moving counterclockwise** (by fifth steps) adds flats to the key signature. The position of the “new flats” is the fourth, starting from the current tonic note (es. 4th of Si = Mi → the Si key has a Mi_b + the flats inherited from the previous tonalities in the cycle).
- **Modal Relationships:** it helps in identifying the major and relative minor keys that share the same key signature. The relative minor of a certain key is the key located a minor third lower (ex. The relative minor of Do is La).
- **Compositional Tool:** the Circle of Fifths can be used to create chord progressions that have a strong sense of movement and resolution, as adjacent keys on the circle are closely related harmonically.

The Circle of Fifths is also a helpful tool in understanding how each mode relates to the others and how changing one note affects the mode. By using the Circle of Fifths, we can order the modes from the brightest to the darkest and notice that the Hamming distance of each neighbor mode is 2, which basically means that 2 tones change ([figure 34](#)).

	0	1	2	3	4	5	6	7	8	9	10	11	
Lydian	1	0	1	0	1	0	1	1	0	1	0	1	Major
Ionian	1	0	1	0	1	1	0	1	0	1	0	1	Major
Mixolydian	1	0	1	0	1	1	0	1	0	1	1	0	Major
Dorian	1	0	1	1	0	1	0	1	0	1	1	0	Minor
Aeolian	1	0	1	1	0	1	0	1	1	0	1	0	Minor
Phrygian	1	1	0	1	0	1	0	1	1	0	1	0	Minor
Locrian	1	1	0	1	0	1	1	0	1	0	1	0	Minor

Figure 34

5.4 QUESTIONS FROM PREVIOUS EXAMS

Q4. Describe how intervals and their inversions relate to each other and how intervals can be stacked together to form triads and related inversions. Explain how to obtain the various types of chords (and mood qualities) through such combinations. Do you have any suggestions on how to generalize this to quadriads?

Q4. Define and describe tonal intervals, and specify when they are major, minor, tritone, and perfect. How do you choose and order two intervals to stack on top of each other to form major, minor and diminished triads? How do you choose and order 3 intervals to stack on top of each other to form major (delta), dominant, minor 7, half-diminished and diminished tetrads?

Q4. If you select a series of major scales starting from notes that follow the ascending cycle of fifths, how does the corresponding key signature changes? What if you follow a descending cycle of fifths? How do you determine which sharps or flats will appear on the key signature?

Q4. How is the Janko Keyboard (uniform keyboard) organized and what are its main features and properties?

Q5. What is the key signature and how does the “Circle (or Cycle) of fifths” help us find it (and the number of sharps or flats in it)? Please use a graphic representation.

6 RHYTHMIC STRUCTURES AND MODELS

6.1 THE CONCEPT OF RHYTHM

The concept of rhythm, although complex and somewhat subjective, has roots tracing back to the hominid era. Fundamentally, rhythm is based on a beat, often exemplified by natural rhythms such as the pace of walking or the human heartbeat, which typically ranges between 60 and 100 beats per minute (BPM).

To understand rhythm, it's essential to identify the 'pulse' or 'tactus' or 'beat.' This element is a sequence of consistent, distinct stimuli of short duration. It's important to note that the pulse doesn't necessarily represent the fastest or slowest component of the rhythm but is perceived as the fundamental element. This concept of a metric level is crucial in dance, poetry, and music, where there's an effort to maintain it consistently.

Once the beat is established, rhythm involves divisions and subdivisions within a measure. For instance, in a measure with four beats, each quarter note (beat) can be further divided into eighth and sixteenth notes, referred to as division levels. Additionally, there are multiple levels like half notes or whole notes. The periodicity of a piece of music also includes stressed and unstressed beats, and the alternation between these creates a sense of movement.

Syncopation is another aspect of rhythm, where the placement of upbeats and downbeats is shifted from the original rhythm, adding complexity. Even in intricate rhythms, the meter can be broken down into duple (two beats per measure) and triple (three beats per measure) pulses, either by adding to or dividing the basic beat structure. This breakdown helps in understanding and appreciating the complexity and nuances of rhythm in music.

6.2 MUSICAL METER

In the study of music, rhythm plays a crucial role and is structured in a hierarchical fashion across various levels or time scales. This structure, known as musical meter, encompasses three distinct levels:

1. **Tactus or Beat:** This is the fundamental level, akin to the rate at which one might tap their foot to the music. It is the steady pulse that underlies the rhythm and is usually measured in beats per minute (BPM). For a rhythm to be coherent, the tempo, which is the rate of the tactus or beat, should have a slight variation over time.
2. **Tatum:** This represents the shortest durational values in music that are still more than incidentally encountered. It is the grid on which all rhythmic events are based, with every duration being an integer multiple of the tatum period. This ensures that the rhythm fits within the subdivisions of the beat.
3. **Measure:** Related to the harmonic changes or the length of a rhythmic pattern, the measure is a segment of time defined by a given number of beats. Each beat within the measure is represented by the tactus.

The tempo of a piece is the pace at which it progresses; this is the speed of the tactus pulse.

6.2.1 Difference between Rhythm and Metric

- **Rhythms:** A rhythm in music refers to the pattern of sounds and silences in time. The term "complex rhythms" typically refers to rhythms that are more intricate, often involving syncopation or irregular patterns.
- **Meters (Time Signatures):** A meter, or time signature, in music indicates how many beats are in each measure and what note value constitutes one beat.

6.3 METRIC STRUCTURE

The metric structure in music refers to how beats are organized into regular, repeating units known as measures or bars. This structure is fundamental to the rhythm and timing of a piece of music. Here's a breakdown of the key components:

- **Beat:** The beat is the basic unit of time in music, a regular pulse that you would tap your foot to when listening to a song. It's what gives music its sense of rhythm and pace.
- **Time Signature:** The metric structure is often indicated at the beginning of a piece by a time signature. The time signature consists of two numbers, one on top of the other, like a fraction. The top number indicates how many beats are in each measure. The bottom number indicates the note value that represents one beat. For example, in a 4/4 time signature, there are 4 beats per measure, and the quarter note gets one beat. In a 3/4 time signature, there are 3 beats per measure, and again the quarter note gets one beat.
- **Measures/Bars:** Measures, or bars, are the segments in written music that provide a visual structure to the rhythm. Each measure contains a set number of beats, as determined by the time signature. The boundaries of each measure are marked by vertical lines called bar lines.

There are three types of metric structures in music:

- **Metrical Rhythm (Divisive Rhythm):** This structure is where each time value is a multiple or fraction of the beat, with accents that periodically recur, providing systematic grouping into measures ([ex1](#) , [ex2](#)).
- **Measured Rhythm (Additive Rhythm):** Similar to metrical rhythm, but in this case, time values are multiples or fractions of a specified time unit, and the accents do not necessarily occur at regular intervals within the cycle ([ex](#)).
- **Free Rhythm:** This is a more flexible structure that does not conform to periodicity. An example of free rhythm is found in Christian chants, which have a basic pulse but do not follow a strict metric structure ([ex](#)).

FIGURE MUSICAL	Valore	PAUSE
Semibreve	4/4	SEMIBREVE
Minima	2/4	MINIMA
Seminima	1/4	SEMINIMA
Croma	1/8	CROMA
Semicroma	1/16	SEMICROMA
Biscroma	1/32	BISCROMA
Semibiscroma	1/64	SEMIBISCROMA

Figure 35

6.4 METRIC MODULATION

[Metric modulation](#) in music is a compositional technique involving a change in the pulse rate (tempo) and rhythm of a piece through the use of a common rhythmic bridge ([ex.](#)). It's a sophisticated method of changing the meter (time signature) where the new tempo is related to the old one by some subdivision of the basic pulse. Unlike a simple tempo change, this modulation is achieved by reinterpreting a common rhythmic value in both the original and new tempo. For instance, a set of eighth notes in one tempo might be equated to triplet eighth notes in another.

The composer usually indicates this change by writing a specific equation-like notation. For example, an eighth note equals a triplet eighth note, indicating that the tempo should change so that what was previously played as an eighth note is now played with the duration of a triplet eighth note.



Figure 36

This technique creates a smooth and often surprising transition between different rhythmic feels. It can add complexity and interest to the music, changing the listener's perception of the beat and the pacing of the piece.

Metric modulation is often used in modern classical music and some genres of contemporary music, like progressive rock, [metal](#) ([amazing song](#)) or jazz. It's particularly favoured for its ability to seamlessly bridge contrasting rhythmic sections.

6.4.1 Simplification of Periodic Metric Modulations

When rhythmic modulations occur in a periodic and regular fashion, it is indeed possible to notate them under a single global rhythmic signature. This is achieved by using a composite meter that can accommodate the different rhythmic groupings. For example, if a piece alternates between 3/4 and 4/4 time signatures, one might consider consolidating this into a repeating 7/4 time signature, where the pattern of beats might alternate between groupings of 3 and 4 (or vice versa).

6.5 POLYRHYTHM AND POLIMETER

Polyrhythm and polymeter are both musical concepts related to rhythm, but they refer to different aspects of rhythm and have distinct characteristics:

- **Polyrhythm:** Polyrhythm involves the layering of multiple rhythmic patterns or subdivisions within a single meter or time signature. It is characterized by having different rhythms occurring simultaneously but within the same overall beat or pulse.
- **Polymeter:** Polymeter is a more general concept that can include polyrhythm as one of its elements. It involves the simultaneous use of multiple meters or time signatures in a musical composition or performance. In polymeter, different sections or instruments may be in distinct meters, and these meters may have their own separate beats or pulses. Polymeter can create a sense of tension and rhythmic complexity in music, as listeners may hear conflicting time signatures happening simultaneously.

