The Metaphysics of Tonal Harmony

A thesis submitted to The University of Manchester for the degree of
Doctor of Philosophy
in the Faculty of Humanities

2022

Luiz Leal

School of Social Sciences
Department of Philosophy

Table of Contents

Word count: 90,000 words

Abstract

The present study is an enquiry into the metaphysics of tonal harmony. Specifically, it is aimed at accounting for the ontology and the aesthetics of chords. Part I concerns the ontology of chords. Since chords are sounds, I begin by considering the nature of sound. In Chapter One, I characterise the 'problem of sound' in terms of the current philosophical debate, which is marked by two main dichotomies: on the one hand, the Wave and the Event View are taken to be irreconcilable; on the other hand, the Event and the Object View are also seen as incompatible. Conversely, in Chapter Two I propose a thesis of 'sonic reconciliation', which is aimed at integrating the fundamentals of those views. In light of my conciliatory approach, I proceed to examine the nature of chords *qua* sound events in Chapter Three, where I establish the necessary conditions for chord events to exist. This is followed by an analysis of the psychoacoustics and phenomenology of chord perception in Chapter Four, under which I give an account of the nature of chords *qua* Gestalt-based auditory objects. Yet, given that chords are distinctive musical entities within the domain of tonal harmony, in Chapter Five I provide an account of their nature *qua* musical individuals, which further substantiates my treatment of the ontology of chords.

In turn, Part II is focused on the aesthetics of chords. In Chapter Six, I introduce the 'problem of consonance' as concerning what kind of property consonance is: on the one hand, consonance is treated primarily as a psychoacoustic property; on the other hand, it is taken to be a musical property. The problem thus consists in the conflict between psychoacoustic and musical views on consonance (and dissonance). In response, I propose the thesis of the dual phenomenology of consonance/dissonance as pertaining to two different components of 'musical consonance': the sensory component and 'harmony' (i.e. the distinctively musical experience of chords). After examining the psychoacoustic basis of consonance/dissonance perception in Chapter Seven, in Chapter Eight I account for the aesthetics of chords *qua* isolated auditory objects and *qua* musical individuals within a distinctively musical context, whilst also indicating how the psychoacoustics of harmony and the musical experience of it may be reconciled. In Chapter Nine, I substantiate the thesis of the dual phenomenology of consonance/dissonance further by arguing that the sensory component is best accounted for under the notion of response-dependence, whereas 'harmony' is best understood under aesthetic supervenience. In Chapter Ten, I examine the instability of chord tokens *vis-à-vis* tuning and temperament differences, which have historically proven to be more intractable than the problem of consonance itself.

Declaration

No portion of this work has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Acknowledgements

This work comes with a deep sense of gratitude. I would like to thank Dr David Liggins, Prof James Garratt and Dr Emily Caddick Bourne for their support and advice in writing this thesis. Their time and patience have been much appreciated. I would also like to acknowledge Prof Julian Dodd's contribution during the earlier part of the course.

This research project has been generously funded by the Arts and Humanities Research Council (AHRC) via the North West Consortium Doctoral Training Partnership (NWCDTP), for which I am most grateful.

I would also like to thank my parents, Armando and Maria de Lurdes, for their unwavering encouragement and support; and my friend Dominic Berry – for the good companions that philosophy, music *and* poetry most certainly make.

Dedication

To the fundamental One that is never missing.

About the author

Luiz Leal is a PhD candidate in the Department of Philosophy at the University of Manchester. Prior to embarking on the PhD, he was awarded a Master's Degree in Philosophy from Heythrop College, University of London. Luiz has worked in education for over a decade. He is an amateur pianist, having attended music school in his youth.

Overture

The present study is an enquiry into the metaphysics of tonal harmony. Specifically, it is aimed at accounting for the ontology and the aesthetics of chords. As such, this thesis lies at the intersection between philosophy and music theory, and is thus fundamentally the product of an interdisciplinary effort. The approach that I have sought to convey in the forthcoming chapters draws from those two disciplines as well as other subject areas such as psychoacoustics and music psychology, with a view to incorporating not only the most relevant literature but also the latest research on the perception of tonal harmony. Hence, while the presentation of the material can be said to bear a distinctly philosophical tone, it is hoped that the questions raised here and the arguments developed in response may help bring an understanding of the metaphysics of tonal harmony to a broader audience. Still, I take the present work to be primarily intended as an original contribution to the philosophy of music.

Although the philosophy of music has recently benefited from a surge in publications, much of what has been written on musical ontology and aesthetics tends to be focused on the nature of works of music as well as performances thereof. Conversely, little has been said of the internal elements of those works, such as harmony – i.e. the 'vertical aspect of music', as it is commonly described. My ambition in this thesis is thus to make some headway in addressing that imbalance. It is in that spirit that I aim to provide a substantial philosophical treatment of the ontology and the aesthetics of tonal harmony. Correspondingly, this work has been structured into two parts.

As indicated by its title, this is a work in metaphysics, and its object is the nature of chords – which are taken to be the 'building blocks' of tonal harmony. While Part I concerns the ontology of chords, in Part II I offer an account of the aesthetics of chords with a narrow focus on the properties of consonance and dissonance, in particular. Although I seek to address some of the most typical metaphysical questions involving ontology, aesthetics and the nature of perception, it is not my intention to enter into a debate around the merits of metaphysics, or of what purposes a metaphysical account should serve. Rather, I take my proposal to be workable whatever the *raison-d'être* of metaphysics turns out to be. Furthermore, as an enquiry into the nature of *tonal* harmony, this study will be strictly confined to an understanding of it within Western music, with particular focus on its most traditional iteration. Hence, I do not intend to account for cross-cultural differences in the experience of harmony.

Yet, it should be noted that – as Dmitri Tymoczko points out – the term 'tonal' is "contested territory".[1] This is because there is no universal agreement as to which traditions the notion of tonal harmony is meant to encompass. For the sake of clarity, I will adopt the broad definition of 'tonal harmony' offered by Kostka et al. as referring to "the harmonic style of music composed during the period from about 1650 to 1900".[2] That said, in my account of the aesthetics of harmony *vis-à-vis* some historical divergences in tuning and temperament, I will also consider earlier traditions in harmonic theory stemming from mediaeval practices and stretching as far back as Antiquity.

[1] Tymoczko, D., *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice*, Oxford: OUP, 2011, p. 3.
[2] Kostka, S., Payne, D., Almén, B., *Tonal Harmony*, New York: McGraw-Hill, 2018, p. xiv.

Part I

Ontology

Introduction

Chords are sounds. My account of the metaphysics of tonal harmony therefore begins with an analysis of the nature of sounds. Granted, one may speak of the audible qualities of sounds and how long they last; one may also make claims about how they came to be, what kind of objects may be involved in the process, and even suggest where such objects may be located. But the most fundamental of the questions – namely, what *kind* of entities those sounds are – seems far from being settled.

Hence, I will first consider the ontology of sound. The sound debate has recently been reinvigorated by a growing support for non-property theories of sound, which include both wave-based and event theories. After rejecting the 'Property View', I will examine the 'Wave View' and the 'Event View', in particular. Subsequently, I will describe what I will refer to as 'the problem of sound' as having two horns: on the one hand, the Wave and Event views have been portrayed as rival accounts; on the other hand, the Event View is also said to be incompatible with the view that sounds are objects (i.e. the 'Object View').

After characterising the problem of sound, in Chapter Two I will propose a thesis of 'sonic reconciliation', which is a third-way approach aimed at integrating the core intuitions underlying those views. Specifically, I will argue that it is possible to reconcile the fundamentals of the Wave View and the Event View, just as it is possible to reconcile the Event View with the Object View. Most importantly, I will maintain that my proposal is not only feasible but particularly illuminating for an account of the ontology of chords, which constitutes the primary focus of Part I.

My proposal in the ensuing chapters is thus threefold. In line with the thesis of sonic reconciliation, I shall first consider the nature of chords *qua* sound events in Chapter Three. Here, I will characterise the grounding base of chords and seek to establish the necessary conditions for chord events to exist. This will be followed by an analysis of the psychoacoustics and the phenomenology of chord perception in Chapter Four, where I give an account of the nature of chords *qua* Gestalt-based auditory objects. Yet, since chords are taken to be distinctively musical entities in the domain of tonal harmony, I shall also give an account of their nature *qua* musical individuals in Chapter Five, which further substantiates my treatment of the ontology of chords. Given the broad scope of the task ahead, and for ease of exposition, I shall outline my main arguments at the beginning of each chapter, whilst attempting to show their interdependence throughout Part I.

Chapter One: The Problem of Sound

The contemporary philosophical debate on sound is best understood in terms of the contrast between Property and non-Property views. While some have argued that sound belongs to the class of sensory qualities, non-Property theorists tend to characterise it either as a wave or an event. These, however, are often taken to be competing accounts, such that the 'Wave View' and the 'Event View' have come to be portrayed as irreconcilable. In this chapter, I will argue that the problem of sound arises from the nature of that debate. To that end, I shall first provide an overview of the current literature on sound ontology. Since the 'Property View' is usually seen as the traditional philosophical account, I shall indicate why conceiving of sounds in terms of property instantiation is a misguided approach. After providing an outline of the Wave View and of the most prominent iterations of the Event View in Section 3, I shall attempt to identify some reasons as to why a wholesale adoption of either view runs into difficulties – whilst also considering the most relevant counterarguments offered by some of their key advocates. It is only in Section 4, however, that I aim to specify what the problem of sound constitutes, in view of the shortcomings of upholding what would seem to be irreconcilable views.

1. The sound debate

There are some common intuitions about sounds that any philosophical account should aim to capture. These involve not only what can be said about sounds themselves but also of their relationship to their sources. The latter, in particular, is fundamental to what is usually referred to as the traditional philosophical view on sound, under which sounds are understood to be properties of objects. This view has nonetheless been rejected by those who tend to follow the received account from acoustics, whereby sounds are taken to be reducible to waves. Others, however, have more recently come to suggest that sounds are best construed as events, claiming that they are essentially time-taking entities. Hence, the ontology of sound has become the subject of intense dispute in the philosophical literature. The purpose of this section is to provide an overview of the three main strands of that debate, each of which shall be subsequently examined more closely in Sections 2 and 3.

One way of conceiving of the nature of sounds is by considering their relationship to their sources. This approach is particularly favoured by advocates of the view that sounds are properties of objects. Despite variations across different property theories of sound, there are some underlying intuitions that seem to be shared by most of them. First, Property theorists tend to appeal to one aspect of the phenomenology of auditory perception in particular, namely, that we hear objects as well as sounds. On this view, sounds must be ontologically tied to objects just as properties are to their bearers. Secondly, property theories are invariably based on the analogy between colour and sound; since colours are generally treated as properties of objects, then analogously that must be what sounds are.[1] Another common argument made by Property theorists is that sounds are properties because they are repeatable. Such view appeals to our folk phenomenology that sounds are reidentifiable, i.e. that like objects will make like sounds.[2] Given these intuitions, some have argued that the best way to account for the nature of sounds is to say that they are properties of objects. Taken together, these claims could thus be said to constitute the backbone of what Casey O'Callaghan refers to as the 'Property View'.[3]

While Property theorists focus on the relationship between sounds and sources, non-Property theorists emphasise certain aspects of sound perception which point to a different characterisation altogether. Under one type of proposal, sounds are said to be caused by certain events, and to be capable of conveying information about what made them happen. Underlying this assumption is the view that sounds have duration: they seem to come about as a result of some anterior event, they last for a certain period of time, and they tend to fade away. In addition, for the length of their duration, we are able to track variations in sounds. They may become loud or quiet, shrill or low-pitched, continuous or intermittent. They may be heard as blending in with other sounds, or as being masked by them, only to 'return' to their distinctive quality. When we think of sounds that way, it is their temporal properties that stand out. Hence, some philosophers take sounds to belong to a particular ontological category that best captures their essentially temporal nature; namely, that of events.[4] This broad characterisation is thus distinctive of the 'Event View' of sound.

[1] This is not an uncontroversial claim. In addition to being extensive, the literature on the ontology of colour is fraught with rivalling theories so the analogy is not as straightforward as some theorists would have it. Yet, this particular debate lies beyond the scope of the present chapter.
[2] The repeatability of sounds – or the possibility of 're-encountering' them – is arguably not so easily accounted for by non-property theories. This will be considered in greater detail in Sections 3 and 4.
[3] This terminology is found in O'Callaghan, C., 'Sounds and Events', In Nudds, M., O'Callaghan, C. (eds.), *Sounds and Perception*, Oxford: OUP, 2009, p. 27.
[4] That said, Roger Scruton questions whether we should speak of 'events' or 'processes', and whether there is a relevant distinction between them for the case of sounds. Scruton, R., *Understanding Music*, London: Bloomsbury, 2009, p. 20.

Conversely, if we go beyond the phenomenology of auditory experience and consider the physical process underlying sound production, we may instead choose to focus on the all-pervasive element in that process, that is, vibratory motion. This observation is what has led others to support a different kind of proposal, which is underpinned by the scientific account of sounds as pressure waves. On this view, sounds are disturbances in a medium, and they may or may not be heard as such depending on whether or not (a) there is an elastic medium actually being disturbed, e.g. air; and (b) there is some sort of observer able to perceive (or somehow audibly register) those disturbances. To that extent, the common practice is thus to speak of *sound waves* and their respective physical properties – such as frequency, amplitude and intensity – along with their relationships to the properties of the objects involved as well as those of the medium at stake. The most obvious consequences of this 'Wave View' are familiar, especially those concerning the absence of sound in vacuum and acoustic phenomena such as that of echoes, for instance. Essentially, this science-based account of sounds as waves – given in terms of the nature of their production and transmission by means of an elastic medium – provides the Wave View (as well as any theory of sound) with the fundamentals of the physical process underlying the mechanics of sound waves.

Yet, if that is as far as the scientific account goes, one may question how our folk phenomenology of auditory experience may be reconciled with the view that sounds are identical to pressure waves in a medium. Granted, one may assume that what physicists have to say about the nature of sound is sufficient only insofar as they are able to provide a description of the mechanics of sound wave production and propagation. But to think of sounds strictly in those terms may seem counterintuitive to some. If sounds are reducible to pressure waves, then an important question arises concerning the nature of auditory perception. Specifically, since we do not seem to experience sounds *as* pressure waves, then there must be something else that can be said about them. This being the case, it would appear that more can be said about the ontology of sound – in which case there is also a role for the metaphysician in such an investigation.

It is in light of those intuitions about the nature of sound and auditory experience that, I believe, the current philosophical debate is best understood. Most importantly for this chapter, they underpin some fundamental claims that have resulted in conflicting views. These can be divided into two main camps, depending on whether sounds are taken to be properties or not – as shall be considered in the next two sections.

2. Sounds are not properties

Since the so-called Property View is said to be the traditional philosophical account of sound, it seems appropriate to consider it first. It is worth noting, however, that the term 'Property View' is somewhat misleading, in that it suggests that there is a unifying conception of sounds as properties – which is not the case.[5] Although this view is frequently associated with John Locke, it has more recently been given new iterations in the works of John Kulvicki and Jason Leddington, amongst others. In this section, I will argue that sounds are not properties. While I shall discuss some of the claims made in the context of those works, it is not my intention to provide a comprehensive critical assessment of any account in particular. Rather, my ambition here is limited to identifying three main reasons as to why the view that sounds are properties – broadly conceived – is misguided.

It is commonplace in philosophy to associate the Property View with Locke's theory of secondary qualities. In *An Essay Concerning Human Understanding*, he defines 'quality' as "the power to produce any Idea in our mind", and ascribes it to "the Subject wherein that power is".[6] He then goes on to distinguish between primary and secondary qualities: while the former are the real and original ones – in that they are "utterly inseparable from the Body, in what estate soever it be" – secondary qualities consist in powers that are "in any Body, by Reason of *its* insensible primary Qualities, to operate after a peculiar manner on any of our Senses, and thereby *produce in us the different Ideas* of several Colours, Sounds, Smells, Tasts [*sic*], etc.".[7] Essentially, what Locke proposes is a unifying treatment of those putative sensible properties by rendering them all dependent on the real insensible qualities of objects. Most importantly, he emphasises that secondary qualities are "in truth nothing in the Objects themselves".[8] It is generally under this view that the traditional philosophical account of sounds as properties is frequently understood.[9]

---

[5] Pendaran Roberts has recently identified at least eleven respectable property theories of sound. Roberts, P., 'Turning up the Volume on the Property View of Sound', *Inquiry*, 60, 4, 2017, p. 344.
[6] Locke, J., *An Essay Concerning Humane Understanding*, London: Printed for Thomas Dring and Samuel Manship, 1694, p. 60.
[7] Ibid., pp. 60, 64. Locke also identifies a third type of qualities – vaguely designated simply as 'Powers' – which is not particularly relevant for the present discussion.
[8] Ibid., p. 61.
[9] That said, Robert Pasnau has argued that the Aristotelian distinction between common and proper sensibles is a more suitable iteration of the contrast between primary and secondary qualities. In his view, while primary qualities are best understood as "those sensible qualities that fall under (or just are) a determinable kind of sensible quality that can be readily detected through various sensory modalities", secondary qualities are those "that fall under (or just are) a determinable kind of sensible quality to which we have ready access only through a distinctive sort of sensation produced by a specialised sense". Pasnau, R., 'Sensible Qualities: The Case of Sound', *Journal of the History of Philosophy*, Vol. 38, 1, 2000, p. 28; and Pasnau, R., 'A Theory of Secondary Qualities', *Philosophy and Phenomenological Research*, 73, 3, 2006, pp. 585, 589.

One way of construing the relationship between sound (as one of those powers or secondary qualities described above) and an object is in terms of the notion of disposition. This understanding has most recently gained currency in Kulvicki's account of sounds as stable properties of objects. In his view, objects *have* sounds. By stable properties, Kulvicki means that sounds are objects' dispositions to vibrate in response to mechanical stimulation.[10] Specifically, he identifies sound with the object's natural frequencies of vibration – i.e. a medium-independent property which is revealed as and when the object is mechanically stimulated. Hence, on this view, sound is not a vibrating event or a pressure wave but an object's stable disposition to vibrate at a given frequency, which is only manifested under certain stimulus conditions.

This type of dispositionalist view[11] is not only compatible with Locke's taxonomy above but is equally conformable to some of the most common intuitions mentioned at the beginning of this chapter. Its fundamental claim is that sounds are dispositions, i.e. they are properties that are manifested if objects are appropriately stimulated. These dispositions are "relatively stable across media – air, water, even the vacuum – across many kinds of stimulation, and they endure even when the object is unstimulated and does not vibrate".[12] Furthermore, under Kulvicki's account, one may claim that it is possible to hear objects as well as sounds, in that "objects that have distinct natural frequencies tend to sound different, and we can identify and recognise objects on the basis of how they sound".[13] This being the case, such view is also compatible with the claim that sounds are repeatable. Arguably, one advantage of this variant of dispositionalism is that it is best equipped to explain the phenomenon of auditory constancy – i.e. that objects appear to sound similarly on different occasions and, according to Kulvicki, even across different media.

That said, this view runs into difficulties when we consider what seems to be a contrast between sound *qua* disposition and sound *qua* percept. Dispositionalist theorists like Kulvicki stress that sounds are present in objects as dispositions, such that the medium is solely the way through which they become audible. The corollary of this view is that sounds exist even when the medium is absent, or even when there is no way in which the

---

[10] Kulvicki, J., 'Sound Stimulants: Defending the Stable Disposition View', In Stokes, D., Biggs, S. (eds.), *Perception and Its Modalities*, Oxford: OUP, 2014, p. 206.
[11] As Roberts points out, dispositionalist views may focus on the object's vibration – such as Kulvicki's – or on the ensuing waves, instead. In the second case, sounds are taken to be dispositions to produce sound waves under certain generally specified conditions. Roberts, p. 345.
[12] Kulvicki writes of sounds as 'enduring' even in the absence of a stimulus. This seems like an unusual characterisation, in that we do not think of properties as having persistence conditions. The question of sound persistence will be considered in Chapter Two (Section 4). Kulvicki, J., 'The Nature of Noise', *Philosophers' Imprint*, 8, 11, 2008, p. 5.
[13] Ibid., p. 5.

sound may be perceived. Yet, this means that the medium is a necessary condition for sound to be revealed,[14] but not for it to *exist*. Hence, the implication is that there is a difference between sound *qua* disposition (which exists whether manifestly or not) and sound as that which we are able to hear.

Kulvicki tries to counter this objection by stating that the sound an object has and the sound perceived by the observer are the same. He argues that "there are not two kinds of sound – sounds made and sounds had – there are just sounds and the occasions in which we can hear them".[15] In his view, the sounds we hear are identical to the stable dispositions that objects have, which are revealed by mechanical stimuli and made audible when in the presence of a medium. However, in maintaining that claim, Kulvicki seems to conflate the disposition with the actual event of an object vibrating after being stimulated. On the one hand, he argues that his conception of sounds as stable dispositions should not entail that an object "makes a disposition to vibrate when it is thwacked"; on the other hand, he also states that "objects make sounds, or sound off, when they vibrate audibly in a medium: when we can hear them".[16]

But what these claims seem to reveal is that the variant of dispositionalism found in Kulvicki's account tells the story of sound production 'backwards', as it were. By maintaining that sounds *qua* stable dispositions are identical with what we hear, Kulvicki seems to suggest that sounds exist prior to there being any vibratory motion whatsoever, as though an object's natural frequency were a necessary and sufficient condition for its sound to exist. This would render the presence of a medium a contingent factor which only serves the purpose of revealing the sound the object already has, whilst also dismissing the role of the observer in perceiving it. Most importantly, despite Kulvicki's attempt to avoid the charge of postulating two different entities, it seems that his account does just that. It draws a line between the sound *had* by an object *qua* property – i.e. the stable disposition to vibrate – and the sound *made* by the object as it vibrates – i.e. the sound that we are able to perceive unless the medium is wanting. Furthermore, by identifying sounds *qua* dispositions with an object's natural frequencies of vibration, his view seems to conflate the physical properties that constitute the dispositional base with the disposition itself, as if they were 'yet-to-be-revealed' sounds.

---

[14] The view that the medium 'reveals' the sound (and that sounds exist in a vacuum but are not revealed as such) is not exclusive to property theories of sound. Roberto Casati and Jérôme Dokic, for instance, have subscribed to such a view, even though they are Event theorists – as shall be considered in the next section.
[15] Ibid., p. 6.
[16] Ibid., p. 6.

This first objection also raises an important question. If sound is both the disposition and the perceived entity as it is revealed by the medium, then a further issue of conflation arises, this time in the semantic dimension. Briefly, it seems at best unclear what the actual referent for 'sound' should be. By emphasising that sounds are qualities of objects that may or may not be revealed, dispositionalist theorists would seem to favour the view that 'sound' primarily picks out an object's property of being such as to sound off at a particular frequency. If this is accepted, then 'sound' may or may not be used to refer to the percept, depending on whether or not a manifestation of that disposition is forthcoming. Hence, if this interpretation is correct, the locution 'the sound of the clock', for instance, fundamentally refers to sound *qua* disposition. Yet, this similarly leads to the counterintuitive consequence as expressed above; in the example given, it does not seem at all clear whether the ticking sound is a sound that a given clock has as a stable disposition to vibrate, or whether it is the sound made by the clock as it vibrates.

That said, the observations made thus far would only seem applicable to the dispositionalist interpretation of the Lockean view that sounds are powers belonging to objects – a variant of which is found in Kulvicki's account of sound dispositionalism, as seen above. Conversely, other recent property theories offer alternative views.[17] One such account is that proposed by Leddington, who claims that Property theorists tend to ignore an important aspect of our descriptions of objects' sounds – namely, that we normally speak of sound *sources*, instead. Once again, the question of the nature of the relationship between sounds and sources resurfaces, which is at the root of the second reason why I take the Property View to be flawed.

Property theorists are not in agreement as to what sounds are meant to be properties of; for instance, while Kulvicki writes of objects as having sounds, Leddington proposes that sounds are properties of events, instead. Under his 'Event-Property View', sounds are audible properties of their event sources – an approach which he describes as a variant of Locke's view. Although his main target are two distinct event theories of sound – which shall be considered in the next section – Leddington proposes an 'ockhamisation' of the Property View by taking sounds to be "*constituted* by the audible qualities of their sources"– i.e., pitch, timbre, loudness and duration.[18]

[17] For instance, Jonathan Cohen suggests that sounds could be properties of "spatio-temporal regions that are occupied by (and so not identical to) portions of the sound-transmitting medium", such that "the property theorist might suggest that the sound is exemplified by the spatio-temporal region occupied by its source". Although Cohen only briefly hints at this as a possibility, his suggestion indicates that the claim that objects themselves *have* sounds may not be as simple as Kulvicki would have it. Cohen, J., 'Sounds and Temporality', In Zimmerman, D. W. (ed.), *Oxford Studies in Metaphysics*, Vol. 5, Oxford: OUP, 2010, p. 306.
[18] Leddington, J., 'Sounds fully simplified', *Analysis*, 79, 4, 2019, p. 624.

Arguably, one of the advantages of Leddington's Event-Property view is that it can accommodate the claim that sources cause sounds. As seen earlier, one of the problems with dispositionalist accounts of sound – which fall under what Leddington dubs as the 'Object-Property View' – is that they are not so easily conformable to the common intuition (and corresponding linguistic descriptions) that objects *make* sounds, rather than having them. Conversely, Leddington claims that his Event-Property View is able to address that difficulty by considering the two ways in which we use the term 'sound': on one reading, we use it to refer to a repeatable universal, whereby distinct events are said to have the same sound; but, on another reading, 'sound' would refer to a particular instance of that universal. In his view, it is under the semantics of the latter that we speak of event sources as making sounds – i.e. sound-instances, rather than sounds as universals.

Yet, this is precisely where Leddington's thinking encounters some difficulties. For the metaphysics of causation to work satisfactorily in his account, he needs what he calls a 'relaxed view' whereby not only events but also properties *qua* instances are eligible for the role of causal relata. In his words, sounds – understood as the audible qualities themselves – "qualify the very events that cause them to be instantiated".[19] Unlike what Leddington may be inclined to accept, however, this would constitute too much of an infelicity to be so easily dismissed. This is because the view that events cause their own properties would require a rather unorthodox account of property instantiation – rather than just a liberal approach to possible candidates for causal relata. Furthermore, by reducing sounds to audible qualities which are normally ascribed to pressure waves, Leddington's proposal could be otherwise seen as a veiled argument in support of the Wave View, instead.

Over and above that, Leddington's proposal seems unmotivated. If we were to conceive of sounds as consisting precisely in the audible properties that are usually ascribed to those putative sound-instances, a more obvious way of simplifying the ontology of sound would be to stand by the physicists' description of sounds as waves, since it is these that bear the corresponding physical properties underlying our experience of pitch, loudness and timbre.[20] O'Callaghan seems to share this view when he states that "whether audible qualities are dispositions, physical properties, primitive qualities, or projected sensory qualities, their instances depend upon the physical properties that explain the occurrence of audible quality experiences"[21] – properties which are ordinarily associated with waves, instead.

---

[19] Ibid., p. 625.
[20] The case of timbre is particularly complex, as shall be indicated in Chapter Three (Section 2).
[21] O'Callaghan, C., *Sounds: a Philosophical Theory*, Oxford: OUP, 2007, p. 74.

Hence, unlike what Leddington suggests, whether we take sound sources to be objects or events, it would be a strange consequence of his view that we should understand them as causally making their own properties. Similarly, contrary to what Kulvicki has proposed, it would be just as counterintuitive to speak of sources as having sounds, on whatever count, in that they emit sounds. Whether they consist of objects or events, sources do not have sounds – they make sounds. In Chapter Two, I shall argue that this is not a trivial matter of semantics, but rather indicates that the essence of the relationship between sound and source is a causal one, and not a case of property instantiation.[22]

That said, while the objections above are targeted at attempts to account for the nature of sounds as properties in terms of their relationship to their sources, one might argue that their scope is too narrow to undermine the essence of the Property View as broadly conceived. Hence, my last argument against the Property View is based on the very claim that sounds are *properties*, whatever their sources turn out to be. Put simply, I take this claim to be both unmotivated and counterintuitive.

First, it is unmotivated because the ontological tie between sounds and sources that Property theorists seek to maintain is unnecessary for the ontology of sound to be made comprehensible. Both Locke and Kulvicki accept that it is the presence of certain physical properties in objects – the 'real' primary qualities or the natural frequencies of objects, respectively to their accounts – that constitute the base for those powers/dispositions. However, under an uncontroversial understanding of causation as well as the received account from acoustics, those intermediary properties become redundant. In other words, the presence of certain physical properties is both necessary and sufficient to explain why certain objects have certain acoustic properties – such as their natural frequencies – but these do not consist in latent sounds that may be perceived under certain conditions. Rather, the perceived sound is the consequence of the auditory system's ability to encode the information contained in vibratory motion, which results from the transfer of energy between the object and an elastic medium when the former is caused to vibrate. This, however, does not require the presence of any additional powers or dispositions. Understood within this framework, the traditional philosophical account can be said to postulate more entities than necessary, whilst simultaneously attempting to oversimplify the process of sound production by construing sounds as properties of their sources, instead.

[22] Leddington's reference to sound-instances mentioned earlier seems correct in pointing out that the locution 'the sound of ...' picks out two different categories of 'sound'. Yet, since I do not share his view that sounds are properties, in Chapter Five (Section 1) I will characterise the multiple instantiation of sound in terms of the type/token distinction under my analysis of chords *qua* musical individuals, instead.

Secondly, the view that sounds are properties is counterintuitive. The main difficulty that besets the Property View lies in the very conception of sound as a property, whether one takes it to be a power that it has to produce certain sensations or perceptions in an observer (Locke), or a disposition to vibrate in a certain way (Kulvicki), or even as a constellation of audible qualities of events (Leddington). The reason why I take property theories to be counterintuitive concerns the nature of what it means for something to have a certain property – a question that has also been considered by Roberts in his attempt to revive the Lockean view of sound. As he himself points out, a property is a way that something is.[23] But this is precisely what is so counterintuitive about the Property View: sound is *not* a way that something is; rather, it is the effect of what that something is doing, which occurs as a result of the mechanics of wave production and propagation.

Indeed, the widespread use of the verb 'to make' when describing how sounds come about points to the intuition that sounds are not properties of objects. Yet, this is not simply a matter of linguistic practices – whether one claims that sounds are made or 'had' by objects – but it concerns a crucial distinction between the way things are (i.e. properties) and what objects happen to be doing (as a consequence of vibratory motion). The intuition that objects make sounds indicates an underlying causal process that enables sounds to come into being, rather than being manifestations of an object's pre-existent quality. In short, sounds are not ways things are – they are not properties of objects.

If the objections raised above are successful, then the sound of a clock in the earlier example should not be treated as a property thereof. It is not a quality that is revealed when the clock strikes five, for instance, neither is it a disposition that it has to vibrate in a way that we perceive as 'tick-tock'. It is not a power that the clock has to produce a certain sensation in our auditory system, or a property of the event of the clock's hand reaching the number five on the display. Clocks of the same manufacturing standard may be made to ring at the same time; the sounds they make may bear similarities but they are not repeatable universals or instantiations of a single property. In assessing the similarities and differences between those sounds, we do so by treating them as individuals in their own right, as bearing certain distinctive properties. It is in view of these claims and on the basis of the objections presented above that I take the Property View to be misguided, however it may come to be dressed.

[23] Roberts, p. 339.

3. Non-property theories of sound

If the objections raised in the previous section succeed in undermining the essence of property theories, then two other ways of conceiving of the nature of sound would seem to remain. On the one hand, sounds are taken to be waves. The Wave View has recently gained support in the works of Roy Sorensen and Matthew Nudds, having also been previously considered by George Berkeley. On the other hand, some philosophers have highlighted the distinctively temporal nature of sounds and characterised them as events. Different iterations of the Event View have been proposed by O'Callaghan, Scruton as well as Roberto Casati and Jérôme Dokic. The aim of this section is to provide an outline of the Wave and Event views, with reference to their respective works. This, in turn, will inform my characterisation of the problem of sound in Section 4.

3.1. Sounds as waves

As indicated earlier, one of the ways in which one may seek to understand the nature of sound is by considering what can be said of the mechanics of sound wave propagation. In pursuing such an account, one is likely to identify sounds with disturbances in an elastic medium. The standard scientific account is familiar: sounds consist in successive changes in the equilibrium of the medium, with alternated phases of compression and rarefaction (high and low pressure, respectively). Upon reaching an observer, these disturbances are 'interpreted' by the auditory system as a percept with distinctive pitch, loudness and timbre.

Despite the familiarity and ubiquity of this science-based view, the last quarter of a century has seen a renewed interest in the metaphysics of sound, on the basis that the Wave View does not seem to account for the nature of sounds *as we hear them*. In addition, what is all too frequently characterised as the scientific view is itself a less nuanced version of what psychoacousticians understand sound to be. For instance, in an introductory guide on acoustics for the hearing and speech sciences, Charles E. Speaks provides a twofold definition of sound: from a physical perspective, sounds are waves; by contrast, from a psychological perspective sound may also be construed as that which we are able to perceive by means of the auditory system.[24]

[24] Speaks, C. E., *Introduction to Sound: Acoustics for the Hearing and Speech Sciences*, San Diego, CA: Singular Publishing, 1999, p. 3.

Perhaps unsurprisingly, it is the psychological perspective that is fertile ground for philosophical enquiry. In pursuing that view, one may choose to question the apparent simplicity of the scientific account. Yet, what Speaks's distinction does not explicitly convey is the tacit acknowledgement that the physical base has ontological priority over the psychological response. This is arguably the underlying principle enshrined in the Wave View, which has more recently been articulated in the works of Sorensen and Nudds but may also be found in Berkeley's statement of the argument from vacuums.

One of the well-known empirical facts about sounds is that they do not occur in a vacuum. In philosophical terms, the argument from vacuums can be traced back to the first of Berkeley's *Dialogues between Hylas and Philonous*. In a reply to Philonous, Hylas states it as follows: "when any motion is raised in the air, we perceive a sound greater or lesser, in proportion to the air's motion; but without some motion in the air we never hear any sound at all ... [since] it is this very motion in the external air that produces in the mind the sensation of sound". This being the case, Hylas concludes that "a bell struck in the exhausted receiver of an air-pump sends forth no sound".[25] If sounds are waves,[26] then striking the bell in a vacuum makes no sound; that is to say, in the absence of a medium there cannot be any sound being produced by the bell, since there cannot be any disturbance of the medium if there is none. Hence, unlike the claim made by Property theorists, the argument from vacuums entails that the presence of a medium is a necessary condition not only for wave propagation but for the very existence of sound. Furthermore, it arguably provides the Wave View with observable evidence for individuating sounds in terms of pressure waves in an elastic medium.

Despite its widespread acceptance as a scientific fact, the corollary of the argument from vacuums has found some resistance across both sides of the sound debate. This is particularly the case with those who favour the thesis of revelation. As seen in Section 2, Kulvicki argues that the absence of a medium does not entail that there is no sound, but solely that the object's sound cannot be revealed as such. Similarly, Event theorists such as Casati and Dokic have also maintained that vibrating objects resonate independently from the presence or absence of a medium, to the extent that being immersed in it does nothing but reveal the occurring sound to the observer. Hence, if the role of the medium is solely that of revealing the sounds of objects, then the argument from vacuums does not entail that sounds can only ever exist in the presence of an elastic medium.

[25] Berkeley, G., *Three Dialogues between Hylas and Philonous*, New York: Bobbs-Merrill, 1954, pp. 20-21.
[26] It should be noted that Berkeley never endorsed this view, given the immaterialist nature of his philosophy.

However, as O'Callaghan has pointed out, there are "good philosophical reasons" for accepting the argument from vacuums, on the basis that there is "no suitable ascription of audible qualities to sounds" in those circumstances.[27] Yet, he rejects the claim that sounds are located in the medium, which he regards as one of the shortcomings of the Wave View. O'Callaghan argues that such a claim indicates that the auditory system falls prey to a radical illusion concerning sound location. Specifically, if sounds are waves, we would then constantly misperceive sounds as located at a distance from us. This is because, in O'Callaghan's view, auditory experience informs us that sounds are *distally* located, i.e. at or near their sources, rather than in the medium.[28] Echoing the same concern, Pasnau maintains that the standard scientific view is incoherent: while it states that sounds travel through the medium, it also holds that sounds that are caused at a distance appear to us as being at a distance, rather than travelling towards us.[29] Thus, if distal theorists are correct, then the Wave View is an error theory of auditory experience.

Indeed, this issue of sound location has to an extent dominated the sound debate. In response to the charge of locational inconsistency, Sorensen has defended the Wave View by seeking to explain away the apparent perceptual illusion described above. Drawing on a phenomenological distinction between locating by centres and by edges, he claims that the latter is not available to auditory perception because the 'edges' of sounds are unknown.[30] Instead, similarly to the visual representation of sound as a series of concentric circles emitted by a source at its centre, he argues that the 'shape' of a sound is best construed as a fast-growing sphere that rapidly encompasses those that are able to perceive it as such. However, unlike our ability to perceive the edges of water waves rippling on in a pond, the speed at which sound waves are propagated – as well as their invisibility – prevents us from experiencing the edges of sounds. Hence, the upshot of Sorensen's argument is that we can only perceive sound location in terms of *source* location. As a result, we should only speak of it in those terms. However, stated as such, it is not clear how Sorensen's counterargument would defuse the charge of locational inconsistency, in that distal theorists would not disagree with that observation. Nevertheless, it is in his comparison between the spatial phenomenology of seismic perception and the case of audition that his defence of a wave-based, medial theory of sound becomes more refined.

---

27 O'Callaghan, 2007, p. 51.

28 Ibid., pp. 30-31. 'Distal' theorists argue that sounds are located at a distance from the observer, whereas 'medial' theorists take sounds to be located in the medium.

29 Pasnau, R., 'What is Sound?', *The Philosophical Quarterly*, 49, 196, 1999, p. 311.

30 Sorensen, R., 'Hearing Silence: The Perception and Introspection of Absences', In Nudds, M., O'Callaghan, C. (eds.), *Sounds and Perception*, Oxford: OUP, 2009, p. 138.

Sorensen claims that we experience sounds much in the same way as we experience earthquakes. "If the earthquake is the train of seismic waves emanating from the hypocentre", he writes, "then the quake is in its medium and so encompasses a wide area. These waves are moving *away* from the hypocentre. One wave front briefly heads *toward* the epicentre but then spreads out from there".[31] Albeit on a significantly different scale of magnitude, the acoustic and seismic cases are, in his view, similar because their phenomena overlap: just as we are not able to perceive the edges of the quake, neither do we experience the edges of sounds. Hence, for Sorensen, the orientation towards the centre (or the source, in the case of sound) results from our inability to perceive the spatial boundaries of quakes (as of sounds). In either case, he claims that their actual location is, nonetheless, in the succession of waves that pass through the medium.

Similar concerns regarding the implications of the spatial phenomenology of auditory experience for the Wave View have been considered by Nudds, who offers an alternative solution to the problem of locational inconsistency. In 'Sounds and Space', he questions whether sounds can be at all located in space. Unlike Sorensen, however, Nudds attempts to reject the challenges posed by distal theorists against the Wave View by claiming that there is nothing intrinsically spatial about the nature of sounds, for they cannot be spatially individuated. "Sounds", he maintains, "do not have any *intrinsic* spatial significance and do not have any spatial structure".[32] Although this may seem *prima facie* counterintuitive – since it is arguably commonsense to assume that auditory experience provides us with an awareness of sounds' spatial properties – Nudds's response to this kind of objection is unequivocal: spatial location is not a property of sounds but, rather, of their sources.

In his own attempt at providing an account of the metaphysics of sound, Nudds emphasises the physical aspects of sound production and transmission that give rise to the experience of what we call 'sound'. He describes sounds as auditory objects represented in our experience as "patterns or structures of frequency components instantiated by the sound waves that are detected by the ears".[33] Under his view, upon reaching an observer those frequency components are processed by the auditory system, which interprets the acoustic properties of those vibrations and attributes them to a particular source. Although Nudds describes sounds as instantiated where we are – in the sense that they are heard where we are – he argues that sounds themselves are not located anywhere in space.

---

[31] Ibid., pp. 138-139.
[32] Nudds, M., 'Sounds and Space', In Nudds, M., O'Callaghan, C. (eds.), *Sounds and Perception*, Oxford: OUP, 2009, p. 80.
[33] Ibid., p. 75.

In order to account for the view that sounds may be encountered not in space but at multiple times, Nudds describes them as belonging to the ontological category of 'particularised types' or 'abstract individuals'.[34] Yet, this seems to raise a few difficulties. Although Nudds's thesis of the non-spatiality of sound provides a solution to the problem of sound location, by conceiving of sounds as particularised types that are instantiated by waves his proposal seems vulnerable to similar objections as those directed against the Property View. This is because it could be construed as taking sounds to be properties of waves, instead. In addition, it is somewhat unclear to what extent those categories shed light on the nature of sound, in that Nudds does not substantiate them any further.

That said, the empirically-based explanation that Nudds offers to account for his thesis of the non-spatiality of sounds is a strong argument against distal theories. Rather than being allocated to a single source in virtue of the alleged spatial properties of sounds, he notes that the frequency components of waves are grouped by the auditory system in accordance with non-spatial cues (e.g. shared onset times and harmonic relationships), which reflect the physical nature of the sound source (i.e. its mass and elasticity) and of the sound-producing event (such as the force involved). For that reason, Nudds maintains that our experience of sounds must be understood in terms of the function of the auditory system, which is that of informing us about sound sources and their properties, as well as the nature of the event that produced the original disturbance.[35]

This functionalist approach to the phenomenology of auditory experience is thus the cornerstone of Nudds's argument for the non-spatiality of sounds. In his view, the function of auditory perception is not to tell us about sounds themselves. Rather, it is to extract and interpret information about the object and/or the sound-producing event by grouping together "all and only the frequency components that are likely to have been produced by the same source"[36] – a process which, he insists, is not intrinsically spatial. Consequently, any spatial awareness that we may have in auditory perception must be the result either of a bi-modal experience (i.e. involving an additional sense, most likely vision)[37] or of a purely auditory experience of a particular sound as having been produced by a source that is located somewhere in space. In either case, Nudds's conclusion is that it is the sources that have spatial properties, not sounds; the latter, he argues, only do so in a contingent way, to the extent that sounds "need not have any spatial properties at all".[38]

---

[34] Ibid., p. 76.
[35] Ibid., pp. 78-82.
[36] Nudds, M., 'What are Auditory Objects?' *Review of Philosophy and Psychology*, 1, 1, 2010, p. 116.
[37] Nudds, M., 'Experiencing the Production of Sounds', *European Journal of Philosophy*, 9, 2, 2001, p. 220.
[38] Nudds, 2009, p. 77.