In the next paragraphs we will analyze these two concepts in detail.

6.6 POLYRHYTHM

6.6.1 Polyrythms

A [polyrhythm](#) is the simultaneous combination of contrasting rhythms in a musical composition. More specifically, it refers to the presence of two or more independent rhythms that are played at the same time. Polyrhythms can be distinguished from irrational rhythms (tuplets, *figure 37*), which can occur within the context of a single part (*this is the difference between “normal” polyrhythms and cross-rhythms*). If the notes from different parts come together at irregular intervals or create a syncopated pattern that suggests a combination of different rhythmic cycles, then you are likely looking at a polyrhythm.



Figure 37

6.6.2 Hemiola

An hemiola is a specific type of polyrhythm in which two bars in a simple triple time (such as 3/4) are articulated as if they were three bars in a simple duple time (such as 2/4). The hemiola typically creates a temporary shift in the perceived metrical structure. It does not change the actual time signature but alters how the rhythm is perceived within the existing meter. You can see an hemiola example in *figure 38* (you can also [listen to it](#)).

Hemiola Example

Mozart Piano Sonata K332



Figure 38

6.6.3 Cross-Rhythm

[Cross-rhythm](#) is a specific kind of polyrhythm where the rhythmic conflict or interplay is the main feature. It involves the systematic layering of rhythms so that they cross each other's natural boundary beats. In cross rhythms, the contrasting rhythms often share a common unit or pulse divided in different ways (*figure 39*).

For instance, a common type of cross rhythm is the "three against two" ([3:2](#)) rhythm, where one rhythm plays three beats in the same time span that another plays two beats.

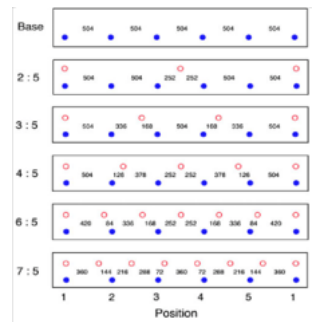


Figure 39

6.7 POLIMETERS

As we have seen in previous paragraphs, polimeters involve the simultaneous use of multiple meters or time signatures in a musical composition or performance

6.7.1 Signature Simplification for Polimeters (Polimeters in "Standard" Western Notation)

When reading music that contains multiple time signatures, it can be difficult to interpret the rhythms. To simplify this, a formula is used to combine two rhythms into a single ratio. The unification is done following the steps below:

- 1) We start from the two original rhythms: $\frac{N_1}{D_1} : \frac{N_2}{D_2}$.
- 2) We can apply the least common multiple (LCM) to the denominator: $\frac{N_1 * K_1}{D} : \frac{N_2 * K_2}{D}$.
- 3) And then we can apply the LCM also at the numerator: $\frac{N}{D}$.

$$\frac{N_1}{D_1} : \frac{N_2}{D_2} \rightarrow \frac{N}{D}$$

Figure 40

6.7.2 Polimeters in Brian Ferneyhough's Notation (BFN)

When the polymetric structure is too complex to be simplified, we can use the BFN. The BFN uses a fractional notation to represent the relationship between two different rhythmic layers or meters. The notation $\frac{m}{n} : \frac{i}{j}$ represents two separate rhythmic layers, where:

- The fraction $\frac{m}{n}$ indicates that the first rhythm, *R1*, consists of ***m* beats** in a measure, in which each beat has the **duration of $\frac{1}{n}$** of a whole note. For instance, if *R1* is notated as $\frac{4}{4}$, it means there are four quarter-note beats in one measure.
- The fraction $\frac{i}{j}$ signifies that the second rhythm, *R2*, consists of ***i* beats** per measure, in which each beat has the **duration of $\frac{1}{j}$** of a whole note. For instance, if *R2* is notated as $\frac{7}{8}$, it implies there are seven eighth-note beats within the same duration as one measure of *R1*.

The BFN also indicates the repetition cycles of each rhythm. *R1* repeats every four of its own beats, and *R2* repeats every seven of its own beats. This creates a complex interlocking pattern where **the two rhythms coincide only after a specific number of cycles**.

6.7.2.1 Special case from a past exam: BFN for polyrhythms

It's important to keep in mind that, usually, BFN is used to notate polimeters and not polyrhythms (as we have seen [here](#), *Polimeters involve multiple meters occurring simultaneously, while polyrhythms involve multiple rhythms occurring within the same metric framework*). However, even though it may not make much sense, we can use the BFN to describe a simple polyrhythm.

For a polyrhythm of *K* against *N*, the BFN would be $\frac{K}{K} : \frac{N}{N}$. This indicates that each rhythm plays in its own time, with *K* beats in a measure for *R1* and *N* beats in a measure for *R2*, each beat being a fraction ($1/K$ or $1/N$) of a whole note. For instance, a 3:2 polyrhythm can be expressed as $\frac{3}{3} : \frac{2}{2}$.

6.8 ALGEBRAIC STRUCTURE OF RHYTHMS

Rhythms can be examined through the lens of [algebraic structures known as groups](#). A group in this context is a set of rhythmic elements combined according to four specific properties:

- **Closure:** This property states that if you take any two rhythmic elements within a set (for instance, two different rhythmic patterns or beats), and combine them (e.g., play one after the other), the result must also be a rhythmic element that is part of the same set. This means the set of rhythms is "closed" under the operation of combination; you can't produce a rhythm that isn't already included in the set by combining rhythms from within the set.
- **Associativity:** Associativity ensures that when combining more than two rhythmic elements, the order in which the combinations occur does not change the final result. Mathematically, this means if you have three rhythms A, B, and C, then $(A \oplus B) \oplus C$ will yield the same rhythmic pattern as $A \oplus (B \oplus C)$. This property is crucial for the structure of musical compositions, where multiple rhythms can be layered or sequenced in various ways.
- **Identity:** The identity property asserts the existence of a unique rhythm in the set that, when combined with any other rhythm, leaves the other rhythm unchanged. In music, this could be interpreted as a basic pulse or a measure of silence which, when added to a rhythm, doesn't alter the inherent pattern or timing of that rhythm.
- **Invertibility:** For each rhythm in the set, there must be another rhythm that acts as its inverse. When a rhythm is combined with its inverse, the result is the identity element of the set. In a musical sense, this could be a rhythm that "negates" or "balances out" another to bring the music back to a neutral state, such as a rest or a return to the basic pulse.

A **group** is a collection of objects G , together with one operation $\oplus$, which has the following properties:

- **Associativity:** $a \oplus (b \oplus c) = (a \oplus b) \oplus c$
- **Identity:** There is an element $e \in G$ such that $e \oplus g = g \oplus e = g$ for all $g \in G$
- **Inverse:** For every $g \in G$ there exists $g^{-1} \in G$ such that $g \oplus g^{-1} = g^{-1} \oplus g = e$.

In music, these groups represent temporal structures and can be visualized using notations or graphical representations such as rhythmic wheels.

6.8.1 Substructures and Symmetry in Polyrythms

In the context of rhythms analyzed through the framework of algebraic structures, the concepts of subgroups and cosets are used to describe more nuanced (*sfumato*) rhythmic patterns, such as those found in polyrythms. Let's delve into a more detailed explanation:

- **Subgroups:** A subgroup is a smaller set of elements within a larger group that itself forms a group, under the same operation as the original group. In terms of rhythm, if you have a set of rhythmic patterns that form a group, a subgroup would be a subset of these patterns that also follows the group's properties (closure, associativity, identity, and invertibility). This means that the rhythms in the subgroup can be combined in various ways and still produce rhythms that are part of the subgroup, essentially functioning as a standalone rhythmic entity within the larger structure.
- **Cosets:** [Cosets](#) are related to subgroups but with a twist—literally. Imagine taking all the rhythms in a subgroup and shifting them in time; that's essentially what a coset is. You're not changing the rhythms themselves, but you're displacing them, so they start at a different point in the measure. This shift can cause the rhythms to align differently with the main beat, often leading to syncopation, which is the emphasis on beats that are not usually accented in the standard meter. In rhythmic terms, if you start every rhythm in the subgroup on the off-beat instead of the downbeat, you've created a coset.
- **Normal Groups and Symmetry:** A group is said to be "normal" if all of its cosets (both left and right cosets) are the same. This means that no matter how you shift the subgroup's elements (rhythms), you end up with the same set of patterns. In music, this would lead to a symmetrical polyrhythm where the shifted patterns (cosets) and the original patterns (subgroup) align in a way that feels balanced and regular. However, in many musical applications, the groups are not normal, leading to asymmetrical and complex rhythms. The shifting (cosets)

does not result in the same patterns, which can produce the irregular and intriguing rhythms characteristic of polyrhythmic music.

6.9 THE RHYTHMIC WHEELS

The [rhythmic wheel](#) (*figure 41*) is a visual aid designed to simplify the understanding of rhythms in music. It's particularly valuable for those tackling complex rhythmic patterns in composition, performance, or education. The tool is structured as follows:

- **Circular Design:** The wheel is circular, symbolizing a single measure of music. This circle is equally segmented, with each segment representing a specific rhythmic duration, akin to the numbers on a clock face.
- **Subdivisions:** These segments depict various rhythmic values, from simple ones like whole notes to complex ones like triplets or quintuplets. Each segment visually represents the duration of a note or rest in the musical measure.