The problem of sound location is thus addressed differently by Nudds and Sorensen, despite the fact that they are both taken to be advocates of the Wave View. While Sorensen argues for a medial theory, Nudds favours the thesis of the non-spatiality of sounds. Another important contrast between their views is that, unlike Sorensen, Nudds is not committed to the claim that sounds are identical with waves, restricting his interpretation to the notion of the multiple instantiation of sounds by waves. To that extent, Nudds believes that the conjunction of the two theses – namely, that of multiple instantiation and that of non-spatiality – enables the Wave View to address the problem of locational inconsistency, whilst also conforming to our normal ways of experiencing and individuating sounds.

Nevertheless, those who are reluctant to accept that what we hear can be sufficiently characterised by an understanding of the mechanics of wave production/propagation and of the frequency-grouping process have pointed to other weaknesses of the Wave View. For instance, Casati and Dokic have argued that it leaves at least three problems unresolved. First, it does not address a broader concern around the difficulty of distinguishing between audible and inaudible vibrations – which they take to be best accounted for by the notion of revelation, as seen earlier. In their view, this inability indicates that the question of audibility cannot be settled by acoustics alone. Second, the Wave View fails to account for cases in which the properties of the waves do not match our experience of the sound heard. For instance, a loud noise heard at a distance is still heard as being loud, despite the fact that the amplitude of the respective waves that reach the observer is reduced. Third, it does not consider the possibility that sounds are disturbances that happen to certain objects – a claim that forms part of their own theory of sound,[39] as shall be outlined in the next section.

Two further difficulties besetting the Wave View have been identified by O'Callaghan, who criticises it for its metaphysical obscurity.[40] First, he suggests that it could imply that sounds are properties of the medium. This is because what we refer to as 'waves' is nothing but patterns of alternated phases of pressure variation of sections of the medium over a period of time. According to O'Callaghan, in identifying sounds and their qualities with the pressure waves and their physical properties, Wave theorists are effectively construing waves as properties of the medium. Such an approach would therefore fall prey to similar challenges as those faced by Property theorists.[41]

---

[39] Casati, R., Dokic, J., *La Philosophie du Son*, Nîmes: Chambon, 1994. Some individual chapters were later given an English translation in 2009, so any specific references to them will be made accordingly.
[40] O'Callaghan, 2007, p. 24.
[41] This is in line with my objection to Nudds's category of abstract particulars/particularised types above.

Secondly, O'Callaghan questions whether we should speak of sound waves as the *objects* of auditory perception. He concedes that waves – taken as particulars rather than as properties of the medium – could be said to have certain object-like characteristics, such as spatial boundaries and the ability to survive qualitative changes. Yet, O'Callaghan argues that waves are not only ontologically dependent upon the medium but, crucially, that they happen to the medium. To that extent, he states that "the existence, propagation, and boundaries of the wave depend on processes that occur in and essentially involve a medium"; hence, "to highlight the medium dependence of the wave and its attributes is to highlight the wave's event-like characteristics".[42] In his view, the pressure waves that Wave theorists identify with sounds are thus better understood as events. However, even if this claim is accepted, it is not clear why we should take waves to be events that happen *to* the medium.

But if O'Callaghan's analysis is correct, one may wonder how the Event View came to be pitched as a rival to the Wave View, rather than solely as an alternative to the Property View, instead. Indeed, it is surprising that all of the recent proponents of event theories – Casati and Dokic, Scruton, and O'Callaghan himself – have without exception rejected the Wave View. In its place, they have offered different iterations of the Event View – as shall be considered in the next section.

3.2. Sounds as events

As indicated earlier, the renewed interest in the philosophy of sound is reflected in the growing support for the view that sounds are events, which I will seek to outline in greater detail in this section. The theoretical roots of the Event View can be traced back to Casati and Dokic's *La Philosophie du Son* (1994) and Scruton's *The Aesthetics of Music* (1997), but it has gained further support in O'Callaghan's *Sounds: a Philosophical Theory* (2007).[43] However, Event theorists have come to understand sounds *qua* events differently: while Casati, Dokic and O'Callaghan have taken a physicalist approach, Scruton has emphasised the phenomenal character of sounds, instead. The purpose of this section is thus to provide a broad characterisation of their individual accounts, and of the ways in which they differ.

---

[42] Ibid., p. 26.
[43] Scruton, R., *The Aesthetics of Music*, Oxford: OUP, 1997. Full reference details for the other two titles have been provided in earlier footnotes.

In their seminal work, Casati and Dokic set out to defend the view that sounds are not qualities of objects, but events that happen to them where they are. For example, bell chimes happen to the bell and are located where the bell is. Their distal event theory is centred on the role of the source, on the basis that sounds "involve primarily the space occupied by the resonant object".[44] On this view, unlike waves, sounds do not propagate across sections of the medium. Rather, just as the initial disturbance happens to the source and does not move away from it, neither do sounds. Hence, they cannot be identified with disturbances in the medium. As seen earlier, Casati and Dokic maintain that the medium solely reveals the sound as an audible percept. In their view, whether or not the medium is present, if there is vibratory motion at the source then the event of sound is taking place. Using the example above, they maintain that a bell that is struck in a vacuum makes a sound, even if it cannot be heard.

As indicated in the previous section, Casati and Dokic are not troubled by the argument from vacuums. Using the example of a tuning fork in a vacuum-jar, they argue that it is reasonable to conceive of the tuning fork as making a sound provided it continues to vibrate within it. In their view, inserting air in the jar would only reveal the sound that the tuning fork is making.[45] However, whilst dissociating sounds from waves, Casati and Dokic both acknowledge that we can only acquire perceptual information about sounds in case the medium is present to reveal them. Yet, this once again raises the question as to what kind of percepts sounds are. In response, Casati and Dokic argue that "even if sounds have phenomenal qualities, these qualities are not essential to their definition".[46]

Such response is in line with their physicalist iteration of the Event View, whereby sounds are taken to be physical entities that may or may not be made audible. For instance, under their proposal both infrasounds and ultrasounds would be classed as sounds even though they are not audible entities – but they maintain that they would not perform the role that audible sounds have as "perceptual intermediates".[47] This is because, on their view, we are only able to hear sources in virtue of hearing the sounds that happen to the relevant objects – which is not the case with infra/ultrasounds. By associating sounds with their sources from a physicalist yet medium-independent perspective, Casati and Dokic thus attempt to explain why sounds can inform us about the nature of the objects that produced them without depending on the nature of the medium at stake.

---

[44] Casati, R., Dokic, J., *La Philosophie du Son*, Nîmes: Chambon, 1994, English translation, 2009, p. 11, URL = https://jeannicod.ccsd.cnrs.fr/ijn_00420039.
[45] Ibid., p. 8.
[46] Ibid., p. 13.
[47] Ibid., p. 12.

However, the view that there are occurring yet unheard sounds has, perhaps unsurprisingly, attracted some criticism. By taking sounds to be physical entities and maintaining that sounds may occur without being audible, Casati and Dokic's view seems to dissociate sound from perceptual experience. In response, Scruton argued that their defence of physicalism means that they "bravely accept the consequence that sound is essentially non-phenomenal, in other words that what a sound essentially is has nothing to do with how it sounds".[48] Hence, Casati and Dokic's event theory seems to clash not only against the received scientific view that sounds do not occur in a vacuum but also against the folk phenomenology of auditory experience about what makes a sound a 'sound'.

Although these concerns would seem to undermine any attempt at providing a physicalist basis to the Event View, they have not deterred O'Callaghan from defending it. Specifically, he takes sounds to be non-mental particulars which stand in causal relations and have both duration and spatial location, as do waves.[49] Similarly to Casati and Dokic, O'Callaghan defends a distal theory of sound and maintains that the category of 'event' best captures its distinctively temporal nature. This is because sounds are time-taking particulars that have distinctive identity, individuation and persistence conditions.[50] O'Callaghan maintains that the Event View is able to accommodate the fact that sounds are primarily described as happenings and occurrences (e.g. the sound of a collision). In other words, he suggests that the emphasis on duration is evident in the way we think and speak of sounds, which would make the Event View conformable to our linguistic intuitions as to how sounds should be described. In O'Callaghan's view, these linguistic intuitions mirror the perception that sounds are essentially temporal entities, which is one of the aspects of sound ontology that sets them apart from other objects of perception.

In addition to accounting for the temporal nature of sounds, O'Callaghan argues that the Event View can equally accommodate the spatial dimension of sound perception. He claims that the distal view conforms to the folk phenomenology of auditory experience, in that "sounds are heard to be roughly where the events that cause them take place" – a claim which he describes as "empirically supported, introspectively discernible, and sometimes revealed in ordinary language".[51] In addition, he indicates that, by construing sounds as particulars (as opposed to properties), the Event theorist recognises sounds as property bearers, since they appear to have certain audible qualities themselves.

---

[48] Scruton, 2009, p. 23.
[49] O'Callaghan, 2007, p. 59.
[50] Ibid., p. 10.
[51] Ibid., p. 33.

One of O'Callaghan's proposed desiderata that the Event View is meant to meet concerns the causal relations behind the phenomenon of sound production. This is unsurprising, in that his physicalist account requires an explanation of sounds' causal role if they are taken to be distinct from waves; but his account is markedly different from Casati and Dokic's in this respect. He argues that sounds are "events of oscillating or interacting bodies disturbing or setting a surrounding medium into wave motion", and that they are "causally intermediate between ordinary, everyday events and travelling sound waves".[52] On this view, the causal relata are events that involve both the sound source and the medium in which it is located. By construing sounds as causally intermediate, O'Callaghan may be taken to suggest that sound is the medium-disturbing event. As such, it is "not constituted by those two events, but it shares constituents – the object and the medium – with those events".[53] Hence, in his view, the sound event is part of a larger event that encompasses the event of the initial disturbance as well as the waves. These, in turn, are the effect of the disturbing event that is the sound, and thus carry information about it which may then be processed by the auditory system. Also using a tuning fork as example, O'Callaghan states that "the event of the tuning fork's disturbing the medium" is what its sound is.[54]

Yet, the main point of departure between the two event theories considered thus far concerns the very nature of the sound event within the causal process. In 'Hearing Properties, Effects or Parts?', O'Callaghan refines his position by advocating that sounds should not be seen as causally related to audible sources; rather, sounds are audible entities related *mereologically* to them. In other words, they must be understood in terms of a part-whole relation, in that "sounds audibly are constituent *parts* of everyday audible events, such as collisions and vibrations, which involve material bodies".[55] One of the advantages of this view, according to O'Callaghan, is that it preserves the independence of sounds from their sources: just as one is able to perceive parts without perceiving wholes, one is equally able to hear sounds without hearing the broader event encompassing it. In addition, it accounts for the perception that sounds may appear as bound to their sources, given that audible sounds *qua* parts and audible sources *qua* wholes "are not *wholly* distinct".[56]

---

[52] Ibid., pp. 59-60.
[53] Ibid., p. 66.
[54] In spite of this somewhat contorted description of the sound made by a vibrating tuning fork, O'Callaghan still maintained this characterisation in a recent work. Writing of the disturbing event of the tuning fork, he states that "its creating the disturbance constitutes the tuning fork's sounding". O'Callaghan, C., *Beyond Vision*, Oxford: OUP, 2017, p. 23.
[55] O'Callaghan, C., 'Hearing Properties, Effects or Parts?', *Proceedings of the Aristotelian Society*, New Series, 111, 2011, p. 395.
[56] Ibid., p. 396.

Contrastingly, Casati and Dokic (along with Elvira Di Bona) have proposed that this 'parthood version' of the 'distal-event' view of sound can be further simplified. In 'The Ockhamisation of Event Sources of Sound', they suggest that the Event View is best understood in terms of an identity relation. Rather than a proper part of the event source, they take sound to be *identical* to it. For instance, in their example of a collision, sound is not a part of the collision (as O'Callaghan would have it) but the event of the collision itself. In other words, they take the collision and the sound to be one and the same thing – a view which they consider to be a "metaphysical ockhamisation" of O'Callaghan's account.[57] This is because, as they state it, "there is no need of preserving the distinction between sound and event source, [and so] talk of 'event source' is redundant".[58]

This is an important development in Casati and Dokic's event theory – and with extended ontological commitments – as they propose that hearing sounds is tantamount to hearing distal audible events. Yet, their attempt at a 'metaphysical ockhamisation' has recently been challenged by a Property theorist. In his defence of the Event-Property View, Leddington points out that one of the major infelicities of the 'Identity View' espoused by them is the implication that experiencing events by means of other sensory modalities (primarily vision) is equivalent to saying that sounds may be seen. In this case, in witnessing a collision one sees a sound – which Leddington has appropriately described as a bizarre consequence of their theory.[59] Similarly, the Identity View entails that deaf people are able to experience certain sounds as such by perceiving certain events taking place (e.g. when standing next to loudspeakers). Hence, the upshot of Leddington's objection is that what Casati et al. might be proposing goes against some of our most common intuitions about sounds, so they do not give a satisfactory account of the ontology of sound *as such*.

Nevertheless, the Event View has also gained support in Scruton's work; but his iteration of it considerably differs from those of Casati et al. and O'Callaghan, respectively. Although he echoes them in dismissing the Property View, Scruton rejects their physicalist approach and, instead, conceives of sounds as existing in an analogous manner to secondary qualities. Specifically, his argument is that sounds should be classed as secondary objects instead, seeing as they are inextricably dependent upon sensory perception. For instance, he states that just as rainbows are *visibilia* (i.e. secondary objects of visual perception), sounds are *audibilia* (i.e. secondary objects of auditory perception).[60]

---

[57] Casati, R., Di Bona, E., Dokic, J., 'The Ockhamisation of the Event Sources of Sound', *Analysis*, 73, 3, 2013, p. 464.
[58] Ibid., p. 463.
[59] Leddington, p. 625.
[60] Scruton, 2009, p. 24.

Yet, even though it is in virtue of vibratory motion in the medium that we perceive sounds as *audibilia*, Scruton argues that we hear them as events, such that "all that we hear when we hear sounds are the secondary properties of sound events".[61] He thus rejects the physicalist approach to the nature of sound shared by the other Event theorists. In his view, they seem rather more concerned with identifying the location of the sound event in space so that the distal theory has the upper hand over the medial account.

Another key difference between their event theories concerns the roles of the sound source and the medium. As seen earlier, Casati et al. and O'Callaghan share the view that sounds do not travel through the medium, but should instead be construed either as identical to the event source (according to Casati et al.) or as part of that event (as argued by O'Callaghan). Conversely, Scruton once again holds a different view on the nature of sound *qua* event. He argues that sounds should be conceived as "independently existing events, located in a region of space... [that] can be detached completely from their source".[62] As such, he claims that they should be classed as 'pure' events, in that they do not happen to the source, or the medium, or the auditory system of the observer. As he understands it, the sound event is pure in the sense that it does not happen to anything.

Although Scruton's category of pure events may seem obscure to some, he indicates certain principles under which they may be understood. Specifically, he describes pure events as being ordered by what he calls principles of aggregation and disaggregation, "whereby events can be decomposed into smaller events, joined up into large ones, and accorded precise relations in time, all by reference to the events themselves, and independently of any physical objects involving in producing them".[63] This, for Scruton, is precisely what happens in the case of a distinctively musical experience. Specifically, he argues that this enables us to hear musical tones as completely detached from their sources, as if they belonged to a different order. To hear sounds in this way, he states, is to perceive them *acousmatically*, rather than acoustically.[64] In particular, Scruton sees the independence of sounds from sources as an advantage of his view, in that it does not become a target for the objections raised against the physicalist approach taken by Casati et al. and O'Callaghan. However, in spite of divergences between their event theories, the premise that affords them a unified front against the Property View is that temporality is an essential aspect of the nature of sound.

[61] Scruton, 1997, p. 39.
[62] Scruton, 2009, pp. 21-22.
[63] Ibid., p. 28.
[64] Ibid., p. 30. Scruton's acousmatic thesis will be considered at greater length in Chapter Five (Section 2).

As seen earlier, the claim that sounds are distinctively temporal entities conforms to the common intuition that sounds last. Yet, in his most recent challenge to the Event View, Kulvicki argues that the claim that sounds have duration is based on "a kind of perceptual seeming" whereby we infer that sounds last only because they *seem* to last over a period of time.[65] This foregone conclusion is, for Kulvicki, based on the assumption that the events we hear take time, and that this coincides with the duration of the respective auditory experience. But these claims, he suggests, do not entail that sounds themselves have durations; rather, they are "altogether neutral with respect to whether sounds, or merely the episodes in which sounds can be heard, have durations".[66] However, this would seem to challenge the ontology of events itself, in that the same claim could equally be made of visual experience – i.e. that the events we normally witness may not have durations, either.

As can be seen, although the contributions from Event theorists have reinvigorated the philosophy of sound, their views have been the recipient of substantial criticism across the sound debate. Similarly, several objections have been raised against the Wave View. However, beyond the assessment of the merits and pitfalls of individual accounts is a more important problem that needs addressing – namely, the claim that the Wave View and the Event View are irreconcilable. It is the task of the next section to point out why these views have been portrayed as such, and how they compound the 'problem of sound'.

4. Statement of the problem

Having considered a range of non-property theories of sound under the Wave and Event views, the aim of this final section is to provide a characterisation of what I will refer to as the 'problem of sound'. I am going to argue that this problem arises from the very nature of the debate. Specifically, it concerns the ways in which the ontology of sound has been overshadowed by the question of sound location and that of multiple instantiation. In order to understand why the two views above have come to be portrayed as irreconcilable, I will indicate why theorists have tended to favour exclusively one view to the detriment of the other. Furthermore, I will consider another way in which the sound debate has been framed, whereby the Event View is taken to be incompatible with the view that sounds are objects (i.e. the 'Object View'). Hence, I will describe the problem of sound as having two horns, in that it comprises both the Wave-Event and the Object-Event dichotomies.

---

[65] Kulvicki, 2014, p. 210.
[66] Ibid., p. 210.

First, as indicated in Section 3, the Wave and Event views have been treated as irreconcilable. Although those who defend the Event View do not reject the Wave View for exactly the same reasons, there seem to be certain broadly shared concerns within that camp. From the Event theorist's perspective, one of the difficulties with the Wave View is that we do not experience sounds as pressure waves. Hence, to reduce them to their acoustic base is either to mischaracterise the sounds that we hear or at least offer an impoverished account of the objects of audition. Furthermore, some have alluded to the fact that we may be able to experience pressure waves but not hear any sound at all; as per the earlier example, a deaf person standing next to loudspeakers may experience the pressure waves in the air around them and still not hear a single sound. Beyond these counterintuitive consequences, there are also those specific concerns regarding the spatial phenomenology of auditory experience as well as questions around the ontology of waves themselves which were considered in Section 3.1.

Similarly, the adequacy of the Event View has also been challenged, albeit most frequently by Property theorists instead of advocates of the Wave View. One of the main objections against the Event View is that the notional distinctiveness of sound's temporality is either questionable or at least not something that can be inferred from certain aspects of the nature of audition. Furthermore, the task of identifying what event in the causal chain of sound production and auditory perception is the right candidate for what we hear, and how such event is related to space, may prove too intractable to some. Over and above that, another difficulty besetting event theories concerns their ability to provide a satisfactory account of what sounds are independently of waves. Specifically, one might object that they do not offer much beyond just proposing what ontological category best captures their distinctively temporal nature. Although this may primarily be seen as an issue for the ontology of events instead, it could be argued that describing sounds as such does not tell us anything of explanatory significance about the nature of sound altogether.

However, contrary to claim that the Wave and Event views are irreconcilable, I want to argue that this first horn of the problem of sound stems from the fact that the question of the ontology of sound has been conflated with the problem of sound location. As seen earlier, the issue of locational inconsistency seems to be at the heart of the dispute between those who support either the Wave or the Event View. Yet, this conflict is fundamentally one between medial and distal theorists, instead. Hence, Wave and Event theorists – with the exception of Nudds and Scruton – seem to have portrayed the two views as incompatible because they have treated them as merged with theories of sound location.

But the way of presenting the sound debate described thus far is not unrivalled. Instead of taking it to arise from the rivalry between property and non-property theories, Dokic proposes that the debate is best understood in terms of the dichotomy between the Event View and the 'Object View' – which is the second horn of the problem of sound, as I am describing it. In 'Two Ontologies of Sound', he states that sounds are either unrepeatable events or repeatable objects. The former view, favoured by Dokic, conforms to the claim that sounds are distinctively temporal entities. As such, he adds, sounds "do not exist as wholes at any proper time of their duration",[67] so they must have temporal parts.[68] Most importantly, as per the nature of events, sounds would be unrepeatable.

Conversely, on the alternative conception, sounds are taken to be particular kinds of objects that can exist as wholes in more than one place and time and thus have multiple occurrences. On this iteration of the Object View, sounds are instances of auditory types that may or may not be instantiated; hence, they are taken to be repeatable entities. Despite pitching the two conceptions as incompatible, Dokic acknowledges that they share two principles, namely: that sounds are particulars, rather than universals; and that sounds "are not ordinary space-occupying objects [in that] they lack spatial parts".[69] But he maintains that the real bone of contention concerns the question of the repeatability of sounds.

Dokic's presentation of the sound debate raises further concerns when considered in parallel with the Wave-Event polarisation. Briefly, if sounds are repeatable objects, then questions around their relationship to their sources as well as the pressure waves still remain. Similarly, if one takes sounds to be unrepeatable events, explaining the common intuition that sounds may be re-encountered would prove challenging. Hence, Dokic's dichotomy also raises difficulties for accounting for the phenomenology of auditory experience and the relationship between sounds and their acoustic base. Most importantly, however, under Dokic's proposal the Event View and the Object View are irreconcilable. Yet, in Chapter Two I will argue that it is possible to conceive of sounds both as events and as objects, whilst subsequently addressing the question of their multiple instantiation at different times over the ensuing chapters. Specifically, I will show that the case of chords is not only illustrative of that possibility but also a particularly illuminating one.

---

[67] Dokic, J., 'Two Ontologies of Sound', *The Monist*, 90, 3, 2007, p. 391.

[68] The claim that sounds have temporal parts will be considered in Chapter Two (Section 4), as part of my account of sound persistence. For the purposes of this section, I am adopting the definition of temporal parts offered by Katherine Hawley in her characterisation of perdurance theory, under which objects are spread out not only in space but also in time. As she describes it, objects have "different spatial parts in different parts of the spatial region which they occupy, and they have different temporal parts in different parts of the temporal interval they occupy". Hawley, K., *How Things Persist*, Oxford: Clarendon Press, 2001, p. 10.

[69] Dokic, pp. 391-392.

As can be seen, it is not difficult to grasp why the philosophy of sound has recently come to be characterised as an arena of irreconcilable views. On the one hand, there is an apparent rivalry between the Wave View and the Event View and, on the other, a supposed opposition between the Event View and the Object View. It is precisely in the confluence of these dichotomies that I take the problem of sound to consist. Specifically, I take it to be fundamentally of an intellectual character, as arising from the way the debate has been framed, rather than a metaphysical problem as such. By emphasising certain specificities of their views by way of contrast to other theories, Wave and Event theorists seem to have lost sight of the real opposition at stake – namely, that between property and non-property theories of sound.[70] Similarly, casting the debate in terms of the Object-Event dichotomy only seems to shift the rivalry under the criterion of the repeatability of sound. Yet, alternative views such as Scruton's indicate that this is not an intractable problem.[71] Finally, the fact that most of the objections directed against the Event View come from Property theorists rather than Wave View advocates is itself an indicator that a potential Wave-Event reconciliation is not only possible but, in my view, desirable. It is, therefore, under this ambition that I shall propose my account of the ontology of sound in the next chapter.

[70] This is also the view expressed by Roberts, an advocate of the Property View himself. Roberts, p. 337.
[71] As indicated in Section 3.2, Scruton takes sounds to be secondary objects as well as pure events.

Chapter Two: Towards a Sonic Reconciliation

As seen in the previous chapter, the philosophy of sound has come to be portrayed as an arena of rivalling accounts. This, I argued, has created a problem that is essentially of an intellectual order, rather than a metaphysical one. While there are important differences between accounts under the same view, what some Wave and Event theorists seem to neglect is that what they are fundamentally opposed to is the Property View. Instead, those divergences seem to have prevented non-Property theorists from proposing a 'third way' under which their views may be reconciled. In this chapter, I aim to defend an alternative conception that would enable such reconciliation, in an attempt at addressing the problem of sound as outlined in Chapter One. Specifically, I shall provide an account of the ontology of sound that reconciles the fundamentals of the Wave View with those of the Event View, whilst also showing that the latter is compatible with the Object View.

1. A sonic reconciliation

In Chapter One, I indicated that the sound debate should not be construed in terms of several competing accounts but chiefly as a dichotomy between property and non-property views. If sound is not a property – as I have argued – then it must be some kind of particular. Yet, as seen in that chapter, different claims have been made as to how the nature of sound is best understood. In this section, I aim to outline how my proposal will constitute a 'third way' between the Wave and Event views, in an attempt at reconciling their most fundamental claims under a unified account.

The cornerstone of my proposal in this chapter is the thesis of 'sonic reconciliation', which is twofold in nature. First, I will maintain that my account is able to address the first horn of the problem of sound by accommodating the core intuitions underscoring both the Wave and the Event View. My argument here is mostly indebted to the accounts put forward by O'Callaghan and Nudds as outlined in the previous chapter. But, contrary to their stances, my main motivation is to point out that the Wave-Event dichotomy is false. The second component of my thesis of sonic reconciliation concerns the second horn of the problem of sound, i.e. the claim that the Object and Event views are incompatible. Contrary to Dokic's approach, I will argue that those views may also be reconciled under my proposal – which therefore consists in a double conciliatory effort.

The present chapter is thus intended not only as a contribution to the sound debate but also as an attempt at a solution to the problem of sound arising from it. Unlike the way in which the debate has been framed, my proposal is that (a) it is possible to conceive of sound *both* in terms of the Wave View *and* the Event View, just as (b) it is possible to conceive of 'sound streams' *both* as auditory objects *and* as persisting events. In defending these claims, I will aim to meet the following desiderata: (i) if (a), then an explanation of the relationship between sounds *qua* events and the pressure waves must be given; (ii) if (b), then such an account must explain how treating sounds both as objects and events is compatible with our phenomenology of auditory experience; and (iii) if (a) and (b), then we must account for the semantic implications for the way we speak of sounds.

Since sounds are not properties, I am going to account for their nature as particulars possessing distinctive identity, individuation and persistence conditions. In order to uphold (a), I will characterise the identity conditions of sound in terms of a grounding relation. Specifically, I shall argue that sounds are events that are grounded by pressure waves. After attempting to meet (i) in Section 2, I will seek to defend (b) under my account of sound individuation in Section 3. Contrary to Dokic's Object-Event dichotomy, I will show how my thesis is able to meet (ii) by distinguishing between sounds *qua* events that are grounded by pressure waves and the stream that results from the auditory grouping of those sounds by the auditory system. It is on that basis that I aim to satisfy (iii) in the last section; I shall argue that what we ordinarily refer to as 'sounds' is best understood as a stream of sounds, instead. I will defend this view by treating sounds as the temporal parts of the stream – such that it is the *stream* that I take to be the proper object of auditory experience.

In accounting for the ontology of sound, I will take the following claims to be uncontroversial for both Wave and Event theorists: (1) sounds are not properties but particulars, instead; (2) sounds are caused by an initial disturbance that constitutes the sound-producing event; (3) sound waves are caused by the vibratory motion of the objects involved in the sound-producing event; (4) the presence of an elastic medium is a necessary condition for wave propagation and sound transmission; and (5) the presence of an observer is a necessary condition for an auditory appearance to occur. Having addressed (1) in the previous chapter, I will propose that (2) and (3) enable us to understand the relationship between sound events and their sources. Furthermore, I will maintain that (4) and (5) contain the necessary conditions for sounds not only to be experienced as auditory objects but to *exist*. The conjunction of these propositions, I believe, provides the basis for an understanding of sound's identity, individuation and persistence conditions.

2. Identity

The starting point of my proposal concerns the identity conditions of sound. As seen in Chapter One, the views offered by non-Property theorists seem to preclude the possibility that we may understand the nature of sound both in terms of waves and events. Conversely, I am going to argue that sound's identity is best understood via an integrated approach which reconciles the fundamentals of those views – which is my attempt at addressing the first horn of the problem of sound. Specifically, I am going to argue that sounds are events that are grounded by pressure waves. My account in this section will first comprise certain physical aspects of sound production/transmission. Subsequently, I will consider the phenomenological dimension of our acquaintance with sounds' properties.

The first part of my argument concerns the way in which sounds may be understood to be part of the fabric of reality, as it were. To that end, it seems fundamental to consider the physical conditions under which sounds come into being. Briefly, sounds are caused by an initial disturbance that sets the surrounding medium in vibratory motion, which is detected by an observer. Another way of describing this physical phenomenon has been expressed above in (2) and (3). Perhaps unsurprisingly, both claims are aligned with the standard view of sound received from acoustics. Most importantly, their conjunction is also compatible with the backbone of the views espoused by Wave and Event theorists. Yet, if we consider them disjunctively, (2) and (3) may be taken as two different conceptions of the nature of sound, in support of either the Event or the Wave View, respectively.

Contrary to the disjunctive approach, I want to argue that it is counterintuitive to think of sounds *either* as waves *or* as events. Put simply, just as sounds have event-like characteristics, the same can be said of pressure waves. What we describe as a 'wave' is something that happens within the medium; more technically, it is a pattern of pressure variation that occurs due to particle displacement. The disturbance that is propagated is fundamentally a change to which the medium is subject, in that it would otherwise remain in a state of equilibrium – to which it will eventually return. As such, the wave has duration; it has a beginning and an end in time. Over and above that, the wave is constituted by the medium, since the compression phase consists in a tighter agglomeration of particles, whereas rarefaction means that the particles are spread further away from one another. The multiple micro-events of particle movement can thus be seen as sub-events occurring as part of a larger event that we refer to as 'waves'. On the basis of these claims, it seems reasonable to suggest that sound waves have an event-like nature.

At first glance, these remarks about the nature of waves would seem to point to a veiled defence of the Wave View. This may be seen to be the case especially if one considers that the properties that are commonly associated with sounds – for instance, pitch and loudness – are taken to be the audible counterparts of the waves' frequency and amplitude/intensity, respectively. Yet, such an interpretation would be misguided for two main reasons, the first of which concerns the issue of property instantiation.

By identifying sounds with waves, the Wave theorist is assuming that sounds must have whatever properties the respective waves have. However, unlike the apparently neat correspondence between acoustic and audible properties, the case of timbre (i.e. the distinctive quality of the sound heard) raises a particular difficulty. As Stephen Handel notes, at present "no known acoustic invariants can be said to underlie timbre"; for instance, in musical perception "the cues that determine timbre quality are interdependent ... [and] depend on context: the duration, intensity, and frequency of the notes, the set of comparison sounds, the task, and the experience of the subjects all determine the outcomes".[1] For that reason, the audible property of timbre cannot be solely understood in terms of the waves' acoustic properties. Hence, under the principle of the indiscernibility of identicals, one would have difficulty in treating sounds as identical to waves.

Secondly, the Wave View is fraught with a certain degree of obscurity concerning the perception of sounds' properties. In identifying sounds with waves, the Wave theorist is committed to the view that sounds pass through the medium, just as waves do. However, as O'Callaghan has pointed out, the problem is that this medial view entails that sounds have certain spatial boundaries as they move through the medium – in which case, he claims, we must be under some sort of illusion when we perceive sounds the way we do. In his words, we would "mistake an experience of the spatial boundaries of a sound for an experience of the duration and temporal boundaries of that sound".[2] But the Wave theorist does not have to be committed to the view that sounds are identical or reducible to waves. While Sorensen endorses it, Nudds proposes an alternative, more nuanced account of sounds in terms of their multiple instantiation in time.

As seen in Chapter One, Nudds takes sounds to be patterns or structures of frequency components that are instantiated by the waves. But he also stresses that, "since people at different places, who hear the same sound are not, and need not be, affected by the same instantiation of frequency components, we cannot identify particular sounds with

[1] Handel, S., 'Timbre Perception and Auditory Object Identification', In Moore, B. C. (ed.), *Hearing*, San Diego, CA: Academic Press, 1995, p. 433.
[2] O'Callaghan, 2007, p. 44.

instances of a pattern or structural type".[3] Consequently, he argues that sounds are better understood as abstract individuals – a view which is not without difficulties, as noted in the previous chapter. Nevertheless, what this variant of the Wave View points to is the possibility that (2) and (3) above may both be taken as saying something about the nature of sound without committing to the view that sounds are waves.

That said, sounds and waves seem to be connected in virtue of some of the properties they bear. In spite of the case of timbre, there is an empirical basis for defending the correlation between audible and acoustic properties, such as that between pitch and frequency. Both O'Callaghan and Nudds seem to treat this correlation as empirical fact: while O'Callaghan maintains that "the dependence of audible qualities upon physical characteristics needs to be preserved",[4] Nudds suggests that "our auditory experience is of, or represents, the sources of sounds and their properties as well as sounds and their properties".[5] However, sounds and waves do not seem to share the *same* kind of properties, since audible properties are phenomenal qualities and acoustic properties are physical quantities. Furthermore, they do not share spatial properties, either.[6] Hence, they are not identical in logical terms, and neither can they be said to be phenomenologically the same in that we do not experience sounds as waves.

If sounds and waves are not the same entities, then it would seem that the Event View has the upper hand over the Wave View. Yet, Nudds rejects the former, claiming that it presents us with a misguided direction of explanation in accounting for our epistemic access to sound sources. If the Event View is correct, he states, then "it would be reasonable to suppose that an account of auditory perception should begin with an explanation of how we perceive sound events, and that an explanation of our perception of the sources of sounds should be in terms of the perception of those sound events".[7] The problem is that, according to Nudds, the converse is precisely the case. Specifically, his rejection of the Event View rests upon two main assumptions, namely: that the putative sound events must be distinguished from the sound-producing events; and that it is only in terms of our experience of the sources of sounds that we can explain how we perceive sounds themselves. Hence, on this view, the phenomenology of the auditory experience of sound sources takes precedence over the ontology of sound *per se*.

---

[3] Nudds, 2009, p. 76.
[4] O'Callaghan, 2007, p. 76.
[5] Nudds, 2009, p. 70.
[6] It is worth remembering that, as seen in the previous chapter, Nudds maintains that sounds do not have any spatial properties at all.
[7] Nudds, 2013, pp. 281-282.

The phenomenological dimension of auditory perception is at the centre of the second part of my account in this section. As common ground assertions, (2) and (3) also indicate two generic facts about how sounds come to be: they are preceded by an event that would seem to explain their inception, and the relationship between sound and that event seems to be a causal one. From the conjunction of (2) and (3), I take the sound-producing event to refer to the initial disturbance, which then causes the vibratory motion of the medium that we refer to as waves. Similarly, this causal process can also be said to underlie sound occurrences: their onset is brought about as a result of vibratory motion happening to objects (i.e. the initial disturbance) as well as within the medium (i.e. the vibratory event). Sound is thus an event that is caused by the sound-producing event.

This claim bears consequences not only for the present section but also for my account of sound individuation and persistence in subsequent sections. One of them concerns a key difficulty for Event theorists, namely, that of specifying which event is the right candidate to be classed as the sound event. As seen in Chapter One, Event theorists are not in agreement: while O'Callaghan argues that the sound event is a proper part of the sound-producing event (i.e. the Parthood View), Casati et al. maintain that sound is the sound-producing event itself (i.e. the Identity View). By contrast, although Nudds accepts the language of events as applied to what happens to the sound source, he correctly distinguishes between the initial disturbance and the vibratory event that ensues. Nudds claims that ordinary sound-producing events "cause vibratory events (or perhaps have vibratory events as parts), but are not themselves vibratory events".[8] Although the Identity View is dismissed under this claim, the same is not the case for the Parthood View. Yet, since the initial disturbance and the vibratory event are changes that affect the object and the medium respectively, Nudds indicates that they must be accounted for individually.

Although I do not take Nudds's objection to undermine the Event View *as such*, I agree that the sound-producing event and the vibratory event must be accounted for individually. Hence, as indicated earlier, I take the vibratory event to refer to an event that happens in the medium: it is a change from a state of equilibrium to a certain pattern of vibratory motion that affects the medium, caused by the initial disturbance at the source. Hence, we should instead treat the *waves* as the vibratory event. This conforms to (3), in that sound waves are caused by the vibratory motion of the objects involved in the sound-producing event. In addition, it also reflects the event-like characteristics of waves – which, I believe, can illuminate our understanding of the nature of the sound-wave relation.

---

[8] Nudds, 2014, pp. 467-468.

Yet, as seen earlier, it is sounds that appear to us in auditory experience and not waves, since we do not experience them as such. This obscurity of the sound-wave relation is, perhaps, the main difficulty in the search for the right candidate for sound. For instance, if we take sounds to be the vibratory events instead of waves, the Event View then becomes vulnerable to challenges from both Property theorists and Wave View advocates: on the one hand, it could be argued that sounds are instead properties of the medium; on the other hand, since it is vibratory motion that is at stake, this may lead one to infer that sounds are reducible to it. Given that neither of these outcomes is desirable for the Event View, I am proposing that it is the waves that should be classed as the vibratory event.

It is in light of that claim that I want to account for the sound-wave relation in a way that conforms to (2) and (3) but without entailing identity between sounds and waves. A few observable facts seem relevant; for instance, sounds and waves are caused by the same initial disturbance, and some of their properties can be correlated. But, while variations in the waves' properties affect the sounds that we hear, any perceived changes to audible qualities do not entail that waves will also vary accordingly as a result – which indicates an asymmetric relation. Furthermore, it seems that the nature of the relationship between them is non-causal. As Paul Audi notes, causation would seemingly require that facts about the causal relata are "wholly distinct", which is something that cannot be said of those concerning the nature of sounds and that of waves.[9] Hence, if sounds are not identical with the waves, there must be another way to explain the empirically established parallels between them. Given the event-like nature of waves – as outlined above – and the apparent non-causal nature of the sound-wave relation, I want to argue that the latter is best understood in terms of the metaphysics of grounding.

Under my proposal, sounds are events that are grounded by pressure waves. In making this claim, I am drawing on two fundamental aspects of the metaphysics of grounding in order to account for the nature of the sound-wave relation. First, echoing Audi's view, I take the grounding relation to be, by definition, a non-causal relation of determination. To use one of his examples, a ball rolls down an inclined plane in virtue of being spherical, but being spherical does not cause it to have such a disposition.[10] Similarly, I want to argue that the sound-wave relation is such the nature of sound may be accounted for in virtue of certain facts about the nature of waves; but since waves do not cause sounds, I do not take this type of determination to be given in terms of causal explanations.

[9] Audi, P., 'Grounding: Toward a Theory of the "In-virtue-of" Relation', *The Journal of Philosophy*, 109, 12, 2012, p. 687.
[10] Ibid., pp. 687-688.

Secondly, while there is some debate as to what the right candidates for grounding relata are,[11] I will echo a growing consensus that grounding holds between facts.[12] On this view, David Liggins offers one of the possible formulations of the grounding relation: "the fact that *a* bears [relation] R to *b* might be grounded by the fact that *a* has the property of Fness".[13] Audi gives a similar iteration when he states it as follows: "*a*'s being *F* grounds *a*'s being G" (although his iteration involves facts about the same entity).[14] In the specific case of the sound-wave relation, I hold the view that it is precisely the event-like nature of waves that grounds the fact that sounds are events, so I am adapting the two previous formulations to suggest that *a*'s being *F* grounds *b*'s being *F* – or, put differently, the fact that *a* is *F* grounds the fact that *b* is *F* – where *a* refers to 'waves', *b* refers to 'sounds' and *F* to 'being event-like'. Under this iteration, both sounds and waves are events, thus belonging to the same category. That said, although I take facts about sounds and facts about waves to be the relata, my formulation is not intended as a commitment to a broader view that grounding relata must belong to the same ontological category.[15]

As stated earlier, the sound-wave relation is asymmetric. This is easily accommodated by my account of sounds as grounded by waves, in that one of the key principles of a grounding relation is asymmetry.[16] For instance, the fact that waves have a certain physical property (i.e. a certain frequency in Hertz) might ground the fact that sounds are heard as having a certain pitch, but the converse is obviously not the case. This gives waves ontological priority over sounds, thus making it compatible with a physicalist reading of the Wave View. Yet, if facts about sounds are grounded by facts about waves – as opposed to being reducible to them – then facts about the waves' properties may be necessary but insufficient to explain certain facts about sounds' properties. As seen earlier, this is precisely the case with timbre. Hence, my proposal would conform to the intuition underscoring the Event View that sounds' properties are irreducible to the waves' properties. The notion of grounding is therefore well suited to describing the asymmetry of the sound-wave relation whilst accounting for the correlation of their distinctive properties.

---

[11] Sylvia Barnett argues that this question has been largely neglected in the literature. Barnett, S. C., *What Sort of Entities Does Grounding Relate?* (PhD thesis), Manchester: University of Manchester, 2018, p. 8.

[12] Some of the scholars who have recently supported this view include Paul Audi and David Liggins (with respective references in the subsequent footnotes) as well as Sylvia Barnett (as per the previous footnote).

[13] Liggins, D., 'Grounding and the Indispensability Argument', *Synthese*, 193, 2016, p. 532.

[14] Audi, p. 686.

[15] As noted by Michael Clark and David Liggins, under a flat theory of grounding the relata must belong to the same ontological category, whereas dimensioned theories allow for them to belong to different categories. Clark, M., Liggins, D., 'Recent Work on Grounding', *Analysis*, 72, 4, 2012, p. 818.

[16] Trogdon, K., 'An Introduction to Grounding', In Hoeltje, M., Schnieder, B., Steinberg, A. (eds.), *Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, Response-dependence*, Munich: Philosophia, 2013, p. 106.