By placing notes and rests on the wheel according to their durations, users can visually decode how different rhythmic elements interact (*figure 42*). This setup is particularly effective for understanding complex concepts such as syncopation and polyrhythms.

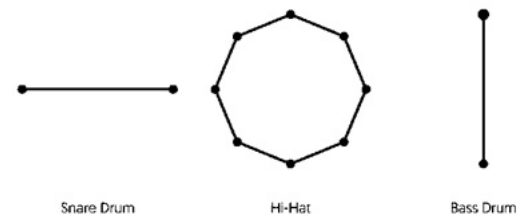


Figure 41

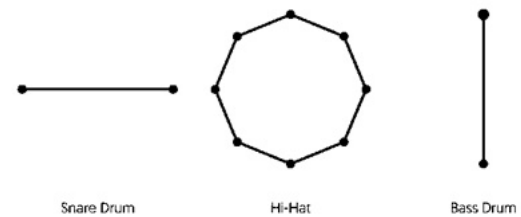


Figure 42

However, certain limitations exist with the rhythmic wheel, particularly in representing non-standard or highly complex rhythms. For example:

- **Irregular Time Signatures:** Time signatures that don't conform to standard patterns (like 4/4 or 3/4) can be challenging to represent. Examples include 5/4, 7/8, or other asymmetrical time signatures where beats per measure are not easily divisible into equal parts.
- **Changing Meter:** Music that frequently changes time signatures can be difficult to depict on a single rhythmic wheel, as each segment is typically fixed to represent a specific time division.
- **Highly Complex Rhythms:** Rhythms involving intricate subdivisions, such as nested tuplets (a triplet within a triplet, for instance), or rhythms that stretch over multiple measures in non-repetitive patterns, may not be easily visualizable on the wheel.

6.10 MODERN WESTERN NOTATION & BRIAN FERNEYHOUGH'S NOTATION: INSIGHTS

As we have seen in a previous paragraph, in Western music notation the metric structure is notated at the beginning of a piece of music and consists of two numbers, one on top of the other. The top number indicates how many beats are in a measure, and the bottom number indicates the note value that represents one beat. This system of notation allows musicians to understand the rhythmic framework of a composition. However, it has some limitations, for example:

- **Complexity in Irregular Rhythms:** Time signatures are less adept at handling complex, changing, or irregular rhythms which might require frequent changes in the time signature, making the music more difficult to read.
- **Polymeters and Polyrythms:** When multiple rhythmic layers are present (polymeters or polyrythms), standard notation can become cluttered and complex, making it difficult for the performer to read.

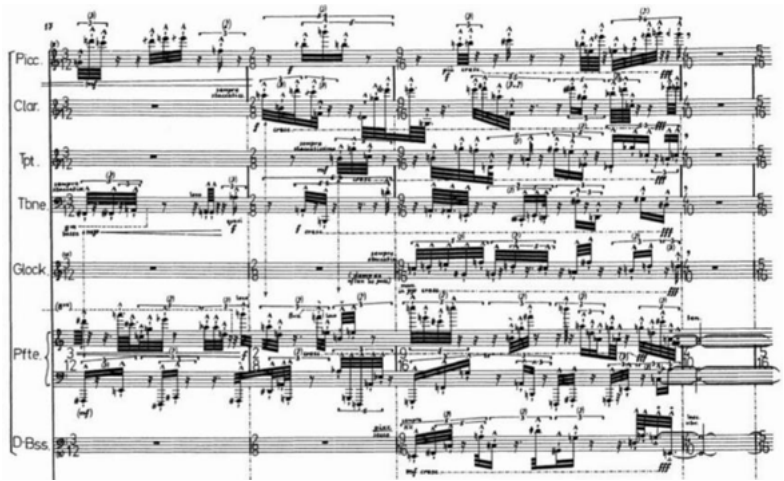


Figure 43

To overcome these problems, it is possible to use the **Brian Ferneyhough notation** (figure 43).

It is not a distinct language separate from traditional Western musical notation, but rather an extension and intensification of it. Here are some key characteristics:

- **Extreme detail and precision:** Ferneyhough's scores often contain an extraordinary amount of detail. This includes specific instructions for dynamics, articulation, and timing, often with a level of precision that goes beyond what is typical in standard notation.
- **Time Signatures and Rhythmic Groupings:** In standard notation, time signatures are used to organize music into consistent, repeating patterns of strong and weak beats. Ferneyhough's compositions frequently employ constantly changing time signatures and complex rhythmic groupings that do not adhere to traditional patterns, making the sense of a regular beat or pulse often elusive.
- **Tuplets and Polyrythms:** While standard notation uses tuplets (like triplets or quintuplets) to indicate a different number of notes played in the space of a given beat, Ferneyhough takes this to an extreme level. He often uses nested tuplets (figure 44) or complex polyrythms, where multiple layers of different rhythmic divisions occur simultaneously, making the rhythm extraordinarily complex.
- **Microtonality:** His use of microtones (intervals smaller than a semitone) requires additional notation symbols beyond the standard Western chromatic scale.
- **Microtiming:** Ferneyhough's notation sometimes includes microtiming - very subtle variations in timing that go beyond the standard divisions of beats in Western notation. This can involve fractions of standard rhythmic values, demanding an extremely high level of precision and control from the performer.
- **Graphic Elements and Interpretation:** Standard notation tends to be quite prescriptive, with a clear indication of pitch, duration, and timing. Ferneyhough's notation, on the other hand, can include more graphic elements or ambiguous symbols that require a high degree of interpretation from the performer. The notation might suggest a range of possibilities rather than a single, defined outcome.



Figure 44

However, it's important to note that the complexity of Ferneyhough's notation can also be seen as a disadvantage. It may be exceedingly difficult to interpret and perform, potentially limiting the accessibility of his music both to performers and audiences.

6.11 RHYTHM AND METER IN POETRY

Poetry shares similarities with music, particularly in its rhythmic structure which alternates between stressed and unstressed syllables. This rhythmic pattern often influences the melody and harmony in songs. In poetry, this rhythm is measured using a unit called a 'foot', which consists of one stressed syllable and at least one unstressed syllable. Common types of feet include two-syllable forms like iambic, Trochaic, and Spondaic, and three-syllable forms like Anapaestic and Dactylic. Poetic lines are categorized based on the number of feet they contain, such as dimeter, trimeter, etc. For instance, an iambic pentameter line contains five iambic feet, totalling ten syllables. While rhythms are often periodic, they can be visually represented in a "rhythmic wheel". However, in drum machines or software, a box system is used for representation, as it's simpler to display rhythms in a rectangular format than a circular one.

6.12 QUESTIONS FROM PREVIOUS EXAMS

Q3. Describe what rhythmic modulation is and explain its role in music enjoyment. Assuming that rhythmic modulations occurs in a periodic and regular fashion, is there a way to consolidate the rhythm under a single global rhythmic signature? Please explain with an example based on two alternating signatures.

Q3. List, define and describe the basic elements that constitute the meter in rhythms. Explain what is often used to create rhythmic surprises at both metric (mensural) and structural level. Explain what metric modulation is in the construction of rhythmic structures, why is it used and how.

Q3. Define the basic elements of the metric structure: pulse, beat, tatum, measure, and tempo.

Q3. What defines and characterizes the metric structure of a musical piece? Can you briefly summarize how the modern western notation for metric structure works? Can you point out some of its limits?

Q3. Describe the fundamental components of a metric structure and their relevance in our perception of rhythm. Describe the rhythmic wheel as a tool for representing rhythms. How are the components of the metric structure represented with this tool? What kinds of rhythmic structures can be represented with this tool and what kinds cannot be represented with it?

Q3. What is a polymeter? Describe Brian Ferneyhough's notation for polymetric structures and specify the meaning of the numerators and denominators appearing in the fractions. When looking at this notation, how do we recognize that the corresponding polymetric structure is, in fact, a polyrhythm?

Q3. Describe in general terms what a polyrhythm is. In particular, describe the notion of systemic polyrhythm. Is the concept of polymetric structure more general than that of polyrhythmic structure? If so, how? Which of the two can be implemented using a standard drum machine (or represented using a "rhythmic wheel")? Please justify your answers.

Q5. Define and describe the Brian Ferneyhough's Notation (BFN) for polymeters. How would you express a 2-against-3 rhythm of a composite hemiola using BFN?

7 HARMONY

7.1 RELATION BETWEEN MELODY AND HARMONIC/RHYTHMIC STRUCTURES

A melody is a sequence of musical notes that are perceived as a coherent and organized musical idea. It is typically the most prominent and memorable aspect of a piece of music, serving as its main theme or tune. It often spans across multiple layers in music, including harmonic, rhythmic, and structural aspects:

7.1.1 Harmonic Interaction

Melodies interact with underlying harmonies or chords. The choice of notes in the melody can either align with or create tension against the harmonies, including both chord tones (notes that belong to the current harmony: consonance, harmonious) and non-chord tones (notes that create tension or momentarily deviate from the harmony: dissonance, tense). The balance between these elements affects the melody's character.

When the underlying chords or harmonies change, the melody can adapt to fit these changes, using different notes that align with the new harmony. Modulations (key changes) can introduce entirely new harmonic environments, and melodies often adapt to these shifts by changing their pitch content.

Note: *melody and contour are not the same thing:*

- *melody is a sequence of musical notes played or sung in a specific order, creating a coherent and organized musical idea;*
- *contour refers to the shape or trajectory of a melody's pitch as it rises and falls over time. It can be used to analyze and describe the general shape or pattern of a melody, such as whether it ascends, descends, or follows a more complex pattern.*

7.1.2 Rhythmic Interaction

Melodies can adapt to rhythmic changes by altering the timing and duration of their notes. Syncopation, for example, involves placing melodic accents on offbeats or unexpected moments, creating a rhythmic tension. Changes in tempo (speed) can also affect how a melody is perceived, as the same sequence of notes can feel different at various tempos.

7.1.3 Structural Role

Melodies play a significant structural role in music. They can mark the beginning and end of sections, serve as a recurring theme, or undergo variations and developments throughout a composition.

In classical music, for example, the exposition of a sonata often presents a primary melody, which undergoes transformation and development in the subsequent sections.

7.2 FOUR-PART HARMONY

7.2.1 Four-Part-Harmony Overview

Four-part harmony is a traditional system for organizing chords for four vocal ranges (the ranges values are taken from this [Wikipedia page](#)):

- **Soprano:** from C4 to C6. Typically, it is the range of female singers with the highest pitch range, such as a soprano in a choir or opera. Instruments that can play in the soprano range include the flute, the violin, and the oboe.
- **Alto:** from around A3 to A5. Usually, female singers with a lower range than sopranos, known as altos or contraltos. Mezzo-sopranos also often sing in this range. The alto flute, the viola, and the French horn can represent the alto voice.
- **Tenor:** generally, from roughly C3 to C5. Typically, male singers with a higher pitch range. In opera, a tenor would often be the male lead. The clarinet, the trumpet, and the cello have ranges that can accommodate tenor parts.
- **Bass:** often from about E2 to E4. Male singers with the lowest vocal range, including basses and bass-baritones. The bassoon, the trombone, and the double bass can cover the bass range.

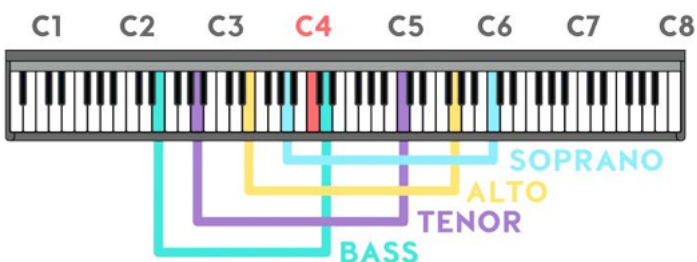


Figure 45

It's important to note that these ranges can vary slightly based on the specific choir or musical context (*figure 45* shows a slightly different range subdivision).

Historically, each vocal range in choral music was assigned a specific clef that indicated the pitch range of that voice part (*figure 46*). The soprano, alto, and tenor parts were notated using different C clefs, which are movable and can be placed on any line of the staff to indicate the position of middle C (C4). The soprano clef would be placed on the first line, the mezzo-soprano on the second line, the alto on the third line, the tenor on the fourth line, and the baritone on the fifth line. The bass part was written in the F clef, which places F below middle C on the fourth line of the staff.

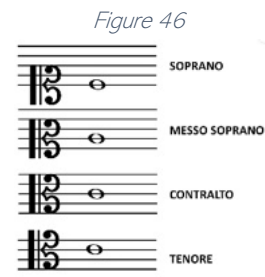


Figure 46

In contemporary music notation (*figure 47*), to simplify reading, the soprano and alto parts are typically written in the treble clef, which is also sometimes referred to as the "G clef" because it circles around the G above middle C. Tenor parts can be written in either the treble or bass clef, but when written in the treble clef, they are usually sung an octave lower than written. This practice is known as "octave transposition." The bass part continues to be notated in the bass clef. This modern system is easier for musicians to read and is widely adopted in most music written today.

Requiem
Dies Irae
Giuseppe Verdi

Presto

Soprano

Alto

Tenor

Bass

Piano

ff

Di - - es

Di es

Figure 47

7.2.2 Rules and Techniques of Four-Part Harmony

In four-part harmony, the four voices move in relation to each other in one of four ways:

- **Static:** Where one or more voices remain the same while others change.
- **Oblique:** Where one voice stays on a single pitch while the others move around it.
- **Contrary:** Where some voices move in opposite directions.
- **Similar:** Where all voices move in the same direction, but not necessarily by the same intervals.

The basic traditional rules for composing a correct four-part harmony are:

- **Spacing:** avoid placing two consecutive voices more than one octave apart, except for tenor and bass, where more leeway (*margine di manovra, libertà*) is given.
- **Crossing:** maintain the traditional order of the voices from highest to lowest without overlap: soprano, alto, tenor, bass.
- **Doubling:** often, one of the voices is doubled, usually the most stable note in the chord. Avoid doubling the leading tone in dominant-function chords (*In music theory, the leading tone is the seventh note of a diatonic scale, one semitone below the tonic. In a dominant-function chord, which typically leads to the tonic chord in a resolution, the leading tone has a strong tendency to move up to the tonic note to complete this resolution*).
- **Parallelism:** Prevent two voices from moving in parallel perfect fifths or octaves to avoid weakening the harmonic structure. Use contrary motion, especially when the bass moves by step.
- **Voicing for Triads:** In four-part writing, consider the overtone series and the natural expansion of intervals. Keep upper voices closer together and allow greater distance between the upper voices and the bass as it descends.
- **Baroque Practices:** In baroque music, special attention is given to the handling of the root, perfect cadences, texture, and voicing. Ensure that the soprano and alto are not more than an octave apart, and the tenor not more than a seventh below the alto.

These principles form the backbone of harmony not only in the vertical sense (through chord voicings) but also horizontally (through voice leading).

7.3 BASIC THEORY

7.3.1 The Diatonic Major Scale and Its Modes

The major scale, a diatonic scale, consists of seven notes with a specific pattern of whole and half steps: five whole steps and two half steps, with the half steps separated by either two or three whole steps. This pattern not only defines the major scale but also its modes. (*as we previously saw, when we begin the scale on different notes of the major scale, we generate the seven diatonic modes: Ionian, Dorian, Phrygian, Lydian, Mixolydian, Aeolian, and Locrian*).

7.4 TONAL PATTERNS

A tonal pattern in music is a sequence of notes that creates a particular melodic or harmonic formula, which can be recognized and repeated in different musical contexts. This pattern is often defined by its interval structure rather than by specific pitches, which allows it to be transposed to various tonal centres within a tonal framework. When adapting a tonal pattern to different scales, it's important to adjust the pattern to fit within the scale's structure. For instance, in a major scale, the pattern must use the scale's intervals (whole and half steps) to maintain the tonality of the scale. In a minor scale, the pattern may need to account for the different forms of the minor scale (natural, harmonic, and melodic) and adjust the intervals accordingly. In practice, musicians often use tonal patterns as building blocks for creating melodies or developing harmonic progressions. They can be a tool for improvisation, composition, and analysis. Some examples of tonal patterns are the II-V-I and the I-VI-II-V progressions, discussed in the next paragraph.

7.4.1 Tetrads and Common Chords Progressions

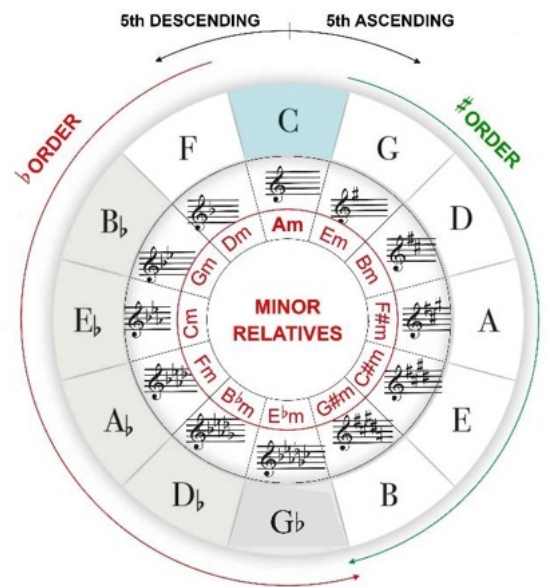
From these modes, we can derive tetrads (seventh chords) by stacking thirds, which is selecting every other note in the scale to form chords. For example, the Ionian mode creates a major 7th chord, the Dorian mode gives us a minor 7th chord, and so forth.

Consider now also the circle of fifth: it can illustrate the concept of harmonic proximity, where chord progressions that sound "better" or more cohesive tend to form more compact shapes within the circle.

Several common chord progressions derive from the circle of fifths:

- The [II-V-I progression](#), (I prefer [this version](#) but the pattern is not as obvious) used in jazz and other music genres, is a progression in which we utilize the second (II, minor 7th), fifth (V, dominant 7th), and first (I, major 7th) chords of the major scale. In this chord progression, the leading tones, like the third and seventh, tend to move forming continuous sub-melodies (the 3rd becomes the 7th in the following chord by staying in the same position, while the 7th becomes the 3rd by going half a step down).
- the [I-VI-II-V progression](#) is a classic chord sequence used in a multitude of songs across various genres, especially in jazz and pop. The chords in this progression are built on the first (I), sixth (VI), second (II), and fifth (V) degrees of the major scale. The VI and II chords are typically minor, while the V chord is typically dominant.