In addition to asymmetry, the sound-wave relation is equally conformable to the other principles that are generally taken to qualify the notion of grounding, namely, transitivity and irreflexivity.[17] While the latter is most conspicuous, since sounds do not ground themselves, the former requires a more in-depth understanding of the nature of acoustic properties. One way to illustrate how the sound-wave relation is transitive is by considering the relationship between those properties and the sounds that we hear. To use a music-related example, just as facts about waves in simple harmonic motion ground facts about the properties of pure tones, the former facts will also ground facts about complex waves, which in turn ground facts about the complex tones that we hear as musical notes.[18]

The sound-wave grounding relation is essential to my thesis of sonic reconciliation. As outlined in Section 1, one of the aims of this chapter is to defend (a), i.e. that it is possible to conceive of sounds in terms of the Wave and the Event View. Hence, my proposal is that sounds are *events* that are grounded by waves. This approach, I believe, has three key advantages. First, it addresses the theoretical limitations of identifying sounds with waves. If facts about sounds are grounded by facts about waves, this enables an understanding of their nature that is not solely based on the mechanics of vibratory motion but also on the phenomenology of auditory perception. Second, it allows for sounds to have distinctive properties which are acquainted with differently from those of their wave counterparts. Third, by conceiving of sounds as events grounded by waves, my proposal provides an alternative to the Identity and Parthood views, by avoiding the pitfalls of the former as well as the obscurity of the latter.

That said, it could be argued that the analysis I have presented thus far does not address the nuances of the accounts outlined in Chapter One. Granted, my argument has focused on the dialectic between the claims made by Nudds and O'Callaghan, in particular, to the detriment of alternative views. However, this is because of certain implications of my assumption that the presence of an elastic medium is a necessary condition for wave propagation, which was expressed earlier under (4). Specifically, this renders Casati and Dokic's event physicalist view a less credible theory of sound. For instance, if we consider (4) in light of the sound-wave grounding relation, it seems clear that the medium is indeed a necessary condition not only for wave propagation but for the existence of sound itself. If facts about sounds are grounded by facts about waves, and these can only occur in the presence of a medium, Casati and Dokic's main thesis that sounds are physical entities that exist independently of the presence of a medium becomes untenable.

[17]Ibid., p. 106.

[18] The acoustic nature of pure and complex tones will be considered at greater length in the next chapter.

A similar casualty occurs on the Wave View front. By identifying sounds with pressure waves as well as maintaining that silence is the absence of waves, Sorensen's account falls prey to some counterintuitive consequences. For instance, when considering the case of destructive interference (which occurs when sound waves are out of phase such that they cancel one another out), Sorensen argues that "if the destructive interference is perfect, then the hearer will report hearing nothing. Or he might take himself to be hearing silence. But he is actually hearing two sounds that sum to zero".[19] Yet, this conclusion is not only acoustically misguided but also phenomenologically counterintuitive. First, as noted by Speaks, when the destructive interference is perfect "the cancellation means that at that moment, the medium is not vibrating".[20] If we once again consider (4) in view of my claim that sounds are events that are grounded by waves, and seeing as these consist in vibratory motion in the medium, if the latter is not vibrating then there can be no sound. Furthermore, the claim that we *hear* two sounds that sum to zero when no vibration occurs seems to be a strange consequence of Sorensen's account, since no sound at all is heard.

Nevertheless, another variant of the Event View previously considered should not be so easily dismissed. In defending his acousmatic thesis, Scruton rejects any attempt at conceiving of sound in terms of what happens to the medium, or even at the source. Although his category of pure events certainly raises questions, in characterising sound as both events and secondary objects he is, to a certain extent, advancing the second aim of the present chapter. This, I believe, is better accounted for under an analysis of the individuation of sounds, as shall be considered in the forthcoming section.

3. Individuation

Under the proposal outlined in Section 2, sound's existence is ontologically dependent upon that of the waves, but sounds' properties are not reducible to waves' properties in that they are not identical but stand in a grounding relation. Yet, in order to provide a satisfactory account of the nature of sound, we must consider how sounds are detected by the auditory system and how we experience them. Having previously addressed (a), in this section I aim to substantiate (b), i.e. that we can conceive of sounds both as objects and as events – thus addressing the second horn of the problem of sound. To that end, I shall offer an analysis of the individuation conditions of sound which, I believe, meets the desideratum expressed in (ii) and serves as the basis for meeting (iii) in Section 4.

[19] Sorensen, pp. 141-142.
[20] Speaks, p. 269.

In order to give an account of how sounds may be individuated, it is essential to consider the phenomenology of auditory perception. In the previous section, I argued that sounds are events that are grounded by pressure waves. But, as indicated in (5), the presence of some sort of observer is necessary for sound to exist. Specifically, it is in auditory perception that sounds can be said to have the properties that they do, which are different from the waves' properties albeit grounded by them. Briefly, as the properties of the waves (i.e. the vibratory event) are detected by the auditory system, the latter processes those vibrations and assigns certain audible properties to the sound event. Yet, the auditory system operates like a filter; its inner workings determine how sound events – and which of them – will be *experienced* as a percept.

In his account of the phenomenology of auditory experience, Nudds suggests that sounds are best understood as objects, rather than events. This being the case, he argues that we should speak of them in those terms, for that is how we experience sounds. However, although the auditory system is presented with a range of stimuli surrounding the observer, not all of the corresponding sound events are experienced as such. As Nudds points out, auditory perception is fundamentally based on the grouping of the waves' acoustic properties, which is what determines the objects of audition – i.e. the "temporally extended sequences of sounds experienced as a group", as he describes them.[21] In particular, one of the key aspects of auditory grouping is that, depending on their properties, some sounds are given more prominence than others.

Nudds's account of the phenomenology of auditory perception is underpinned by the empirical research on auditory scene analysis proposed by Albert Bregman. On this view, the auditory system bases its grouping activity on multiple cues that operate on several different levels, which may be taken either jointly or individually. For instance, Bregman states that "sounds could stimulate the same frequency neurons, but still be segregated on the basis of other cues ... [such that] when one cue doesn't give useful information, others may still work".[22] Hence, under auditory scene analysis, auditory objects can be said to result from the auditory system's attempts to organise multiple sounds into 'sequences' – which may be either simultaneous or sequential – in terms of the most effective cues of which it can avail itself. Dovetailing with this view, it is on the basis of the grouping of simultaneous/sequential sounds that, I believe, sound individuation must be understood.

---

21 Nudds, 2010, p. 106.
22 Bregman, A. S., 'Progress in Understanding Auditory Scene Analysis', *Music Perception: an Interdisciplinary Journal*, 33, 1, 2015, p. 15.

If the necessary conditions identified in (4) and (5) are met, one may be inclined to suggest that all sound events processed by the auditory system are perceived as audible entities. However, given the function of auditory perception briefly described above, sounds may only contingently be classed as auditory objects. This is for two main reasons. First, in keeping with Bregman's auditory scene analysis, sound events must be detected under some sort of sequence, so that the information encoded in the waves' properties may be appropriately extracted by the auditory system. Second, as Nudds points out, "not all sequences of sounds are auditory objects ... because not all sequences of sounds are experienced as a group. When we experience a sequence as a group we experience the different parts of the sequence as belonging together".[23] Most importantly, what these reasons reveal is that there is a distinction between sound *qua* perceptual input and sound *qua* cognitive output, which forms the basis of my account in this section.

Specifically, there is an important distinction between the auditory perception of sound events and the auditory experience of sounds as we hear them – and this distinction is temporally marked by the process of grouping. If this is correct, we can speak of sounds *qua* perceptual input – i.e. consisting in sound events and their properties, prior to being grouped – and sounds *qua* cognitive output – i.e. the resultant stream, which results from the grouping of sequential/simultaneous sound events. To use a musical example, the sounds made by an orchestra in a concert hall are grouped in such a way by the auditory system that some sounds become more prominent than others, whilst any simultaneously occurring non-musical sounds around the observer may be dismissed and not be perceived as such. The resultant stream – which I have described as the cognitive output – is also an event that results from the grouping of some sound events to the detriment of others. Most importantly, it is the stream that I take to be the auditory object *per se*.

Hence, the first component of my account of sound individuation is that it is the streams that are the proper objects of auditory experience. Sounds are events, not objects, and they are grounded by pressure waves. But they are individuated in auditory perception as the sub-events of a larger group of sounds, which is organised by the auditory system based on the most effective available cues and presented as an auditory object. Consequently, auditory experience involves the perception of a sound stream, instead. Yet, streams are also events – for they comprise sounds as sub-events – and their duration is determined by the auditory system's grouping activity along with the related cognitive response. As such, sound streams may be conceived *both* as events *and* as auditory objects.

[23] Nudds, p. 106.

If my proposal is accepted, we should then question whether the Object-Event dichotomy is an appropriate way of framing the sound debate. As seen in Chapter One, Dokic's distinction between two ontologies of sound – namely, the Object and Event views – arguably eliminates the possibility that the content of auditory experience may be characterised under the concept of object as well as that of event. Yet, under the account I am proposing here, we are indeed able to conceive of sounds as *events* that are grounded by waves and presented as such to the auditory system, just as we are also able to conceive of sound streams that result from the process of auditory grouping as the *objects* of auditory experience. Hence, to maintain that the Event and the Object View are incompatible is a misguided approach, since it portrays these two ways of conceiving of sounds as irreconcilable when, in my view, the contrary is precisely the case.

This attempt at reconciling the Event and Object views is not, however, an entirely unprecedented approach. As stated in Section 2, Scruton's account stands out as an alternative event theory that challenges Dokic's dichotomy, in that he takes sounds to be both secondary objects and pure events. Sounds, he states, are "objects in their own right, bearers of properties" as well as "things that happen but which don't happen *to* anything".[24] To a certain extent, Scruton's view advances some of the elements of my account in this chapter. Specifically, he identifies three aspects of auditory experience that echo my analysis here: first, that we perceive sounds in virtue of vibratory motion – which I have characterised under the sound-wave grounding relation; second, that the resulting sound is an auditory object; and third, that this auditory object is heard as an event, and one that is irreducible to the sound events that constitute it or the waves' properties.[25]

Scruton also considers how the auditory system processes the sound events that we hear and suggests that streams of sounds are organised much in the same way as visual objects are. "In general," he states, "sequences of sounds are 'streamed' in our perception – each allocated to a temporal *Gestalt*, formed according to temporal analogues of the principles for *Gestalt* formation in vision".[26] Similarly, I take the import from Gestalt theory to be a useful tool for understanding how those groupings are undertaken on a phenomenological level – which will be particularly relevant for my analysis of chords *qua* auditory objects in Chapter Four.

---

[24] Scruton, 2009, p. 20.
[25] Scruton, 1997, p. 39.
[26] Scruton, 2009, p. 22. As described by D. W. Hamlyn, these principles stem from "a psychological theory which tried to explain various aspects of psychology in terms of structures (*Gestalten*), particularly in relation to the tendency of forms of perception... The Gestaltists emphasised 'wholes' and structures which could not be broken down into elements". Hamlyn, D. W., 'Gestalt theory', In Honderich, T. (ed.), *The Oxford Companion to Philosophy*, Oxford: OUP, 2005, p. 338.

However, there are some important differences between Scruton's position and my own. While I take sound streams to be the proper objects of auditory experience, Scruton takes sounds – i.e. sound events, in my terminology – to be the secondary objects, instead. This is particularly evident from his claim that "by treating sounds as secondary objects, we restore to them their true nature, as information-bearing events which are organised aurally".[27] This, however, would seem to point to my description of sounds *qua* perceptual input, whereas I take auditory objects to be the stream which is the cognitive output. Furthermore, given the centrality of the acousmatic thesis to his Event View, Scruton is forced to reject one of the key consequences of Bregman's auditory scene analysis, namely, that sounds are grouped in such a way as to inform us about the nature of their sources. Contrary to that view, he argues that "streaming involves attributing to sounds an identity distinct from any process in their source"[28] – which is in line with his pure-event thesis.

Scruton's view that sounds do not happen to the source or the medium seems plausible. Although my argument that sounds are grounded by waves makes my account closer to the physicalist view than his pure-event thesis, this is only to a certain extent: if facts about sounds are grounded by facts about the waves then it is the waves that perform the primary role of information-bearers. However, I take the view that sound events only contingently inform us about their sources. This is because they are not in any way reducible to waves, which are the *de facto* vehicles of source-tracking information. Consequently, on the one hand, it is indeed possible that sound streams may be heard as independent from their sources – which would conform to the acousmatic thesis; but, on the other hand, I do not take them to be Scruton's completely emancipated entities, seeing as the sub-events they comprise (i.e. the sound events) are grounded by waves.

That said, one might argue that my characterisation of streams as both objects and events does not address the question of repeatability raised by Dokic's dichotomy. As seen in Chapter One, he states that, if sounds are events, they must be unrepeatable, whereas if they are objects the same sounds may be re-encountered. However, there are at least two problems with this understanding of the Object View. First, there is nothing essential to the nature of an object which would require sound to have the property of multiple instantiation. Second, the claim that we are able to encounter the *same* sound is counterintuitive, in that we cannot assume that auditory objects are repeatable because they appear to be the same sound again. Rather, under my account we are not able to encounter the same sounds or even the same streams in that I take both to be events.

[27] Scruton, 2009, p. 22.
[28] Ibid., p. 22.

In view of what has been proposed above, sounds may be understood both as objects and as events. But the second component of my thesis of sonic reconciliation has further implications for the way we conceive and speak of sounds, as indicated under (iii). Specifically, if sounds are events, then they are fundamentally temporal entities. Hence, my account of the ontology of sound must address the question of how they persist as such.

4. Persistence

Treating sounds as objects as well as events poses two important questions: how they persist over time, and whether or not they survive qualitative change. In this final section, I aim to make some headway towards addressing these questions. In line with the account I have pursued thus far, I want to argue for a perdurantist view that sounds are the temporal parts of the stream. To substantiate this claim, I will draw on the notion of temporal window to account for the nature of sound *qua* perduring object. To that end, I will maintain that the stream is the entity that survives qualitative change over time, whereas sounds do not. It is on that basis that I will attempt to meet (iii), by arguing that the semantic move towards speaking of streams is not only well-motivated but necessary for my account of the ontology of sound to work.

As seen in Chapter One, the emphasis on sound's temporality is a distinctive aspect of the Event View. If sounds have duration, one should reasonably expect event theories to account for the way in which sounds last, i.e. how they persist through time. Albeit a nuanced debate, there are two main views that are widely held to predominate in the literature, namely, endurantism and perdurantism. Briefly, the key distinction between them is that endurantists take objects to be wholly present at each moment at which they exist, whereas perdurantists argue that objects persist through time by having distinct temporal parts at each moment at which they exist.[29] Yet, there seems to be a certain degree of reluctance on the part of Event theorists to engage with the metaphysics of persistence – which I take to be a weakness of the Event View. For instance, O'Callaghan's view on sound persistence is somewhat vague, which stems from his dismissal of the claim that sounds are objects on the basis that they are not experienced in the same way that ordinary objects are.[30]

[29] Eagle, A., 'Location and Perdurance', In Zimmerman, D. W. (ed.), *Oxford Studies in Metaphysics*, Vol. 5, Oxford: OUP, 2010, p. 53. For a definition of 'temporal parts', please refer to footnote no. 68 on page 35.
[30] O'Callaghan, 2007, p. 8.

One of the reasons why O'Callaghan rejects the Object View is that objects are usually taken to persist as enduring entities, rather than having different temporal parts at different times. Hence, if sounds are events, then the perdurantist view is – he tentatively suggests – best equipped to account for sounds' temporality. However, O'Callaghan does not commit to such a view, stating that he does not want his account of sound to be tied to the persistence debate. Rather, he solely chooses to stress that "sounds are in important respects different from ordinary objects in their ways of extending through time".[31] The problem is that, if sounds are time-taking particulars, then the Event View should be able to account for their persistence conditions; yet, no such account has thus far been provided by the Event theorists identified in Chapter One.

Conversely, I want to argue that my thesis of sonic reconciliation is able to address the question of sound persistence. Specifically, as proposed under (a), my argument that the Wave and Event views may be reconciled allows for the persistence of sounds to be considered in terms of the sound-wave grounding relation, which I proposed in Section 2. Furthermore, my proposal in (b) that sounds may be understood both as events and as objects enables us to conceive of their persistence conditions on the basis of the distinction between sounds *qua* perceptual input and sounds *qua* cognitive output, as outlined in Section 3. This, I believe, also enables my proposal in this section to show how the Object-Event reconciliation is compatible with a perdurantist account of sound.

As seen earlier, the fundamental claim that underpins perdurantism is that objects have temporal parts. Although there are different views as to how objects may be said to be extended through time, I do not intend to enter into a broader discussion on the nature of four-dimensionalism.[32] Rather, I want my account to be neutral on the specificities of the debate, such that sound streams may be said to extend through time on any account of the nature of a perduring entity. Most importantly, however, in order to address the question set out in the introduction to this section as to *how* sounds persist, I want to argue that the persistence conditions of the stream – rather than its component sounds – must be understood as dependent upon the auditory system's ability to group multiple sound events, both sequentially and simultaneously, over a certain period of time.

[31] Ibid., p. 27.
[32] For instance, Theodore Sider's defence of four-dimensionalism comes in the form of stage theory, as opposed to the 'worm view'. Under the former, the objects we normally refer to are instantaneous stages, whereas worm theorists take objects to be temporally extended space-time worms. Whilst acknowledging that four-dimensionalism is a broad church, Sider states that the unifying view is that objects are spatiotemporal entities that have a temporal part at every moment at which they exist. Sider, T., *Four-Dimensionalism: An Ontology of Persistence and Time*, Oxford: Clarendon Press, 2001, pp. 3, 191.

In Section 3, I argued that sounds are the sub-events that constitute the larger sequential/simultaneous group that I have referred to as the stream. In line with that claim, I take sounds to be the temporal parts of the auditory object that is the stream. Although this may seem counterintuitive in terms of the folk phenomenology of auditory perception, this need not be the case. Granted, one might argue that we do not refer to sounds as having temporal parts or as being spread out in time. Instead, we usually speak of them as though we are able to abstract the content of auditory experience in terms of a definite description (i.e. 'the sound of x'), or we emulate what they sound like. But that does not mean that sounds are not extended in time in the way that perduring objects are. Hence, it is possible to conceive of streams as having temporal parts, even if our linguistic practices are not aligned with such a view.

For instance, under my proposal, when we hear the sounds made by chiming bells, what is auditorily experienced is a stream of sounds. Sound events are grouped by the auditory system as temporal parts of the resultant auditory object that we refer to as 'the sound of bells'. This being the case, what that definite description picks out is the stream, and not the individual sound events as they are presented to the auditory system prior to being grouped. Hence, the semantic distinction between 'sounds' and 'stream' mirrors the ontological distinction I set out earlier between sound *qua* perceptual input and sound *qua* cognitive output, respectively. It should be noted, however, that this distinction is not aimed at rendering locutions that take the form of 'the sound of *x*' (e.g. bells) inadequate, insofar as they designate the auditory object that is the stream, instead. It is under these terms that, I believe, my account is able to satisfy the desideratum expressed in (iii).

Yet, in order to substantiate (b) further, it seems important to consider how my perdurantist view of streams is compatible with the phenomenology of auditory perception. Specifically, my argument here is indebted to Nick Young's notion of temporal window (or temporal field). Echoing other scholars, Young argues that perceptual experience is structured by a temporal window, seeing as "the world is not experienced instant by instant, but rather at intervals of about a second and a half or less". As he describes it, the temporal window is the temporal counterpart to the visual spatial field,[33] which is useful for understanding the perception of change and, consequently, that of perduring entities.[34] Although Young is not concerned with the nature of sound in particular, I take the notion of temporal window to be most helpful to understanding the case of auditory objects.

[33] Young, N., 'Hearing Objects and Events', *Philos*, 175, 2018, p. 2941.
[34] Ibid., p. 2945.

Hence, it is in light of the notion of temporal window – and in line with my account of sound individuation in Section 3 – that I take streams to be auditory objects that perdure. Although we may not hear the sounds that are its temporal parts as such, what we experience results from the sequential/simultaneous grouping of sounds as performed by the auditory system within a particular temporal window. This process – as would seem obvious – is itself an event. Thus, given the nature of the auditory process whereby they are 'forged', auditory objects are extended in time even if we may not speak of as them as such. Furthermore, the duration of the temporal window may be experienced by our consciously attending to an auditory object as it is fashioned – as we often do when listening to music.

The considerations above have consequences for the second question identified at the beginning of this section. Specifically, it concerns the nature of sound *vis-à-vis* the perception of change. One of the advantages of adopting Young's notion of temporal window is that it enables the perdurantist view of sound I have defended here to dovetail with my account of sound identity in Section 2. Given the sound-wave grounding relation, the audible properties of sounds – which are, under my proposal, ontologically dependent upon the waves' properties – may be understood as pertaining to different temporal parts of a stream. Most importantly, as indicated above, the stream is forged by the auditory system's grouping activity – a process which itself happens within a temporal window. This conforms to the view that auditory objects – i.e. the resultant streams – are events, due not only to their composition but also the very nature of the processes happening within the relevant temporal window. Hence, it is under this view that I will argue that streams survive qualitative changes to them, whereas sound events – as I have described them – do not.

Sound streams may be perceived as persisting despite changes to the way they sound. One of the ways in which one can conceive of the nature of the auditory experience of change is in terms of the contrast between 'survivalism' and 'non-survivalism'. This distinction has been identified by Cohen as part of his argument against the view that sounds should be treated differently from other sensible qualities due to their distinctively temporal features.[35] Under survivalism, sound is treated as a single entity temporally extended over a period of time, such that only one single sound is heard over a certain interval of time t_0-t_1. If there are any qualitative changes to the sound's properties (e.g. pitch) between t_0 and t_1, then the sound – as a single entity – survives those changes to its pitch. Yet, the survivalist view also allows for the existence of several temporal parts to the sound, which, in turn, do not survive the qualitative changes that occur to them.

[35] This claim underpins what Cohen terms the thesis of 'sound exceptionalism' – a view which he rejects, given that he is an advocate of the Property View. Cohen, p. 303.

By contrast, non-survivalists take several distinct sounds to be heard at t_0, t_1, t_2, etc., such that each of these bears distinctive auditory qualities. These individual sounds would then, over a period of time, constitute a temporally extended stream. Consequently, under non-survivalism, sounds do not survive qualitative changes, for they are the temporal parts of a stream which survives those changes to its temporal parts. In Cohen's words, what we hear in this case is "a number of distinct sounds over interval *T* ... [which] can be distinguished by differences in their pitch (or other auditory) qualities ... [and yet] there is ... the single, temporally extended stream of which these sounds are temporal parts".[36]

Even though the terminology used by Cohen in describing the non-survivalist view *prima facie* overlaps with the one I use in this chapter, I want to argue that it is survivalism that best captures my understanding of streams in light of the notion of temporal window. Although Cohen points to the temporally extended stream and its multiple temporal parts under his characterisation of non-survivalism, I believe that my account is more adequately construed in terms of survivalism because the stream is the auditory object that results from the grouping of the sound events within a given temporal window. In other words, under my proposal, what we hear is the stream as a single entity that is temporally extended, and not the multiple individual sounds over a given interval which Cohen associates with the non-survivalist view. However, if what I propose were to be made conformable to Cohen's characterisation of survivalism, one would have to take his reference to 'sound' as designating what I have described as the 'stream', instead.

It is, therefore, under this conception that I take the desideratum expressed in (iii) to be more satisfactorily met, namely, that we must account for the semantic implications of my thesis in view of the way we normally speak of sounds. Although it may be objected that my proposal is based on an unmotivated distinction between 'sound' and 'stream', such an objection is defused when we consider that both are events, but streams comprise sounds as their temporal parts. Furthermore, given that I take streams to be the proper objects of auditory experience, they must be what survives any perceptible qualitative changes, rather than sounds. Hence, and in view of the above, the semantic move towards speaking of streams is not only well-motivated but necessary for my account of the ontology of sound to work. If what persists is the temporally extended sequence of sounds as a single entity, then sounds – as I have conceived of them – cannot survive qualitative changes, since they are events grounded by pressure waves which, themselves, do not.

---

36 Ibid., p. 313.

The purpose of this chapter has been to defend my solution to the problem of sound outlined in Chapter One. Under my thesis of sonic reconciliation, I have argued that is possible to conceive of sounds both in terms of the Wave and the Event View, just as it is possible to conceive of sound streams both as auditory objects and as persisting events. Having considered the ontology of sounds at large in the first two chapters of this study, it is to the specific case of chords – the principal object of this research – that we now turn.

Chapter Three: Chords *qua* Sound Events

Chords are sounds. The preceding chapters are thus an essential preamble to accounting for their nature. Chords are not properties of musical instruments, but auditory objects in their own right. As such, they have their own identity and individuation conditions, as sounds do. In line with the thesis I defended in Chapter Two, my overall proposal in this chapter and the next is that chords must be understood both as events and as auditory objects. The focus of the present chapter, in particular, is on the view that chords are events. To conceive of them as such, I will argue, is to think of them as the coalescence of sound events that is grounded by the superposition of pressure waves. In order to substantiate this claim, I aim to provide a characterisation of the grounding base of chord events. I will subsequently identify the necessary conditions they must conform to: namely, periodicity, simultaneity and what I am going to refer to as 'mereological amalgamation'. The latter, in particular, will be considered in light of a mereology of events.

1. Defining 'chord'

All chords are sounds, but not all sounds are chords. Beyond this platitude, it is useful to start an analysis of the nature of chords by considering a couple of standard definitions. Chords have been described as "any simultaneous combination of notes, but usually of not fewer than three"[1] or, more simply, "two or more notes sounded together".[2] Yet, neither of these definitions seems satisfactory, and that is for two main reasons. First, in both of them chords are defined with reference to 'notes'. Second, there is no agreement as to whether two simultaneously occurring tones (known as a harmonic interval or a 'dyad') should be deemed 'chords'. Two preliminary observations are thus in order.

To define chords in terms of notes is to presuppose that chords must be understood within a framework of musical notation. Granted, it is the aim of the present study to undertake such task; yet, the purpose of the present chapter is to consider the nature of

---

1 Kennedy, J., Kennedy, M., Rutherford-Johnson, T. (eds.), 'Chord', *The Oxford Dictionary of Music*, OUP, 2013, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199578108.001.0001/acref-9780199578108-e-1857.

2 Fallows, P., 'Chord', In Latham, A. (ed.), *The Oxford Companion to Music*, OUP, 2011, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199579037.001.0001/acref-9780199579037-e-1376

chords *qua* sounds, which is what they primarily are. Hence, rather than notes, we should speak of tones sounding together.[3] This is because, apart from a musical context, a tone may be defined solely under perceptual and acoustic terms. As Harry Olson points out, tones may be described both as a "sound sensation having pitch or a sound wave capable of exciting an auditory sensation having pitch".[4]

Secondly, there is the question of whether dyads should be treated as chords. The divergence between the definitions above reflects a wider lack of consensus as to whether they should be. Indeed, the ambiguity underlying it can sometimes be found within the same work.[5] Yet, some may argue that, whether or not dyads are chords, they may be taken to 'imply' them.[6] Most importantly, since the same acoustic principles concerning triads, tetrads (three- and four-tone chords, respectively) and so forth are also applicable to dyads, that question is innocuous for the purposes of this chapter. However, for the sake of clarity, I shall henceforth take 'chord' to refer to a unit of at least three tones; and I will use the terms 'harmonic interval' and 'dyad' interchangeably, as is the case in the literature.[7]

Unlike the previous definitions, a more suitable way of describing chords is given in an entry from *The Oxford World Encyclopedia*. Here, 'chord' is defined as "the simultaneous occurrence of three or more musical tones of different pitch".[8] This seems to be a more appropriate starting point, in that it encapsulates the intuitions underlying the account I shall put forward in this chapter. By describing chords as 'occurrences', this definition indicates that chords are events. Similarly to the previous entries, it also specifies one of the necessary conditions for chord events to be individuated as such; namely, that of simultaneity. In addition, it describes chords in terms of 'tones' and 'pitch', which are particularly important for any account that incorporates a chord's relationship to its grounding base and the workings of auditory perception. Over and above that, it can be used to establish a distinction between dyads and chords. Hence, the object of the present study seems more neatly captured by that definition, instead.

---

3 Both dictionaries indicate in their respective entries for 'tone' that there is a terminological difference between British and American usage; while the latter favours 'tone', the former does 'note', instead. However, I do not take the two terms to be interchangeable.
4 Olson, H. F., *Music, Physics and Engineering*, New York: Dover, 1967, p. 36.
5 For instance, White states that chords consist of "the simultaneous sounding of two or more notes", but he shortly afterwards distinguishes between dyads and chords when considering their periodicity. White, H. E., *Physics and Music: the Science of Musical Sound*, Philadelphia: Saunders College, 1980, p. 194.
6 I will consider this under the notion of fundamental tracking, as will be detailed in Chapter Four (Section 1).
7 One example of this usage can be found in Barry Parker's work. Parker, B., *Good Vibrations: The Physics of Music*, Baltimore: John Hopkins University Press, 2009, p. 93.
8 'Chord', *The Oxford World Encyclopedia*, Philip's, 2014, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199546091.001.0001/acref-9780199546091-e-2402.

2. The grounding base

Dovetailing with my thesis of sonic reconciliation, I want to argue that chords are events that are grounded by the superposition of pressure waves. The main purpose of this section is thus to offer a characterisation of their grounding base. In line with the definition adopted above, I take chords to be sound events comprising three or more simultaneously occurring complex tones, which are grounded by the superposition of complex waves. Specifically, I will characterise the relationship between complex tones and waves with reference to their fundamental frequency and additional frequency components. Drawing from musical acoustics, this section is thus aimed at identifying the specificities of the grounding base of chord events.

Chords are events that are grounded by the superposition of pressure waves in the medium – which I will take to be air, by default. Perhaps unsurprisingly, the waves that ground chord events are of a complex nature. On a more fundamental level, however, those waves comprise sinusoidal waves in simple harmonic motion (henceforth: SHM). As John Rigden notes, SHM consists of "motion that repeats itself in a definite time interval called the period".[9] Hence, an essential aspect of the grounding base of chord events concerns the repetition rate of waves, i.e. their periodicity.

Pressure waves in SHM give rise to – or, under my account, they ground – what is known as 'pure tones'. The classic example of a pure tone is the sound generated from striking a tuning fork; its vibrating tines produce longitudinal pressure waves in the air that may be measured in cycles per second or Hertz (Hz).[10] Beyond historical variations, under the current system a tuning fork must vibrate at 440 Hz, which is the frequency corresponding to the pitch conventionally assigned to the note A4 on a keyboard.[11] Pure tones, such as those generated by striking tuning forks, may be described as "the simplest of all musical sounds".[12] That said, unlike what may be gleaned from this example, pure tones should not be confused with the musical sounds we typically hear. This is because the sounds we hear as distinctive musical notes are not pure tones, but complex ones.[13]

---

[9] Rigden, J., *Physics and the Sound of Music*, New York: John Wiley & Sons, 1977, pp. 12-13.
[10] The frequency of a pure tone is the rate at which a sinusoidal wave repeats itself. It is inversely proportional to its period, which is the length of time required for the wave to complete one full cycle. Speaks, p. 70.
[11] As will be detailed in Chapter Five, musical notes may be identified by a letter and a number. The letter refers to a musical note (i.e. C, D, E, F, G, A or B) and the number indicates the octave (i.e. the interval between two notes of the same name) in which the note is found (e.g. C2, D4, etc.). If the note is flattened or sharpened, the letter is followed by a specific sign (e.g. A♭ and C♯, for 'flat' and 'sharp' respectively).
[12] Rigden, p. 7.
[13] The nature of the auditory perception of complex tones will be detailed under my account of chords *qua* auditory objects in the next chapter.

Although pure tones are the simplest of all musical sounds, they are not normally encountered in the musical experience. For example, the sound of the note A4 on a piano is not acoustically identical to the pure tone produced by a tuning fork vibrating at 440 Hz. This is because musical notes (such as A4) are *complex* tones, instead. Chords are, therefore, constituted by complex tones, each of which is grounded by a number of component waves in SHM vibrating at different frequencies. These frequency components are known as 'harmonics' or 'partials', which are found in a sequence that is referred to as the overtone series. One of the defining properties of the overtone series is that its members bear a certain numerical relationship to one another and to the lowest frequency of all – which came to be known as the 'fundamental'.[14] Although all frequency components of the overtone series are 'partials', the term 'harmonic' is reserved for those that are an exact whole-number multiple of a fundamental frequency *f* (i.e. 2*f*, 3*f*, 4*f*, 5*f*, etc.) – where the fundamental is referred to as the first harmonic, the first partial/overtone is the second harmonic (2*f*), the second partial/overtone is the third harmonic (3*f*), and so forth.

Hence, although the pitch conventionally assigned to the note A4 is said to correspond to the pitch of the tuning fork, when the respective piano strings are sounded multiple frequencies will be generated in addition to that of the fundamental.[15] If they are integer multiples of *f*, they constitute the harmonic series of A4, where A4 is the note corresponding to the lowest frequency – i.e. the fundamental. In spite of the concomitant presence of the harmonic series in the waveform – and, rather paradoxically, because of it – the auditory system associates the sound emitted by the piano strings with the pitch of the fundamental frequency.[16] In other words, the note A4 can still be said to correspond to the pitch of the fundamental at 440 Hz, which is precisely the frequency of the pure tone produced by the tuning fork. Yet, A4 – as with all other musical tones – is a complex tone.

The fundamental frequency of a complex wave is thus essential for the perception of a musical tone. Alongside the overtone series, it constitutes a distinctive component of complex waves. Under the characterisation of the acoustic base of chords I have provided here, it is these complex waves that ground the complex tones comprised within chords.

---

[14] Although the nature of chords *qua* musical individuals will be considered in Chapter Five, it should be noted that the notion of 'fundamental' is equally important for a musical understanding of chords but it does not refer to *frequencies* as such. Specifically, the fundamental is the tone that occupies the lowest position in a certain chord structure (also known as the 'root'). For instance, C is the fundamental note of a C Major triad in root position, which comprises the tones C, E and G (typically in this order, but it may also be found as C–G–E). For the present purposes, 'fundamental' is to be understood in terms of frequencies, instead.

[15] For instance, if *f* = 440 Hz, the first overtone will be at 880 Hz (2*f*); the second overtone at 1320 Hz(3*f*); the third overtone (4*f*) will be at 1760 Hz, and so forth.

[16] That said, the overtone series is essential to the perception of musical sounds. For instance, as Parker notes, the presence of overtones is one of the main contributing factors to the experience of timbre. Parker, p. 65.

3. Necessary conditions

Having considered the nature of the grounding base of chords above, in this section I will characterise the identity of a chord event in relation to it. Specifically, I will identify three conditions for a chord event to exist. I am going to argue that the identity of a chord event is such that, for all *x*, if *x* is a chord event then the waves that ground it must be both periodic and simultaneous, and their frequencies must be mereologically amalgamated.

First, one of the defining features of the waves that ground most musical sounds is their periodicity.[17] As seen earlier, although pure tones are the simplest of all musical sounds, it is complex tones that are normally encountered in the musical experience. Either way, both sinusoidal waves in SHM (which ground pure tones) and complex waves (which ground complex tones) are *periodic*. Hence, the first necessary condition for the existence of a chord event is periodicity. This property of the waveform can be described in terms of the rate at which a particular wave pattern repeats itself at regular intervals, where the repetition rate is equivalent to the fundamental frequency of said waveform.[18] Under this condition, therefore, musical sounds that result from aperiodic waves cannot form chords.

With the exception of combinations of pure tones – which may be achieved by electronic means, for example – all chords comprise complex tones, but not all complex tones compose chords. This is due to the second necessary condition for the existence of chord events, namely, that of simultaneity. In line with my earlier definition, three or more complex tones can be said to form a chord event if and only if they are simultaneously occurring. This simultaneity requirement, as I am describing it, is determined by the superposition of the complex periodic waves corresponding to each of the component tones of a given chord – which, as I have argued, is entailed by a grounding relation. On this view, therefore, the so-called 'broken chord' – i.e. a chord in which the constituent tones are swiftly played one after another, rather than simultaneously – is, strictly speaking, not a chord but an 'arpeggio'.[19] This is because, in accordance with my proposal in the previous chapter, I take the simultaneity requirement to be best understood under the notion of temporal window – which means that a synchronous onset of tones is essential.

---

[17] Indeed, it is a standard presentation of the fundamentals of acoustics to classify sound waves in terms of their periodicity – i.e. as either periodic or aperiodic – and degree of complexity – i.e. as either sinusoidal or complex. Speaks, pp. 155-156.

[18] White, G. D., Louie, G. J., *The Audio Dictionary*, Seattle: University of Washington Press, 2005, pp. 284-285.

[19] "The notes of a chord 'spread', i.e. played one after the other from the bottom upwards, or from the top downwards". Latham, A., 'Arpeggio', *The Oxford Companion to Music*, 2011, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199579037.001.0001/acref-9780199579037-e-407.

As indicated in Section 2, chord events are grounded by the superposition of pressure waves, the frequencies of which are processed by the auditory system as corresponding to the individual pitches of the complex tones that form a given chord. But, under my proposal, this superposition process must occur within a given temporal window for a chord event to be classed as such. In other words, the simultaneity requirement is only satisfied if the component tones are sounded within the same temporal window. Specifically, their onset times must all coincide at the same time, when the 'initial attack' happens – i.e. the simultaneous sounding of the notes of a chord. Hence, for a sound event to be qualified as a chord, the coalescence of its constituent tones and simultaneous detection of their frequency components must occur within the same temporal window.

In speaking of constituent tones and frequency components, another metaphysical question comes to the fore, which concerns the relationship between a chord event and its components. In order to address that question, I shall conceive of frequency components and complex waves, on the one hand, and of complex tones and chord events, on the other, in terms of a part-whole relation. Given the scope of this chapter, the account I will offer here must be taken strictly in terms of an understanding of chords *qua* sound events.[20] In line with my thesis of sonic reconciliation, I want to argue that a mereology of chord events is best understood by considering that facts about the physical properties of their component tones ground facts about the acoustic properties of the chord as a whole. It is under this view that I will characterise the third necessary condition for chord events to occur – a requirement which I am going to refer to as 'mereological amalgamation'.

Two acoustic principles, in particular, are relevant for understanding why we may speak of 'mereology' and 'amalgamation'. First, under Fourier's theorem, each complex tone that corresponds to a given musical tone is a unique whole comprising component pure tones as parts. As Nicholas Giordano put it, "any musical tone ... can be written as the sum of one and only one combination of pure tones".[21] This sum comprises what came to be known as the Fourier series, and the process of decomposition of a complex periodic waveform into its component sinusoids is referred to as the Fourier analysis. It is, therefore, under the Fourier analysis that the mereological aspect of the third condition must be conceived: just as the theorem is applicable to complex tones, the same can be said of a chord. Hence, the resultant complex waveform of any given chord is also a unique whole which may be decomposed into its component parts and analysed accordingly.

[20] The mereology of chords *qua* auditory objects will be considered in Chapter Four (Section 2).
[21] Giordano, N. J., *Physics of the Piano*, Oxford: OUP, 2010, p. 13.

In addition, the third condition may be seen as a consequence of Ohm's acoustic law. Under this principle, the perception of musical sounds as such is due to the fact that "the human auditory system responds to a complex sound by generating sensations of the separate components of the sound rather than a sensation of a single integrated sound".[22] Yet, as shall be detailed in Chapter Four, the auditory system is able not only to map out different frequencies but also simultaneously group them based on certain cues. It is in view of this process of differentiation and combination that I am going to describe the third condition for chord events to occur as one of 'amalgamation'. To use the chemistry-based analogy, just as metal components are extracted by treating them with mercury and subsequently form an amalgam, a similar process occurs when incident complex waves reach an observer: the auditory system first extracts the frequency components from a waveform, only to 'amalgamate' them into the auditory object we hear as a chord. Hence, it is to this psychoacoustic process of decomposition of complex waves and combination of their frequency components that I refer to as 'mereological amalgamation'. But, in order to substantiate it further as a necessary condition for chord events to occur, three important observations must be made.

First, mereological amalgamation may be understood on two different levels. On strictly physical terms, it concerns the part-whole relation between different frequency components (i.e. the fundamental frequency and its overtones) and a complex wave. As seen earlier, the complex periodic wave that grounds a complex tone is a superposition of several sinusoidal waves in SHM, where the lowest frequency is recognised by the auditory system as the fundamental and the other frequency components are parts of its overtone series (i.e. its partials). By the same token, when complex waves converge into a chord event, they become parts of a further superposition – whereby the resultant waveform is what grounds the chord event *per se*. That is to say, they are parts insofar as their frequency components are 'amalgamated' into to the resultant waveform. Yet, on a metaphysical level, mereological amalgamation is best construed as the part-whole relation between complex tones and the unique whole that is the chord event. This is because the pure tones of the fundamentals – which give the chord's components their individual pitches – are parts of complex tones. In turn, complex tones are parts of a chord event. This not only conforms to the view that pure tones are the most elemental of all musical sounds but also to the transitive aspect of the sound-wave relation, as described in the previous chapter.

[22] Colman, A. M. 'Ohm's acoustic law', *The Oxford Dictionary of Psychology*, 2015, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199657681.001.0001/acref-9780199657681-e-5725.

Secondly, although I have taken the two acoustic principles identified above to constitute the empirical basis for mereological amalgamation, it should be noted that they indicate different operations performed by the auditory system. While the discriminating aspect of auditory perception is captured by Ohm's acoustic law, the auditory system in fact operates the reverse of the Fourier analysis by grouping frequency components, instead. As a consequence of these principles, Tan et al. point out that "we typically hear one sound as having one pitch, but can still make out distinct pitches when multiple sounds form a cluster (as when a pianist plays a chord)".[23] Mereological amalgamation may thus be seen as a two-way process, resulting *both* from the decomposition of complex waveforms *and* the grouping of their frequency components into the unique whole that is a chord event.

Thirdly, since sounds are events and chords are simultaneous combinations of *n* tones (where *n* is at least three), then the nature of the part-whole relation between tone-events and a single chord-event demands further explanation for mereological amalgamation to be workable. Such an explanation has been attempted by Casati and Dokic, under the formulation of their event theory. Indeed, of the key works in contemporary philosophy of sound Casati and Dokic's is alone in dedicating a full chapter to the ontology and phenomenology of chords. Briefly, they describe chords as events where various sounds resonate together within the same spatio-temporal region.[24] Most importantly for this section, however, is their explanation as to why chords must be treated as single events.

The initial premise of Casati and Dokic's argument is that two or more events can share their location in space without occupying it, since they are temporal entities. To illustrate this claim, they contrast the example of a sphere that spins on its axis whilst simultaneously heating itself up with that of a parachutist's descent in a spiral movement. In both examples, two simultaneous events happen to the same object: in the sphere's case, the spinning event and the heating event, while the parachutist's fall comprises both a descending movement and a spiralling one. However, Casati and Dokic point out a key difference between the two. In the case of the sphere, the movement of rotation and the event of heat generation belong to different categories, whereas the case of the parachutist involves two events that fall within the category of movements.[25] Casati and Dokic then argue that the case of chords is analogous to the parachutist's case, since the sound events that form a chord-event belong to one and the same category (i.e. that of complex tones).