Figure 49



7.4.2 Chord Scale Theory

Understanding chord scale theory often begins by comparing the memorization of scales with the memorization of chords. The construction of chords can be seen as a selection process from the major scale, where certain notes are chosen to create harmony. Typically, we might consider a single octave and select the root note (which serves as our tonal center), the third (which determines the chord's major or minor quality), the fifth (which reinforces the tonal center), and the seventh (which adds color). However, we can also extend this process across two octaves, allowing us to incorporate extended tones like the ninth, eleventh, and thirteenth. These extended tones often introduce tension into a chord, adding richness and complexity to the harmony. When these extended tones are brought down an octave, they are no longer "extended" but become "added" tones that change the color or character of the chord without the tension usually associated with them being played in a higher octave. We can consider the role of each note with the tree graph in figure 50.

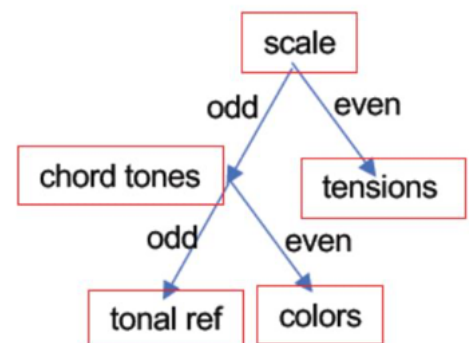


Figure 50

7.4.3 Major Modes Chords

Chord structures in music theory often utilize a concept known as the "avoid note." This term, not a strict rule, refers to a particular note within a mode that can introduce dissonance and is typically used more as a passing tone. For instance, within the Ionian mod, the fourth scale degree is commonly treated as an avoid note due to its tendency to clash with the harmonious nature of the scale.

Interestingly, each mode within the major scale framework generates a distinct type of chord:

- **Ionian Mode:** By sampling every other note (using a 2:1 subsampling method), we get the **C major 7th chord**, characterized by the inclusion of both a major third and a major seventh.
- **Dorian Mode:** This mode produces a **minor 7th chord**, due to the presence of a minor third and a minor seventh.
- **Phrygian Mode:** Typically, this mode forms a **minor 7th chord**. However, it is distinguished by a flat 9th scale degree. Often, the flat 6th is considered an avoid note within this mode.
- **Lydian Mode:** The chord corresponding to this mode is a **major 7th with an augmented 4th**, a note that was once the subject of much debate.
- **Mixolydian Mode:** This mode leads to the creation of the **dominant 7th chord**, with both a major third and a minor seventh. The 11th scale degree acts as the avoid note. Additionally, substituting the 4th scale degree for the major third results in a suspended chord, which, unlike the dominant chord, does not carry a strong resolution impulse.
- **Aeolian Mode:** Also known as the natural minor scale, the Aeolian mode is **typically played over a VI chord** in a major key context.
- **Locrian Mode:** This mode is associated with the **minor 7th flat 5 chord**, also known as the half-diminished chord. In Locrian, the flat 9th is an avoid note.

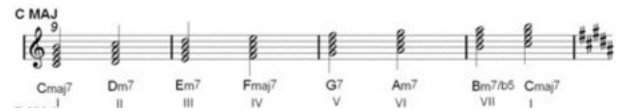


Figure 51

It is worth mentioning that a variant of the Lydian mode exists where the 7th scale degree is flattened, creating a scale that doesn't naturally derive from the modes of the major scale.

7.4.4 Minor Modes Chords

Transitioning to minor modes, the melodic minor scale presents a unique case. It resembles the major scale but incorporates a minor 3rd instead. This scale shouldn't be confused with the Aeolian mode, which is synonymous with the natural minor scale. Notably, the melodic minor scale isn't considered diatonic, as it doesn't maintain the same pattern of whole and half steps found in the major scale.

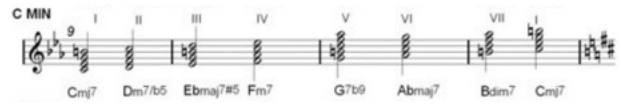


Figure 52

In the major scale, there are two types of diatonic fourths: the tritone and the perfect 4th. Contrastingly, the melodic minor introduces a third variety, the diminished 4th (also known as a major 3rd).

Minor scale modes enjoy a notable absence of avoid notes, indicating a higher degree of freedom. Almost any melodic phrase in one mode can be applicable in another. The modes of the melodic minor are versatile. For example, a II-V progression in melodic minor doesn't necessarily have to resolve to a minor chord; it can seamlessly lead into a major chord as well. The II-V-I progression exists in the minor scale as well, but unlike the major progression where all chords derive from a single scale, the minor progression's chords come from three different melodic minor scales.

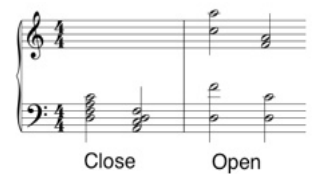
Let's analyze the minor modes chords:

- **Mode 1:** The first mode of the melodic minor scale forms the minor-major 7th chord. This chord is a substitute for minor 7th chords and is used as the tonic in the minor scale, whereas the minor 7th is typically the ii chord in the major scale.
- **Other modes:** the other modes derived from the melodic minor scale lack specific names. Musicians often resort to these chords using "slash chords". These are notated as a triad over a different bass note. For example, the chord symbol "C/D" denotes a C major triad (C, E, G) played over a D bass note.

7.5 VOICINGS

Voicing refers to the vertical arrangement of notes within a chord. It involves altering the order of notes, often described as "dropping" notes to different octaves, to achieve a desired harmonic effect. The way a chord is voiced can significantly influence its impact on the listener, as the distribution of notes across the sonic spectrum changes the chord's texture and color. Different voicings can result in the same chord having a varied presence in a musical context. For instance, a chord voiced with tight, close spacing may produce a sense of tension and compactness, whereas the same chord voiced with wider spacing can offer a sense of openness and airiness.

Figure 53



Another technique in voicing is the omission of certain notes, which is not merely a matter of convenience for the player but a deliberate choice to affect the chord's character. By selectively leaving out notes, a musician can "clean up" the sound of a chord, reducing potential dissonance or creating space for other musical elements. Commonly omitted notes include the perfect fifth, which can be excluded without significantly altering the chord's perceived quality, and sometimes even the root note, particularly when another instrument in the ensemble is providing the harmonic foundation. The omission of notes can be especially useful in complex chords with many extensions, such as 9ths, 11ths, and 13ths, where avoiding certain pitches can prevent muddiness and maintain clarity within the harmony.

The concept of voicing is influenced by the overtone series, which is a natural occurrence in the physics of sound where a fundamental pitch produces a series of harmonics at fixed intervals above it. The first harmonics (partials) are spaced further apart, and as you move to higher harmonics, they become closer together. This natural phenomenon influences how the bass voice in harmony is often written with wider intervals between its notes compared to the upper voices, whose intervals tend to be closer together. This creates a sonorous and well-balanced chord structure that is pleasing to the ear, following the principles observed in the overtone series (*as we have previously seen, in music the overtone/harmonics series is a natural phenomenon where higher frequency notes (harmonics) are produced alongside the fundamental note. These overtones follow a specific pattern, with lower overtones spaced wider apart and higher overtones closer together. In four-part harmony, this concept is applied to chord construction. The bass voice, mirroring lower overtones, is spaced with wider intervals. In contrast, the upper voices use closer intervals, akin to higher overtones*).

7.6 VOICE LEADING

While voicing refers to the vertical arrangement of notes, Voice Leading refers to the horizontal arrangement of notes. Voice leading is the way in which individual melodic lines (or 'voices') move from one note to another in a piece of music. Good voice leading typically involves smooth, logical progressions of notes, minimizing awkward jumps or leaps and maintaining the individuality of each voice. It's crucial in genres like [classical counterpoint](#) but is also a fundamental concept in other forms of music. It focuses on how the melody, harmony, and bass lines interact with each other to create a cohesive musical narrative (analysing them as horizontal meaningful musical phrases, not as vertical chord structures).

7.6.1 Cadence

A [cadence](#) is a sequence of chords that brings a section of music to a sense of resolution or pause. It's akin to a punctuation mark in music. Cadences can be strong (like a full stop), giving a sense of finality (perfect cadence), or weaker (like a comma), suggesting continuation or pause (imperfect cadence). The sense of resolution is given by creating a moment of high tension (dissonance) that resolves in a moment of rest (consonance). This process is showed in the *figure 54*, in which a series of chord build an increasing tension atmosphere that resolves with the last chord (you can listen to it [here](#)).



Figure 54

7.6.2 Clausula Vera (*hint*)

The [Clausula Vera](#), meaning "true close" in Latin, is a specific type of cadential formula used in the polyphonic music of the Renaissance era. It is related to counterpoint and so it is detailed in a [next paragraph](#).