---

[23] Tan, S-L., Pfordresher, P., Harre, R., *Psychology of Music: from Sound to Significance*, New York: Psychology Press, 2010, p. 32.
[24] Casati, R., Dokic, J., *La Philosophie du Son*, Nîmes: Chambon, 1994, p. 95.
[25] Ibid., p. 96.

Consequently, Casati and Dokic infer that we should construe tones and chords in terms of sub-events and larger events, respectively. Still drawing from the parachutist's example, they suggest that, if we have epistemic access to the causal chain that precedes the event of the descent, we may have sufficient reason to consider the descending spiral movement as resulting from two different components, while simultaneously regarding both components as the sub-events that compose the descent. Analogously, they take chords to be sound events that acoustically correspond to the resultant waveform of at least three main components, where the chord would constitute what they have termed the 'event node'. On this view, therefore, the identity of a chord is primarily determined by the structure of a tie between the sub-events, which they suggest could be warranted by certain 'relational factors' or emerging from a causal relation. That said, they do not elaborate on either of those terms or potential explanations, thus failing to provide a satisfactory account of chord identity.[26]

Granted, Casati and Dokic's conception of chord events as wholes comprising complex tones as their sub-events conforms to the view that I have proposed in this chapter. However, despite this element of common ground, their account seems vulnerable to two main weaknesses. First, their characterisation of chords as single entities is vague. While they treat chords as larger events comprising sub-events, they also take them to be akin to event nodes – without offering an explicit understanding of what the latter entails. Second, they describe chords as 'sufficiently linked objects',[27] but fall short of detailing what those sufficient conditions for the ties between chord events and the relevant sub-events would be – which further adds to the obscurity of their view.

Conversely, dovetailing with my account in Section 2, the characterisation of mereological amalgamation that I have proposed here is able to address those concerns. Specifically, the obscurity in Casati and Dokic's account is eliminated when we consider the nature of the grounding base of chord events. To be clear, I have argued that a chord event is grounded by the superposition of simultaneously occurring complex periodic waves, the components of which may be individuated by means of the Fourier analysis, for example. Indeed, it is this very possibility of decomposing its waveform into individual components that warrants the view that tone-events are parts of a single chord-event that is the whole. In addition, the mereological view I have proposed conforms to Ohm's acoustic law, thus equally capturing the differentiating aspect of the auditory perception of chord events.

[26] Ibid., pp. 97-98.
[27] Ibid., p. 98.

In the account of the nature of chords *qua* sound events that I have outlined above, I have sought to establish three necessary conditions for chord events to occur, in view of the nature of their grounding base – namely, periodicity, simultaneity and mereological amalgamation. However, these conditions are in themselves insufficient. In order to provide a full account of chord identity, we must also understand their nature *qua* auditory objects. Indeed, as stated at the beginning of this chapter, the case of chords not only conforms to my thesis of sonic reconciliation but also serves as an illuminating case study for explaining why sounds are best conceived both as events and as objects. It is, therefore, the task of the next chapter to give an account of chords *qua* auditory objects.

Chapter Four: Chords *qua* Auditory Objects

As indicated in the previous chapter, my account of the ontology of chords not only serves to illustrate my thesis of sonic reconciliation but testifies to its explanatory power. Under my proposal, we should conceive of chords both in terms of the superposition of complex periodic waves and as events, given the sound-wave grounding relation. Yet, we should also speak of chords as objects – i.e. the single entities we perceive them to be. In line with my thesis of sonic reconciliation, the view I shall defend in this chapter is that chords are auditory objects in their own right. In order to account for the perception of chords as such, I will examine both the psychoacoustic and phenomenological aspects of chord perception. In Section 1, the psychoacoustic view will be outlined in light of the workings of the auditory system. Subsequently, my account of the phenomenology of chord perception will be based on the claim that chords are best understood as Gestalt-based auditory objects, to the detriment of the notion of the phenomenal transparency of chords.

1. The psychoacoustics of chord perception

To conceive of chords as events is to think of them as temporal entities that are grounded by the superposition of pressure waves. But these sound events are auditorily grouped in certain ways, giving rise to the perception of streams that I have previously characterised as the auditory objects *per se*. In order to substantiate this claim, in this section I will examine how the auditory system 'fashions' chord events into auditory objects. To that end, I will describe the nature of chords *qua* objects as dependent upon the ways in which the auditory system processes simultaneously occurring sound events based on certain properties of sound waves – the most prominent of which being their frequencies. Specifically, while the frequency-selection function of the auditory system may be described as linear, other relevant aspects of chord perception involve non-linear processes, instead. In light of these two auditory mechanisms, my account will be focused on the perception of virtual pitch, on the one hand, and that of aural harmonics and heterodyne components, on the other. I will argue that these psychoacoustic phenomena indicate that chords *qua* auditory objects cannot be reduced to chord events grounded by pressure waves.

In my characterisation of the grounding base of chords in the previous chapter, I sought to highlight its complexity *vis-à-vis* the mechanics of sound wave production and propagation. Perhaps unsurprisingly – as indicated by both the Fourier analysis and Ohm's acoustic law – the physiological mechanisms underlying chord perception are just as complex. Hence, in order to account for it adequately, a further examination of acoustic properties and their implications for chord perception must be made. Two such properties, in particular, are central to the fundamental processes undertaken by the auditory system that are at work in chord perception. Specifically, they concern the fact that each complex tone in a chord has a distinctive *power spectrum* which determines its own *formant region*.

The power (or sound) spectrum is defined in terms of the amount of acoustic energy distributed amongst the different frequency components of the overtone series.[1] In turn, as White notes, the formant region of a complex tone corresponds to "a frequency band in its sound spectrum where sound energy is largely concentrated".[2] The power spectra and formant regions of individual complex tones thus result from specific frequency components and amplitudes of the partials in their respective overtone series. One of the roles of the auditory system in 'detecting' the complex tones that constitute a chord consists precisely in identifying their power spectra and formant regions by extracting those frequencies from the waves' properties. Upon reaching the ear, the waves corresponding to a chord's constituent tones set the eardrum to vibrate in such a way as to match the waveform that results from their superposition. Once those vibrations are transmitted through the middle ear, the waveform arrives intact at the inner ear – which is the part of the auditory system connected to the auditory nerve.

What is particularly striking about the auditory process briefly described above, as Rigden points out, is that "neither the eardrum nor the middle ear knows that they are transmitting a complex vibration consisting of the superposition of many simple harmonic oscillations. When the signal reaches the inner ear, however, the complex tone is separated into its individual frequency components".[3] This, in particular, is a consequence of the physiological role of the cochlea – a spiral structure found within the inner ear that functions as a "frequency-selective mechanism".[4] Since the pitches assigned to the components of a given chord are determined by the frequencies within the formant region of each individual complex tone, this inner-ear process is fundamental to chord perception.

---

[1] Giordano, p. 13.
[2] White, p. 92. This is particularly significant for tones produced by different musical instruments, since each of them will generate distinctive patterns of power spectrum.
[3] Rigden, p. 129.
[4] Olson, p. 245. As seen in Chapter Three, this process of differentiation is described by Ohm's acoustic law.

In order to account for the nature of chords *qua* auditory objects, it is therefore essential to consider how a complex waveform is decomposed as it reaches the cochlea. Specifically, its frequency components are processed on the basilar membrane, which is responsible for the roles of frequency-encoding and signal transduction that are fundamental to auditory perception at large. As Tan et al. describe it, while the former concerns the preservation of the information contained in the original signal that is conveyed by different sets of hair cells, signal transduction involves the conversion of fluid vibration into nerve impulses that are subsequently processed in the auditory cortex.[5]

This brief overview of the workings of the cochlea points to its frequency-selective function, which is crucial to understanding how the auditory system selects the component frequencies of a given chord. The cochlea effectively performs the role of a 'frequency analyser' by mapping out different frequencies on to the basilar membrane – a process Tan et al. refer to as tonotopical organisation.[6] This may be illustrated as follows: a complex periodic wave comprising frequencies of 200 Hz, 400 Hz and 800 Hz, for example, will generate three regions of relatively maximal oscillations which can be mapped out on to individual sets of hair cells at three different places on the basilar membrane.[7] Most importantly, what this reveals is an element of *linearity* in chord perception. In other words, the oscillation of individual sets of hair cells can be correlated with specific frequency components of the incident waveform, such that a linear frequency scale can be established along the cochlea – thus rendering its function 'tally-like', as it were.

Hence, tonotopical organisation seems to be what enables the inner ear to process several different frequencies within a complex waveform rapidly and efficiently.[8] Yet, this mechanism also involves the neurocognitive stage of the process of auditory perception, in that the primary section of the auditory cortex also appears to be tonotopically arranged. As Tan et al. describe it, the latter consists of a series of columns of neurons whereby "each cell in a column is tuned to the neuronal signal from a particular frequency in the auditory input into the cochlea".[9] Consequently, a certain location in the primary auditory cortex is matched to a particular location on the basilar membrane inside the inner ear, such that the auditory system is able to represent pitch spatially in a similar way to the distribution of keys on a keyboard – thus reinforcing the linearity of auditory perception.[10]

---

[5] Tan et al., pp. 47-48.
[6] Ibid., p. 50.
[7] Ibid., pp. 47-48.
[8] According to Olson, the auditory system is able to distinguish around 1,500 separate frequencies and resolve them within the audio-frequency range. Olson, p. 246.
[9] Tan et al., p. 49.
[10] Ibid., p. 50.

One of the ways in which the auditory system operates linearly concerns the assignment of a 'virtual pitch', which is the first psychoacoustic phenomenon that is particularly relevant for my account of chords *qua* auditory objects. Also referred to as the phenomenon of the 'missing fundamental' or 'residue pitch', this identification of a virtual pitch occurs when the lowest frequency of a complex tone happens to be absent from the incident waveform but, paradoxically, its corresponding pitch is still audibly perceived.

Arthur Benade gives a clear example of this psychoacoustic phenomenon. Consider the complex tone corresponding to the note A3 (which is an octave below A4 and has a pitch equivalent to the fundamental frequency of 220 Hz) and its harmonics (integer multiples of the fundamental frequency of A3, i.e. A4 = 440 Hz, E5 = 660Hz, A5 = 880 Hz, C♯6 = 1100 Hz, E6 = 1320 Hz, G6 = 1540 Hz, to name but a few).[11] "If we electronically remove the 220/second fundamental component from this collection", Benade writes, "our ears will nevertheless assign the pitch in accordance with a 220/second repetition rate. If we eliminate the second harmonic component (440/second) as well as the fundamental from our tone, we still have no hesitation assigning the pitch exactly as before".[12] Hence, the removal of the fundamental frequency from the incident waveform does not prevent the auditory system from recognising the other components as the overtones of a given fundamental frequency. In other words, we would still assign the original pitch (in this case, A3) to it, thus 'inferring' the fundamental from its harmonics.[13]

This process – also known as 'fundamental tracking' – is relevant for two main reasons. First, the perception of virtual pitch means that complex tones are not reducible to sound events grounded by pressure waves. As fundamental tracking indicates, their pitches are assigned to the lowest frequencies within their overtone series *even if* their fundamental frequencies are missing.[14] Consequently, neither are chords *qua* auditory objects reducible to chord events, in that they comprise complex tones as their constituent parts. Second, although fundamental tracking stems from the linearity of auditory perception, it also points to other consequences resulting from the intricacies of tonotopical organisation and pitch encoding for chord perception. Put simply, the processes underlying fundamental tracking also indicate the presence of *non-linear* mechanisms operating at inner ear level.

---

[11] The harmonic series is calculated by multiplying the fundamental frequency by a progression of integers, such that the resultant frequencies enable us to place the corresponding tones within the range of octaves.
[12] Benade, A. H., *Fundamentals of Musical Acoustics*, New York: Dover, 1976, p. 66.
[13] Tan et al., p. 33.
[14] As Rigden points out in more technical terms, the auditory system detects the pitches of their constituent complex tones from the periodicity of the resultant superposed waveform by matching them to the frequencies of their respective fundamentals. Rigden, pp. 129-130.

Given the physiology of auditory perception described earlier, one may expect the auditory system to perform its main functions via tonotopical mapping, i.e. by simply matching the input frequencies and cognitive output as though it were a frequency tally. Indeed, the spatial representation of the pitch spectrum that underpins tonotopical mapping would seem to point to that view. However, the auditory system's toolbox, as it were, includes some unorthodox methods of processing perceptual input that are equally important to accounting for chords *qua* auditory objects. Specifically, this is due to the non-linearity of the auditory system, which is central to the argument I shall develop here.

Beyond the linear effects of fundamental tracking, chord perception involves other frequency components that are also absent from the original vibration but which result from non-linear mechanisms, instead. Specifically, these concern the presence of 'aural harmonics' and 'heterodyne components'. In the case of aural harmonics, a pure tone that is sounded loudly is perceived as having a series of frequency components that are, nonetheless, absent from the incident waveform. As Rigden describes it, "if a pure tone of frequency *f* is played softly, only the tone of frequency *f* is perceived. As the intensity of the pure tone *f* is increased, however, additional tones are heard corresponding to frequencies 2*f*, 3*f*, 4*f*, etc".[15] In other words, the ear 'ascribes' an overtone series to an incident wave in SHM (provided it carries sufficient power) even when those partials are not present. Similarly, in the case of heterodyne components, when presented with the convergence of two pure tones the ear generates within itself additional frequency components in response to the corresponding waveform.[16] But, as in the first case, these components – which are also known as 'combination tones' – are absent from the stimuli that reach the ear.

Furthermore, the hypersensitivity of the cochlea[17] means that variations in amplitude and intensity of the frequency components of the incident waveform are likely to produce a mixture of both linear and non-linear effects. As Rigden illustrates it, "when f_1 and f_2 [where $f_2 = \frac{3}{2} f_1$] are played together *loudly*, a tone of frequency ½ f_1 is perceived. This is a combination tone and arises because of non-linear behaviour of the auditory system". Contrastingly, "if f_1 and $f_2 = \frac{3}{2} f_1$ are played together *softly*, a tune of frequency ½ f_1 is perceived, but in this case the auditory system's non-linearity is clearly not involved".[18] This is not only relevant in the case of pure tones but musical sounds at large, including chords.

[15] Rigden, p. 44.
[16] Benade, p. 258.
[17] Olson estimates that the cochlea comprises approximately 4,000 nerve fibres connecting it to the auditory cortex – which, in turn, contain around five hair cells each. Olson, p. 245.
[18] Rigden, p. 72.

The frequency-selecting function of the cochlea therefore has a dual mode of operation: depending on certain properties of the incident waveform, such as amplitude and intensity, the auditory system may behave in a linear and/or a non-linear fashion. As White points out, "the mechanical parts of the ear, up to and beyond the oval window, are linear in their responses to all intensities [such that] any non-linearity lies within the neural system of the cochlea itself".[19] In musical terms, while *pianissimo* tones are processed linearly, *fortissimo* ones cause the auditory system to behave in a non-linear fashion. Consequently, the varying intensity of a chord's frequency components will affect the formant region of its individual complex tones. Perhaps unsurprisingly, these variations in the nature of their formant regions will have an effect on how chords are perceived.

To be clear, there is more to chords *qua* auditory objects than what can be said of chord events grounded by the superposition of complex waves that reach the ear. This is because the auditory system adds new frequency components to the vibrations as and when they are processed in the inner ear – as a direct consequence of its linear and non-linear mechanisms. By linearly selecting the formants of each tone and grouping them with additional, non-linearly generated frequency components, the auditory system thus fashions what I will refer to as the 'chord percept'. Whether it be virtual pitch, aural harmonics or heterodyne components, these frequency components are solely present in the chord *qua* auditory object, in that they result from the inner workings of auditory perception – rather than from the mechanics of wave production and propagation. Hence, the significance of both linear and non-linear processes for chord perception could not be overstated.

This dual nature of the auditory system's *modus operandi* therefore has an essential bearing on the perception of chords. It both converts pressure waves and their properties into audible signals with certain perceptual attributes and, in doing so, it fashions the chord percept by linearly selecting frequencies and grouping them with additional, non-linear ones. To use my earlier terminology, these additional components are incorporated into the *stream* as it is perceived within a given temporal window. Yet, the stream is not reducible to the sub-events it comprises. In other words, the waves that ground a given chord event do not fully account for the auditory object that we perceive as a chord. Hence, the perception of frequency components that are absent from the incident waveform indicates that chords *qua* auditory objects cannot be reduced to chord events.

[19] Under White's assessment of the threshold between linearity and non-linearity, it is estimated that at an intensity level of 30 decibels and at the frequency of 350 Hz, for example, the auditory system "begins to be overloaded" and non-linearity ensues (the same applies to other intensity/frequency levels, such as 50 dB at 1000 Hz, for example). White, p. 188.

2. The phenomenology of chord perception

My account in Section 1 was aimed at describing the nature of chords *qua* auditory objects in light of the psychoacoustics of chord perception. But, in order to substantiate my claim that chords are particulars in their own right, it seems essential to consider not only how chord events are fashioned into chord percepts but also how they are *experienced*. Having previously argued that chord events must be individuated in terms of a coalescence of complex tones and their individual components, in this section I aim to consider how this 'coalescence' may be described in terms of the phenomenology of chord perception. My main argument here is twofold. I will first reject Casati and Dokic's assessment of the notion of phenomenal transparency and subsequently support certain aspects of Scruton's view of chords as *Gestalten*. To that extent, I will argue that chords are best construed as Gestalt-based auditory objects, rather than phenomenally transparent percepts.

In Chapter Three (Section 3), I argued that periodicity, simultaneity and mereological amalgamation are necessary conditions for chord events to be classed as such. In a given temporal window, complex tones *qua* constituent sub-events converge into the chord event grounded by the superposition of the respective waves. I then sought to outline in the previous section how the resultant waveform of said superposition is processed by the auditory system, both in linear and non-linear ways. These arguments are thus fundamental to substantiating my claim that chord events are grouped into the stream that is the auditory object *per se*. Indeed, if chords are to be understood as a coalescence of component frequencies, then considering how the auditory system is able to process it as such is essential. Yet, while the psychoacoustic view is a necessary component of my analysis of the nature of chords *qua* auditory objects, it must be considered in tandem with the phenomenology of chord perception for it to be sufficient in character.

Central to the nature of chords *qua* auditory objects is the perception that they are coalescences of musical sounds. As seen in Chapter Three, this is acoustically represented by the superposition of pressure waves that enter the auditory system as a complex waveform. Given the mereological nature of their amalgamation, just as pure tones and harmonics blend into a single waveform, simultaneous complex tones may also be perceived as fused into a chord. It is this coalescence of tones into a single waveform – plus the frequency components added at inner ear level – that warrants conceiving of a chord as a single entity. But chords are experienced as coalescing sounds with distinctive perceptual properties – and they may come to acquire distinctive musical properties as well.

In their account of the phenomenology of chord perception, Casati and Dokic attempt to describe how chords may be perceived both as single entities and composite objects via an assessment of the notion of phenomenal transparency. Specifically, their argument is based on Edmund Husserl's distinction between auditory and visual objects.[20] Under this distinction, visual objects can only be perceived as having relative transparency. This is because two visual objects cannot occupy the same location in space either without completely masking each other (e.g. if one piece of fabric is laid on top of another, completely covering its surface) or without both being perceived at the same time (e.g. if one piece of fabric is interwoven with another). Conversely, auditory objects are distinctive in that sounds may be 'fused' with one another completely and simultaneously. In Casati and Dokic's view, the case of chords exemplifies the latter phenomenon, seeing as they comprise three or more complex tones that are perceived as fused together and – to use their terminology – are experienced as being 'co-localised'. This "apparent co-localisation", they argue, is an essential characteristic of the nature of chords.[21]

However, Casati and Dokic's account departs from Husserl's with regard to the necessity of a chord's fusion. Under their analysis, although apparent co-localisation is a necessary condition for a chord to be experienced as such, it is insufficient, given that two sounds may simultaneously occur without being heard as a fusion. For instance, in hearing birdsong occurring simultaneously as the sound of road traffic, there is a perceptual distinction between a foreground and a background sound. On that basis, Casati and Dokic take auditory objects to be only contingently transparent to one another. Contrastingly, they argue that chords are not contingently transparent but, instead, that their constituent tones are necessarily located within a "region of connexion" seeing as they share the same "spatio-temporal address".[22] Yet, by failing to defining these terms, their account falls short of giving a satisfactory characterisation of chords as coalescences of complex tones.

Still, even if Casati and Dokic's account of chords as summarised above were sufficiently substantiated, their physicalist approach to the Event View – as outlined in Chapter Two – prevents it from being compatible with my proposal. By construing chords in terms of a necessary 'co-localisation', Casati and Dokic point not only to the fusion of different sounds over time but, primarily, to their actual location in space – as terms like 'region' and 'address' seem to indicate. Yet, as I have argued in that chapter, although auditory experience has spatial content, I do not take sounds to be located in space.

[20] Casati, Dokic, 1994, pp. 101-102.
[21] Ibid., p. 99.
[22] Ibid., p. 102.

That said, Casati and Dokic's assessment seems to raise an important issue regarding the perception of the mereology of chords. Unlike other auditory objects consisting of audible components that may be discriminated in auditory scene analysis, a chord's constituent tones would not be able to be segregated as such. This may, therefore, be construed as an argument for mereological essentialism, whereby the constituent tones are taken to be necessary proper parts of a chord. Arguably, one of the advantages of this view is that it conforms to my account of the mereology of chord events, under which the resultant waves that ground them are unique amalgamations of several frequency components. However, if one subscribes to a common understanding of mereology that takes "any whole as identical with the sum of its parts and consequently identifies any two objects containing all and only the same parts",[23] two difficulties seem to arise.