7.7 HARMONIC (VOICING) AND MELODIC (VOICE LEADING) DISSONANCES

Harmonic dissonance and melodic dissonance are two different concepts in music theory that involve dissonance but occur in different musical contexts and have distinct characteristics:

- **Harmonic Dissonance:** Harmonic dissonance refers to dissonance that occurs when two or more notes or chords are played simultaneously in a harmony, such as in a chord progression or when multiple instruments are playing together. It is related to the tension created by the combination of certain intervals or chords, such as augmented or diminished intervals, tritones, or dissonant chords like the dominant seventh chord. Harmonic dissonance typically results from the interaction of notes within a **vertical or simultaneous structure**, like a chord, and its resolution or resolution tension plays a significant role in traditional Western music theory.
- **Melodic Dissonance:** Melodic dissonance, on the other hand, pertains to dissonance within a single melodic line or a single voice within a composition. It occurs when a melodic line contains intervals that create a sense of tension or clash between the notes. These intervals may include augmented seconds or other dissonant intervals within the melody. Unlike harmonic dissonance, which involves the interaction of multiple notes or chords at the same time, melodic dissonance focuses on the **sequential movement of a single melody** and the dissonant intervals it contains. Melodic dissonance can be used to create drama, emotion, or contrast within a melody and is often resolved by moving to more consonant intervals in subsequent notes.

7.8 COUNTERPOINT

7.8.1 Basic information on Counterpoint

Counterpoint is a term that originates from the Latin "punctus contra punctum," which means "point against point." It is a compositional technique in Western music that has been used since the Middle Ages. Counterpoint involves the interweaving of two or more melodic lines into a harmonious musical tapestry. The lines are independent in rhythm and contour but are harmonically interdependent (Voicing, Voice Leading).

There are four types of motion between two voices in counterpoint:

- Contrary motion: The voices move in opposite directions.
- Oblique motion: One voice stays the same while the other moves.
- Parallel motion: Both voices move in the same direction and distance.
- Similar motion: Both voices move in the same direction but at different intervals.

7.8.2 Clausula Vera

As we have seen in the [Voice Leading Paragraph](#), two very important concept in the horizontal development of a melody are consonance and dissonance. This is true also in Counterpoint; Counterpoint relies on the tension and resolution between dissonance (notes that create tension) and consonance (notes that sound harmonious together). Dissonant intervals resolve to consonant intervals to create a sense of forward motion (cadence).

A specific type of cadence, very common in renaissance counterpoint, is the "[clausula vera](#)".

The term "clausula vera" is Latin for "true close" and is traditionally used to describe a cadential formula that brings a contrapuntal section to a conclusive close. The most common form of clausula vera involves a stepwise (*graduale*) resolution of the upper voices to the nearest chord tones of the final chord, often involving a leading tone resolving up

to the tonic and the supertonic resolving down to the tonic (es. If we are in C key: [Re→Do] && [Si → Do]). This movement creates a strong sense of finality.

7.8.3 Counterpoint species

There are several species or types of counterpoint that have been developed and studied over the centuries, each with its own set of rules and characteristics. The most well-known species of counterpoint are often associated with Johann Fux's treatise "Gradus ad Parnassum" and are categorized as follows:

- **First Species** (Note against Note or 1:1): In first species counterpoint, each note in the added voice (counterpoint) is harmonized with a single note in the given voice (cantus firmus). This is the simplest form of counterpoint and focuses on achieving consonance and avoiding dissonance.
- **Second Species** (Two against One or 2:1): In second species counterpoint, two notes in the added voice are harmonized with one note in the given voice. This introduces the concept of a syncopated rhythm where the added voice moves twice as fast as the given voice.
- **Third Species** (Four against One or 4:1): Third species counterpoint involves four notes in the added voice for every one note in the given voice. This leads to more rhythmic complexity and requires careful handling of dissonances.
- **Fourth Species** (Suspension or 1:1 with syncopation): Fourth species counterpoint combines first species counterpoint with syncopation and suspensions. It involves suspending dissonant intervals until they resolve to consonant ones.
- **Fifth Species**: Fifth species is the most complex. It incorporates elements of the previous species and encourages ornamental and melodic decoration.

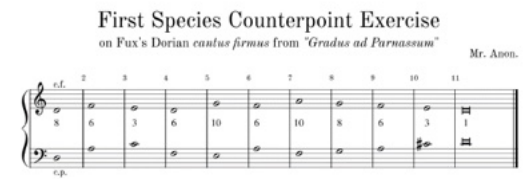


Figure 55



Figure 56

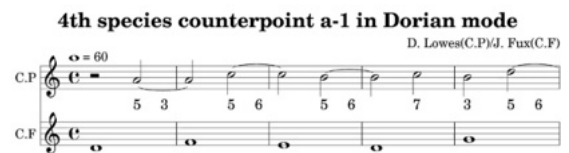


Figure 57

These species of counterpoint provide a structured framework for composers to create harmonically interesting and pleasing music while adhering to certain rules and principles. They were especially influential during the Renaissance and Baroque periods and continue to be studied by composers and musicians today.

7.8.4 Additional Counterpoint Transformation Techniques

In counterpoint, aside from the traditional species rules, composers can create multiple melodic lines using various transformation techniques:

- **Inversion**: This involves flipping the intervallic structure of a melody upside down. If the original melody moves up by a certain interval, the inverted melody will move down by the same interval, and vice versa (figure 58).
- **Retrograde**: A melodic line is played backwards. This means the melody's notes are performed in reverse order from the end to the beginning (figure 58).

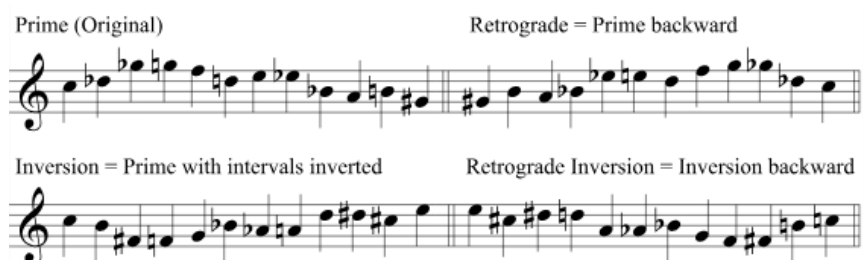


Figure 58

- **Augmentation:** The rhythmic values of the original melody are proportionally increased. For example, if the original melody has a series of quarter notes, in augmentation, these might become half notes or whole notes (*figure 59*).
- **Diminution:** The opposite of augmentation, diminution involves shortening the rhythmic values of the original melody. Quarter notes might be transformed into eighth notes, for instance.
- **Transposition:** This is a relatively straightforward technique where the entire melody is moved up or down in pitch by a consistent interval.
- **Canon:** One voice starts the melody, and another voice enters later with the same melody at a different pitch level or using one of the transformations like inversion or augmentation (*figure 60*).



Figure 59



Figure 60

7.9 QUESTIONS FROM PREVIOUS EXAMS

- Q5.** What is the difference between harmonic dissonance and melodic dissonance? When and how are such concepts relevant?
- Q5.** Define counterpoint and its “vertical” (harmonic) and “horizontal” (melodic) functions. Explain what this has to do with the concept of voice leading (i.e. resolution). Define and describe the notions of melodic and harmonic consonance and their role in counterpoint.
- Q4.** Define and describe what a II-V-I chord progression is. Explain what modes of the major scale are involved in this progression, and what the role of this turnaround is in terms of “voice leading”.
- Q5.** Consider the following sequence of chords/scales:
C ionian, F lydian, E mixolydian, A aeolian, D dorian, G mixolydian, C ionian
 Does this sequence contain any harmonic surprise? The answer will be no if all scales share the same notes. If so, please specify which scale/chord constitutes a surprise, and explain why that scale change was introduced.
- Q4.** What are the notes of a *G* dorian scale? (*hint*: use the image of the keyboard in the bottom, and explain the procedure).
- Q5.** What is a “clausula vera” in counterpoint? And how does it relate to resolutions in harmony?
- Q6.** How do consonance and dissonance change with respect to frequency? And how does this affect the construction of voicings?
- Q5.** Briefly describe in broad terms what counterpoint species are, and how complexity progresses as we go from first to fourth species.

- Q8. How would you characterize and describe a melody? Does it span across multiple layers (harmonic, rhythmic, structural, etc.)? If so, which ones? How would you explain how a melody can adapt to harmonic and/or rhythmic changes?
- Q6. Explain what a tonal pattern is and how you can identify it as it adapts to different reference tonal scales.
- Q4. How are the distances (in half-tones) between white keys distributed in a piano keyboard? Can you explain why such distances are not uniform?
- Q6. Can you briefly explain what counterpoint is? How does it use the notions of melodic and harmonic consonance? What role does "voice leading" play in it?
- Q4. A pattern of notes fully is described by a **periodic sequence of relative distances between consecutive notes, expressed in "steps of a scale"**. Assume that the scale is Ionian in C; and consider the period of relative distances $\{+2, -1, -2, -1\}$ (to be read as "up 2 steps, down 1, down 2, down 1"). Starting from the note "A", list the first 12 notes of the pattern. If your scale were a C Dorian, how would these 12 notes change? Please explain why. *(you can use the image of the keyboard below for reference)*
- Q5. Explain the relationship that exists between scales and chords in harmony. How do you derive a chord (quadriad) from a modal scale? How do you turn (in general terms) this relationship into a duality (between scales and chords)?
- Q6. Explain the difference between harmonic and melodic consonance. Can you offer specific examples of intervals and tonal progressions that are considered dissonant or consonant either way?
- Q6. What are modal scales and how are they named? *Note: if you can't remember them by name, just tag them with their degree.* Now order them from the brightest to the darkest, and explain the procedure that you followed to perform this sorting (do so, if possible, graphically. For your convenience you can refer to the picture of the keyboard below).
- Q6. In addition to the rules set forth for the various species of counterpoint, there are alternate ways to build multiple melodic lines, based on transformations. Please explain what they are and, in broad terms, how they work.
- Q4. How do you define a tonal pattern? How would you modify it in correspondence to a chord change, so that it will be perceived as tonally invariant? Please justify the answer.
- Q5. What are the notes of an aeolian scale in D? *(use the image of the keyboard in the bottom, and explain the procedure)*

8 MUSICAL STRUCTURE

8.1 MOST COMMON MUSICAL STRUCTURES

Musical structure is an integral part of composition, shaping how a piece of music unfolds and impacts the listener. Melodies and musical pieces are typically organized according to specific structures that balance repetition and contrast, creating a coherent and compelling musical narrative.

Let's now analyze the most used musical structures:

- **Binary Structure:** the simplest form of musical structure is the binary structure, which consists of two sections of roughly equal length, labeled A and B. The first section, A, sets up musical ideas or themes, and is usually complemented or answered by the second section, B. In binary form, there is often a key change within the first section that is continued into the second section, and sometimes a return to the original key by the end.
- **Ternary Structure:** a slightly more complex form is the ternary structure, built upon three sections labeled A, B, and A. This form introduces a contrast with the B section that is distinct from A. The final return to A can either replicate the initial section exactly or introduce slight variations to maintain interest and provide a sense of closure.
- **Arch-Shaped (ABA) Structure:** the arch-shaped structure takes the concept of symmetry in musical narrative, often following an A, B, C, B, A pattern. The middle section, C, acts as the apex of the arch, emphasizing the peak of musical tension or thematic development.
- **Rondo Form:** the rondo form is characterized by a principal theme (A) that alternates with one or more contrasting themes or episodes (B, C, D, etc.). These episodes provide contrast and variety while the recurring A section offers a familiar anchor to the listener.
- **Theme and Variations:** It is a musical form that presents a theme and then alters it in subsequent iterations through changes in harmony, melody, rhythm, timbre, or other musical elements. Each variation maintains the essence of the theme while bringing new elements to the piece.
- **Ground Bass and Pedal:** Ground bass, also known as basso ostinato, involves a short, recurring melodic pattern in the bass part, over which the upper parts proceed and vary. The 'pedal' is a sustained note over which harmonies change, often used to create tension and release. These techniques were prevalent during the Baroque period, exemplified by Pachelbel's Canon in D.
- **Call and Response:** A dialogue within music, where a phrase (the 'call') is answered by a subsequent phrase (the 'response'). This form is foundational in African, African-American (like Gospel), and many forms of American music, including blues and jazz.
- **Minuet and Trio Form:** A dance form from the Baroque and Classical periods, often found in symphonies and chamber music, comprises two sections: the minuet (A and B) and the trio (C and D), with the minuet often repeated afterward (A and B). The trio typically features lighter texture and contrasts with the minuet.
- **Strophic, Through-Composed, and Da Capo Aria:** Strophic form repeats the same musical structure with each stanza of the lyrics, commonly found in hymns and folk songs. Through-composed form means new music is written for every stanza, avoiding repetition and reflecting the song's evolving narrative. Da Capo Aria, frequently used in operas, features an ABA structure where the initial section returns after a contrasting middle section.
- **Cyclic Form:** Here, a sequence of chords or melodic pattern is repeated continuously, allowing for variations and developments such as improvisation or changes in dynamics. This form is seen in various types of world music, including African drumming and Indian classical music.
- **Popular Song Forms:** Modern popular music often uses a verse-chorus structure, sometimes with a bridge. Verses maintain the narrative, choruses provide the emotional climax, and bridges offer a contrasting section. Other elements like intros, outros, and instrumental fills may be included.

- **12-bar Blues:** A foundational structure in blues music is based on a specific chord progression over twelve bars, which is repeated throughout the piece. It can be varied with seventh chords, turnarounds, or bebop changes to increase complexity.
- **Sonata Form:** Predominantly used in classical music, the sonata form unfolds in three main parts: exposition (introducing themes), development (exploring and varying these themes through different keys and figures), and recapitulation (returning to the original themes), often concluding with a coda.

In constructing these structures, composers often employ techniques such as the "1-2-3" method for thematic development, counterpoint to enrich the texture, and strategic repetition to reinforce musical ideas.

8.2 STRUCTURAL ANALYSIS

In the previous paragraphs we explained the components of the music, putting them together so as to build the complete songs. Now we want to carry out the reverse process: we want to extrapolate the song's elementary components from so that we can analyze them. To do this, the following concepts are necessary.

8.2.1 Feature Vectors in Music Analysis

In the context of music analysis, a feature vector is a multi-dimensional representation of certain characteristics extracted from a segment of audio. These characteristics, or features, are numerical values that provide information about various aspects of the sound, such as pitch, timbre, rhythm, and dynamics. When we analyze an audio excerpt, we typically break down the audio into short segments known as frames. Each frame might be just a few milliseconds long. The reason for this frame-by-frame analysis is to capture the temporal evolution of the sound, as music is inherently a time-varying phenomenon. Once we have the frames, we extract features from each one. These features might include measurements like spectral content, energy, zero-crossing rate, etc. Every feature extracted from a frame represents one dimension in our feature space. So, if we extract 12 chroma features (which correspond to the 12 different notes in the Western musical scale) and 15 MFCCs (which represent the spectral envelope of a sound), we would have a 27-dimensional feature space (12 chromas + 15 MFCCs).

8.2.2 Chromagram

A chromagram is a tool used in music signal processing that reflects the intensity of the various pitches, or pitch classes, present in a piece of music. In Western music, there are twelve unique pitches, each corresponding to a semitone within an octave. These are the pitch classes that a chromagram maps. The term "chroma" is derived from "chromatic scale," which is a musical scale with twelve pitches, each a semitone apart.

When we analyze an audio signal with a chromagram, we are essentially converting the audio into a simplified form where only the pitch class information is retained while the octave information is discarded. This is because a chromagram is concerned only with the type of notes being played, not the specific octaves they are in. For instance, whether a C note is played in a high octave or a low one, it will be represented in the same position on the chromagram.

The main characteristics of a chromagram are:

- **Harmonic characteristics:** A chromagram can detect chords and harmonic progressions because chords are made up of specific combinations of pitch classes.
- **Melodic characteristics:** Since melodies are sequences of pitches, a chromagram can track the changes in pitch classes over time, thus capturing the melodic contour.
- **Robustness to Timbre and Instrumentation:** The chromagram's focus on pitch class makes it less sensitive to the timbre or color of the sound, which means it can effectively represent the same melody or chord progression regardless of the instrument used to play it. This robustness makes chromagrams particularly useful in applications like music recognition, where the melody or harmony is more important than the specific sound characteristics of the performance.

8.2.3 Mel-Frequency Cepstral Coefficients (MFCCs)

Mel-Frequency Cepstral Coefficients (MFCCs) represent a different aspect of sound. Mel-Frequency Cepstral Coefficients are derived from the spectrum of a musical signal, which is a representation of the signal in the frequency domain.

The process to do to obtain the MFCCs is the following:

- **Log Power Spectrum on a Mel Scale:** The process of computing MFCCs begins by taking the Fourier transform of a signal to get its spectrum. Then, the powers of the spectral components are mapped onto the Mel scale, which is a perceptual scale that more closely mimics the human ear's response to different pitches than the linear frequency scale (see *Pitch and Intensity* section).
- **Cepstral Representation for Timbral Characteristics:** After mapping to the Mel scale, the logarithm of these powers is taken to capture the dynamic range of human hearing. Then, a type of Fourier transform known as the Cosine Transform is applied, resulting in a set of coefficients – the MFCCs. These coefficients provide a compact representation of the sound's timbre, which includes the character or quality of the musical sound that distinguishes different types of voices and instruments.
- **Insight into Timbral Characteristics:** Because they are based on the Mel scale, MFCCs are particularly good at representing the timbre of the audio, capturing nuances that make different instruments sound unique even if they play the same note.

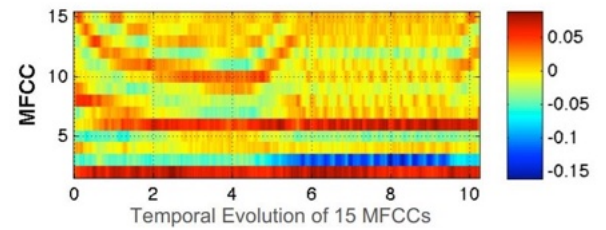


Figure 61

8.2.4 Self-Similarity Matrix (SSM)

The Self-Similarity Matrix (SSM) is a sophisticated tool used to analyze the structure of a musical piece by examining the relationships between different parts of the music. It serves as a map of a musical piece's internal relationships, providing insights into its composition and structure by revealing the patterns of repetition and change over time. Here is an expanded explanation of the concept:

- **Matrix Dimensions:** An SSM is essentially a square matrix, with the size of the matrix being determined by the number of feature vectors that represent the musical piece. If the piece is divided into 'N' segments, then the matrix will be 'N x N', with each row and column corresponding to one of the segments.
- **Pairwise Similarities:** The matrix measures the similarity between every pair of segments in the music. Each element in the matrix represents the degree of similarity between the segments corresponding to that row and column. For instance, the element at the intersection of row i and column j indicates how similar segment i is to segment j.
- **Visualization of Music Structure:** The SSM is a powerful visualization tool. When plotted as an image, the degree of similarity is often represented by colour intensity or shading, with brighter or more intense colours indicating higher degrees of similarity.
- **Highlighting Repetitions and Variations:** Because musical pieces often have repeating themes or motifs, these repetitions show up in the SSM as diagonal stripes. The main diagonal, which compares each segment with itself, will always show maximum similarity. Stripes parallel to the main diagonal indicate repetitions of themes or sections. Variations, which are segments similar but not identical, also show as stripes, but with less intensity than the repetitions.
- **Blocks in the SSM:** Besides diagonal stripes, blocks or clusters within the SSM can represent parts of the music that share similarities with each other, such as a chorus that is repeated with slight variations throughout a song.

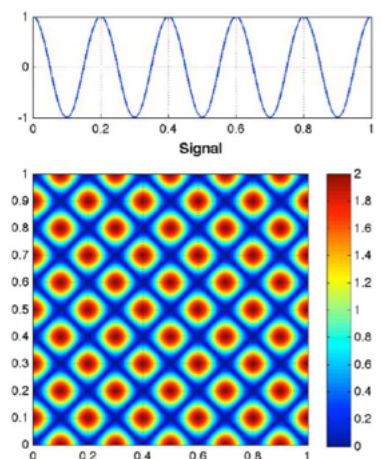


Figure 62

- **Subsequence Identification:** By looking at the patterns of stripes and blocks in the SSM, researchers and musicians can identify which parts of the music have similar content. This is particularly useful for understanding the structure of the piece, such as identifying the verse-chorus-verse pattern in a pop song or the exposition-development-recapitulation structure in a classical sonata.

8.2.4.1 SSM: Examples from previous exams

- **Minuet and trio SSM:** A strong diagonal line from the bottom-left to the top-right, representing the similarity of the piece to itself over time. Two squares of high similarity along this diagonal line separated by a rectangle of lower similarity. These squares represent the A sections (the minuet) which are similar to themselves, while the rectangle represents the B section (the trio), which is different from the A sections. The first and last squares (the A sections) would be more similar to each other than to the central rectangle (the B section), reflecting the return to the minuet theme after the trio.
- **Sonata SSM:** Two large blocks of high similarity at the beginning and end of the matrix (yellow blocks in figure 64), respectively, since they contain the same themes although in different keys. These blocks represent the exposition and recapitulation. A more complex, less uniform area between the exposition and recapitulation blocks, reflecting the thematic and harmonic explorations and transformations taking place. This section would show varying degrees of similarity as themes are manipulated. Areas of lower similarity between the exposition and development, and between the development and recapitulation, corresponding to the transitions between these sections.

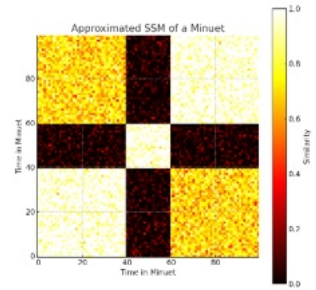


Figure 63

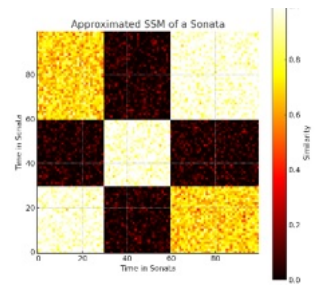


Figure 64

8.2.5 Analytical Methods

Another method to check the similarity of two parts of an audio sequence, is by using analytical methods.

- **Euclidean distance:** The Euclidean distance is the most straightforward way to measure the distance between two points in space. When we talk about feature vectors, each vector is a point in a multi-dimensional space, where each dimension corresponds to a feature of the music. Mathematically, it is expressed as:

$$distance = \sqrt{\sum_{i=1}^N (a_i - b_i)^2}$$

A smaller Euclidean distance between two vectors means they are more similar. In music analysis, this could indicate that two segments of music are similar in terms of their pitch, rhythm, timbre, etc.

- **Dot Product and Cosine Similarity:** The dot product is another measure of similarity between two vectors. It multiplies corresponding elements of the two vectors together and sums up the results. This measure reflects not just the magnitude of the vectors but also their direction relative to each other. When the dot product is normalized by the magnitudes (lengths) of both vectors, the result is the cosine of the angle between them. This normalized dot product is known as cosine similarity. A cosine value of 1 means the vectors are identical, 0 means they are orthogonal (independent), and -1 means they are diametrically opposed.

8.2.6 Phase correlation

(not present in the slides but I found it in a past exam)

Phase correlation is a technique often used in signal processing that involves the frequency domain to identify the relative shift between two similar signals. This method is based on the principle that the phase component of a signal's Fourier transform can provide information about the time alignment of the signal. In the context of analyzing rhythmic

structures, phase correlation can be used to detect and visualize the presence of multiple rhythmic components. Here's how:

- **Fourier Transform:** First, you would take the Fourier transform of the time-domain rhythmic signals to move them into the frequency domain. This transform separates the signals into their constituent frequencies and phases.
- **Phase Difference:** By comparing the phases of these two transformed signals, you can calculate the phase difference at each frequency.
- **Inverse Fourier Transform:** By taking the inverse Fourier transform of the phase difference, you obtain a correlation signal that indicates the time offset between the two rhythmic patterns.

To detect multiple rhythmic components, you would look for peaks in the phase correlation output. Peaks indicate points of high correlation, suggesting that the two rhythmic components have a consistent phase relationship, which can be interpreted as them being in sync or having a stable offset.

If the goal is to visualize two rhythmic components of a polyrhythm, **Analysis Window/Period** should be chosen carefully to encompass at least one full cycle of the most extended rhythmic component. This ensures that the phase correlation captures the relationship between the two rhythms over their entire cycle.

(Note: the Analysis Window is a temporal subset of the signal under investigation. For rhythmic analysis, it would be a specific duration of the musical piece that is being considered).

8.2.6.1 Examples from previous exams

Imagine a polyrhythm consisting of two distinct rhythmic components, where one rhythm has a cycle that completes in a shorter period than the other. If you choose a period for the phase correlation analysis that aligns with the shorter cycle, you're effectively only analyzing the time frame where the shorter rhythm completes its cycle. This period might not encompass a complete cycle of the longer rhythm. The effects of a Skewed Analysis Window are:

- **Incomplete Analysis:** Since the chosen period doesn't cover the full extent of the longer rhythm, the phase correlation would not capture the entire relationship between the two rhythms. You'd only see how the short rhythm behaves within its own cycle without understanding its interaction with the longer rhythm over a more extended period.
- **Misleading Results:** The phase correlation might indicate points of alignment or misalignment, but these would be restricted to the short rhythm's cycle. As a result, you might incorrectly conclude that the rhythms are in phase or out of phase based on incomplete data.
- **Overlooking Polyrhythmic Nature:** The true nature of a polyrhythm lies in how the different rhythmic cycles interact over time. By not considering the full cycle of the longer rhythm, you miss out on observing the composite rhythm's overall pattern, which emerges from the interplay between the short and long cycles.

8.3 QUESTIONS FROM PREVIOUS EXAMS

Q6. Briefly describe the typical structure of a song (sections and their roles). If you were to analyze a song using a Self-Similarity-Matrix (SSM), what would you expect the SSM to look like (in broad terms)? Please sketch a diagram to explain.

Q6. Describe the structure of "minuet and trio". If you were to analyze a song using a Self-Similarity-Matrix (SSM), what would you expect the SSM to look like in broad terms?

- Q3.** Define and describe the phase correlation, and its use as a tool for analyzing rhythmic structures. How do you use it to detect the presence of multiple rhythmic components? If your goal is to visualize two rhythmic components of a polyrhythm, how would you select the analysis window/period of the phase correlation? What happens to the phase correlation if you selected the period in relation to only one of the two components of the rhythmic structure?
- Q6.** Schematically describe the structure of “minuet and trio” while emphasizing what elements are similar to each other. Starting from this, describe what a corresponding Self-Similarity Matrix would look like, with the help of a
- Q3.** Define and describe the phase correlation as a tool for analyzing rhythmic structures. How can it be used to highlight the presence of rhythmic components? If you have multiple rhythmic components, how do you select the analysis period of the phase correlation in order to visualize both simultaneously? If the period is chosen in relation to only one of the two components of the rhythmic structure, what happens to the other component?
- Q8.** A melody can undergo a number of transformations and still be recognizable. Can you describe some? What are the salient features of a melody that remain invariant? and how would you capture them into a model of a melody?
- Q7.** Describe the structure of a sonata form and compare it with the structure of a typical song. Briefly describe what a Self-Similarity Matrix is, and what it would look like (in broad terms) if you used it to analyze the structure of a song (please draw an approximate picture).
- Q7.** What is a Self-Similarity Matrix (SSM)? Please explain what type of variables it uses and how similarity is computed. Explain in broad terms how you would organize a segmentation algorithm based on SSM. Assuming that the different sections of a musical piece tend to differ mostly in their harmonic content and not much in their rhythmic structure, what kind of features you would you tend to use in an SSM-based segmentation algorithm?