First, due to the non-linear mechanisms operating at inner ear level, it would seem incautious to suggest that the chord percept is a sum of its constituent tone-parts when a distinction between which components are originally parts of the incident vibration – i.e. physical properties – and which are non-linear components – e.g. combination tones – becomes indiscernible. Second, if chords *qua* auditory objects are streams heard within a given temporal window, they are individual occurrences and, as per the nature of events, they must be unrepeatable – which could arguably support the essentialist view.

By contrast, a more plausible phenomenology of chord perception has been put forward by Scruton, whose account draws from Gestalt theories. Although its principles have traditionally been associated with visual phenomena, they have come to be applied to other aspects of sensory experience such as sound perception. As seen in Chapter Two, Gestalt theorists seek to explain the nature of percepts as *Gestalten* (i.e. 'structures'), since they take them to be wholes which cannot be reduced to a breakdown of their constituent elements.[24] Furthermore, they hold that "the emergent or whole property of an aggregate and of the elements contributing to that whole influence each other".[25] Hence, for Scruton, chords are best conceived as such structures, comprising several tones that are perceived in unity due to a range of factors – amongst which is that of simultaneity. To conceive of a chord percept as a *Gestalt* is thus to think of it as a structure that is perceived as a single entity, but one in which the constitutive aggregates (i.e. complex tones) are simultaneously held together, influencing one another and being influenced by the organised whole.[26]

---

[23] Lowe, E. J., 'Mereology', In Honderich, T. (ed.), *The Oxford Companion to Philosophy*, Oxford: OUP, 2005, pp. 587-588.
[24] Hamlyn, p. 338. For further detail on Gestalt theories, please refer to footnote no. 26 on page 49.
[25] Tan et al., pp. 77-78.
[26] Scruton, 1997, p. 40.

Similarly to Scruton, I want to argue that chords are Gestalt-based auditory objects, and that is for three main reasons. First, the view that chords are *Gestalten* is compatible with my argument for the non-spatiality of sounds, as outlined in Chapter Two. Although Gestalt theories of perception have tended to be considered in terms of visual phenomena, Gestalt principles such as those of similarity, proximity, closure and good continuation have, as Tan et al. point out, "a more or less direct application to the perception of musical wholes, [in that] tones that are close together in *pitch, time or space* tend to be perceived as a group".[27] Yet, these principles are not intrinsically spatial, so they are no less applicable to temporal entities than they are to space-occupying objects.

Secondly, conceiving of chords as Gestalt-based auditory objects both informs and conforms to my account of the psychoacoustics of chord perception in Section 1. As seen earlier, in addition to selecting frequencies from the incident waveform, the auditory system is able to group them in such a way as to assign pitches to the repetition rates that it identifies. Most importantly, the auditory system's ability to perform tonotopical mapping whilst also 'filling in the gaps' with virtual pitch, for example, may be accounted for in terms of Gestalt principles. As Tan et al. note, these mechanisms indicate that the auditory system "responds similarly (e.g. with a similar cochlear response) to frequencies in close proximity to each other", such that "frequencies that evoke a similar response are said to fit within the same critical band".[28] Hence, the Gestalt principle of proximity is compatible with psychoacoustic phenomena occurring at inner ear level.

Thirdly, the Gestalt approach is more conducive to accounting for the distinction between the experience of harmony and that of polyphony.[29] Although both result from simultaneously occurring musical sounds, Scruton correctly maintains that "the experience of harmony must be distinguished from the experience of 'simultaneities', i.e. the experience of pitched sounds occurring together".[30] Correspondingly, he establishes a distinction between two types of musical coalescence: "chords, in which separate tones are arranged 'vertically', to form a new musical entity; and polyphony, in which the component parts are melodies, rather than tones, and the resulting entity is not a sequence of chords, but a musical movement *through* chords".[31] However, Scruton falls short of weaving this distinction with his account of the nature of chords as *Gestalten*.

---

[27] Tan et al., pp. 78-79.
[28] Ibid., p. 33. The frequency range known as the 'critical band' will be particularly relevant for the discussion around dissonance perception in Part II, so it will be considered in Chapter Seven (Section 4) instead.
[29] Polyphony may be succinctly defined as "a texture arising from the simultaneous combination of two or more melodic lines". Karp, T., *Dictionary of Music*, New York: Dell Publishing, 1973, p. 300.
[30] Scruton, 1997, p. 63.
[31] Ibid., p. 65.

In order to address that weakness in Scruton's account, I take the disparity between the experience of chords and that of polyphony to be the result of different Gestalt principles at work in the phenomenology of musical perception. In a typically polyphonic context – as observed in the practice of counterpoint[32] – different voices (i.e. melodic lines) are meant to be heard as independent of one another, but certain combinations of harmonic intervals may result when the voices overlap. Yet, the auditory system is still able to track them down as independent melodies, rather than identifying the overlapping instances as chord percepts. This ability may thus be accounted for in terms of the Gestalt principle of good continuation as the overriding one, instead. Conversely, the perception of three or more simultaneous tones as belonging within a chord percept may be understood under other Gestalt principles, such as proximity in time (e.g. in terms of the initial attack), similarity of pitch (e.g. when sounded within the same octave), and closure in a given musical context (e.g. in the case of resolution).[33]

The purpose of this chapter has been to provide an account of the nature of chords *qua* auditory objects, which further substantiates my thesis of sonic reconciliation. In my analysis of the psychoacoustics of chord perception, I sought to explain how chord events are processed by the auditory system and why chord percepts are not reducible to them. Having subsequently accounted for the experience of chords *qua* auditory objects in terms of Gestalt principles, in the next chapter I will seek to describe how they may come to bear distinctive musical properties – and how they can be individuated as such.

[32] In *The Oxford Dictionary of Music*, 'counterpoint' is described as "the ability, unique to music, to say two or more things at once comprehensibly". More prosaically, however, it is sometimes described as a synonym for polyphony, although it is best characterised as a specific example thereof. Kennedy, J., Kennedy, M., Rutherford-Johnson, T. (eds.), 'Counterpoint', *The Oxford Dictionary of Music*, 2013, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199578108.001.0001/acref-9780199578108-e-2218.

[33] The notion of resolution will feature more prominently in Chapter Eight (Section 3), but for the current purposes it may be defined as a progression from a dissonant chord either to a consonant chord or a less dissonant one (Karp, p. 325), thus generating an auditory sensation of closure or a certain degree of repose.

Chapter Five: Chords *qua* Musical Individuals

In line with my thesis of sonic reconciliation in Chapter Two, the account of the ontology of chords I have pursued thus far has been centred on the claim that chords are auditory objects as well as events. In Chapter Three, I argued that the nature of chords *qua* sound events is best understood in terms of a grounding relation, as determined by their acoustic base. Subsequently, I sought to account for the nature of chords *qua* auditory objects by considering both psychoacoustic and phenomenological aspects of chord perception. Yet, chords are also taken to be distinctively musical objects in tonal harmony. Hence, an account of the ontology of chords must also include an analysis of their nature *qua* musical individuals, since they bear distinctive musical properties in their own right. It is the purpose of this chapter to provide such an account.

1. Chord identity and individuation

In the preceding chapters, I have defended the view that chords are auditory objects as well as events. But the ontology of chords gains a whole new dimension when one considers their nature *qua* musical individuals. Although chords may come to acquire a number of context-dependent qualities which ascribe to them specific harmonic functions in a given musical phrase, it is with the musical properties of chords conceived as individual entities – i.e. rather than in a specific musical context – that my analysis here is concerned. Hence, in this first section I will provide an analysis of the identity and individuation of chords *qua* musical individuals. To that end, I am going to argue that we can best account for the distinction between chord patterns and the chord events that constitute their instances in terms of the type/token theory. Specifically, I will support this view by considering the intervallic relationships that determine the structure of a chord type.

Composers and musicians alike treat chords as distinct musical individuals. In traditional tonal harmony, they are most commonly found in the form of triads and seventh chords,[1] but they may also appear as ninth, eleventh or thirteenth chords.[2] Triads

[1] Sevenths are the most common expression of a tetrad. They result from extending a triad by an interval of a third, and are known as such because the interval between the root (i.e. the bass note) and the uppermost note is that of a seventh.

may be either major or minor; sevenths may also be major-minor, diminished or half-diminished; and chords comprising five or more tones have increasingly complex tonal qualities. Chords may be in root position (i.e. when the bass note is in the lowest position), or they may be found as an inversion.[3] They may be altered,[4] suspended,[5] and they may also constitute unusual patterns such as that of the famous 'Tristan chord'.[6]

When speaking of a certain chord as having some of the properties mentioned above rather than others, one may be referring to a specific chord event. Consider, for instance, a C Major triad as it is sounded on a piano keyboard. It results from simultaneously pressing the keys corresponding to the notes C, E, and G, in any given octave. As I argued in Chapter Three, the initial attack must happen concomitantly on each individual key so as to generate a chord (as opposed to an arpeggio). If the notes are in that particular order (i.e. C–E–G), the chord is said to be in root position – in that C is the tonic of the C Major scale. In addition to identifying it as a major triad in root position, upon hearing the chord event one is also able to characterise it in terms of its loudness and duration, as well as certain aesthetic properties such as consonance – or, in other cases, the lack thereof.[7]

Conversely, when speaking of a given chord one may instead be referring to a certain chord pattern, rather than the chord event itself. In the previous example, the chord has been specified as a C Major triad in root position. Once the chord is heard, one is also able to specify *which* C Major triad has been sounded. While it could be the triad in C2–E2–G2, on the lower register of the keyboard, it might alternatively be the C6–E6–G6 chord at the opposite end. Although both are classed as a C Major triad, they are different instances of it, aside from being different chord events. Yet, prior to it being sounded, simply by considering its designation one has epistemic access to its structure (i.e. a root-position triad in C), its modality (i.e. major) and its mereological composition (i.e. three tones, for example: C2, E2 and G2). In other words, one is able to know that it is a specific pattern of a C Major triad that is instantiated by one of those chord events. Hence, this reveals a fundamental distinction under which chords *qua* musical individuals must be understood.

---

2 Ninths, elevenths and thirteenth chords comprise five, six, and seven tones, respectively. Chords with eight or more tones are equally possible, albeit rarely used in traditional tonal harmony.

3 For instance, the C Major triad in root position is C–E–G, its first inversion is E–G–C and the second one is G–C–E. Triads only have two inversions, whereas sevenths have three. Chords comprising five or more tones allow for an increasing number of inversion forms.

4 Altered or chromatic chords are those which include one or more tones that do not belong to the key signature of a given musical passage.

5 Suspended or 'sus' chords are those in which a third interval is replaced by either a perfect fourth or a major second in relation to the root. Its irregularity means it cannot be classed as either major or minor.

6 The Tristan chord is a tetrad named after its appearance as the first chord of Wagner's *Tristan und Isolde*, in which it figures as F–B–D♯–G♯. It will be considered further under my analysis of chord types in Section 3.

7 The aesthetics of consonance and dissonance will be considered at length in Chapter Eight.

On the one hand, chords are the sound events that are fashioned into the auditory objects that we hear. As such, they bear both acoustic properties – due to the pressure waves that ground them – and perceptual attributes – depending on how the resultant waveform is processed by the auditory system. But the musical qualities that are ascribed to a chord involve a range of other factors, amongst which are its timbre, its tonality and the specific context in which it occurs. On the other hand, chords are also the patterns of relationships between their constituent tones. In this case, one of the key properties of these patterns is that of multiple instantiation; they may be sounded either within the same octave or across different ones, played on different instruments and at multiple times. Still, their distinctive pattern is able to be identified, and has conventionally been designated by means of a system of musical notation.

Hence, in order to account for the identity and individuation of chords *qua* musical individuals, the first question that must be addressed concerns the nature of the relationship between chord events and chord patterns. While the former are temporal entities grounded by pressure waves – and are, as such, unrepeatable – the latter refer to atemporal structures that represent certain composite abstract objects that *are* repeatable. It is thus essential for an account of the ontology of chords to consider their nature *vis-à-vis* this distinction – which is representative of a larger debate in the metaphysics of sounds. As seen in Chapter Two, this apparent problem was most clearly articulated by Dokic, who identifies two contrasting ontologies of sound. In his view, sounds are either repeatable objects or unrepeatable events. Given that the question of repeatability is particularly relevant for thinking about chords, I have found it more apt to address it here.

In the preceding chapters, I proposed an ontology of chord events that is not only compatible with my account of chords *qua* auditory objects but essential to it. This is because the incident waveform that results from the superposition of complex waves is what enables the auditory system to forge the chord percept as such. Yet, in describing chords as auditory objects which result from the frequency-selecting/grouping process occurring at inner ear level, I have also sought to provide an empirical basis for distinguishing between chord events and chords *qua* auditory objects. That said, the notion of multiple instantiation did not feature in my presentation of either of those accounts, and this is for a simple reason: neither chords *qua* events nor chords *qua* auditory objects – as I have described them – are repeatable entities. As Dokic points out, it follows from the very nature of events that they are unrepeatable; so if that is what sounds are – which both Dokic and I agree is the case – then chords are unrepeatable sound events.

However, in his rejection of the repeatable-object conception (i.e. the Object View), Dokic offers a characterisation of it that is not compatible with my account of chords *qua* auditory objects. This is because what I take to be the repeatable objects are chord *patterns*, instead. By describing sounds as particular kinds of objects that are "capable of having multiple occurrences",[8] Dokic's statement of the Object View is better suited to qualify what I have referred to as chord patterns, as opposed to chord percepts. In other words, whichever way we may come to understand chords as auditory objects, Dokic's dichotomy precludes the possibility that they may be conceived both as objects and events. However, incorporating them both is fundamental to an account of chords *qua* musical individuals, in that they must be understood as capable of picking out *both* the repeatable chord patterns that allow for multiple instantiation *and* the chord events themselves.

Indeed, chord events are particular occurrences and, as such, they are unrepeatable. Yet, beyond the individual circumstances of sound production and in view of the mechanisms of auditory perception, it is an empirical fact that certain patterns of musical sounds may be observed. For instance, despite variations in timbre and loudness, we are able to identify sounds that have the same pitch across a range of instruments and at different times. Albeit not a peculiarly musical phenomenon, this multiple realisability of musical sounds thus raises an important question regarding the ontological category to which they belong.

When considering the ontology of musical sounds at large, Scruton argues that musical individuals could not be treated strictly as concrete particulars, since they are reidentifiable. Although he takes the experience of 'same again' to be "an experience of similarity, and not the 'recognition of an individual' in any strict sense", he concedes that chords are an exception, in that we are able to recognise a certain pattern in them.[9] Granted, although Scruton points to the presence of salient features common to particular sounds such that they may be characterised as types, he stops short of developing an account of chords to that effect – a weakness that this section is intended to address.

Hence, I want to argue that the best way to account for this multiple instantiation of chord patterns is in terms of the type/token theory. To that end, my argument will be conducted in light of Julian Dodd's ontology of musical works. Specifically, he champions the explanatory power of the type/token theory in accounting for the relationship between musical works and performances thereof – which, I believe, can equally illuminate the relationship between chord patterns and chord events.

---

[8] Dokic, p. 392.
[9] Scruton, 1997, p. 106.

Dodd argues that the type/token distinction is best placed to account for the repeatability of musical works "without compromising either our modal intuitions concerning works' occurrences or our presumption that such works are entities in their own right".[10] Under his proposal, any given musical work is a type, and each individual sound-sequence event that instantiates it is a token of it. Analogously, I take chord patterns to be types, such that chord events are tokens that instantiate particular chord types. While my account of chords *qua* musical individuals is not intended to hinge upon any particular ontology of types, I take the analogy between chords and musical works to be particularly useful for illuminating what I mean by describing chord patterns as types.

Under Dodd's characterisation, types are "abstract (i.e. not located in space), unstructured (i.e. without parts), and both modally fixed and temporally inflexible (i.e. items that possess their intrinsic properties necessarily, and which are incapable of change in these properties through time)".[11] Consequently, works of music *qua* types are, on his view, structureless, fixed, unchanging and eternal entities. Yet, the corollary of Dodd's ontology of types of musical works would bear some limitations were it to be extended to the ontology of chords. Unlike works of music, chord patterns *are* structured entities, in that they comprise different parts which stand in specific relationships to one another.[12]

Although this is an important contrast, I want to argue that it does not undermine the strength of the analogy. This is because, unlike what Dodd suggests, being 'structureless' does not seem to be a necessary condition for type identity. Rather, I take the other properties Dodd associates with types (i.e. abstractness as well as modal/ temporal inflexibility) to be the ones that best characterise 'type' as an ontological category – properties which more adequately describe the nature of chord patterns. Furthermore, the most distinctive aspect of type ontology is that it allows for multiple instantiation – a condition which chord patterns conform to *qua* structured types, instead.

Hence, there are two main ways in which we may conceive of chords *qua* musical individuals. When we speak of chord patterns, we are conceiving of them as abstract types; conversely, when we point to a particular chord event, we are referring to a token of a certain chord type. Yet, from the same chord type – such as a C Major triad – more specific patterns may arise – such as C2–E2–G2 or C4–E4–G4, for example. For that reason, I shall refer to those octave-indexed patterns as chord *subtypes*, since they preserve the same structure of a given chord type whilst designating specific pitch classes.

---

[10] Dodd, J., *Works of Music: an Essay in Ontology*, Oxford: OUP, 2007, p. 82.
[11] Ibid., p. 37.
[12] My account of the nature of chord types will be substantiated further in Section 3.

To characterise chords in terms of types, subtypes and tokens is thus to presuppose four conditions which I take to be necessary for chords *qua* musical individuals to exist: (i) chord patterns must conform to a certain structure; (ii) their structure must be determined by certain intervallic relationships between pitch classes; (iii) any given pitch class 'A' must correspond to a certain pitch height 'A*n*', where *n* indicates its octave position;[13] and (iv) in tandem with my characterisation of chord events in Chapter Three, three or more specifiable pitch classes must be simultaneously instantiated for a chord to be individuated *as a single entity*. While (i) and (ii) concern chord types, (iii) and (iv) are necessary conditions for tokens of chord events to be counted as instances of a given chord subtype. It is in view of these conditions – underpinned by the distinction between types/subtypes and tokens – that my account of chords *qua* musical individuals is to be understood.

In order to grasp the nature of chord subtypes, it seems necessary to consider first how their respective tokens may be individuated. As Dodd correctly points out, "it is one thing to provide an answer to the categorial question [by adopting the type/token distinction]; it is quite another to give an account of how musical works are individuated".[14] In line with the earlier analogy, I want to suggest that the individuation of chords as tokens of certain subtypes mirrors that of performances that instantiate certain musical works. In Dodd's view, the simplest and best way to individuate sound-sequence events as performances (tokens) of a given composition (type) is in terms of sonicism – which I take to be helpful in accounting for the multiple instantiation of chord types and subtypes. In his iteration of sonicism,[15] Dodd maintains that a sound-sequence event would only count as a properly formed token of a given musical work – i.e. the norm-type – on the basis of its acoustic qualitative appearance. In his words, in the case of musical works "there is no more to the individuation of such works than how they sound".[16] This, however, creates some instrument-based and context-specific difficulties for the identity of musical works – which Dodd attempts to address in his work.[17]

---

[13] This is a more formal version of the characterisation I provided earlier in Chapter Three (specifically, in footnote no. 11 on page 59). It should be noted that octaves are numbered with reference to the C scale, so chord pattern components may be found across octaves (e.g. a G major triad instantiated in G4–B4–D5).
[14] Ibid., p. 201.
[15] The term 'sonicism' has been used in reference to different theses in the philosophical literature on sound. For instance, under Leddington's account it means that all we can hear is sounds, such that objects can only be heard *in* or *in virtue of* hearing a sound. Leddington, J., 'What We Hear', In Brown, R.(ed.), *Consciousness Inside and Out*, Dordrecht: Springer, 2013, p. 327.
[16] Dodd, p. 6.
[17] Dodd devotes two chapters of his book to countering said difficulties, which he identifies as two contrasting theses: one of 'instrumentalism', whereby musical works written for different instruments cannot be identical; and that of 'contextualism', which renders individual compositions tied to specific musico-historical contexts. (Dodd, pp. 202-203). However, neither of these presents a problem for the case of chords, as shall be subsequently argued.

Yet, I want to argue that the conditions I have identified above preclude this kind of objection in the case of chords. This is because neither instrumental differences nor circumstantial aspects of sound production should prevent a chord event from being identified as a token of a chord type/subtype. The only constraint would be that the instrument be uniformly tuned to a given standard – a requirement that I take to be implicit to my account in this chapter and throughout the present study. This being the case, my argument for the individuation of chords as tokens of certain chord subtypes is compatible with Dodd's iteration of sonicism. In other words, it is the acoustic qualitative appearance of a chord event that enables us to identify it as an instance of a certain chord subtype.

To be clear, unlike the case of musical works, neither the timbre of an individual instrument nor the musical context in which it occurs would seem to present difficulties for identifying specific chord events as tokens of chord subtypes.[18] Granted, the harmonic function of C–E–G as the first degree of the C Major diatonic triad scale[19] is different from that of the same chord pattern when it occupies the fifth degree – i.e. the dominant triad – of the F Major diatonic scale. However, for the purpose of individuating a chord event as a C Major triad in root position, the harmonic function it performs is not relevant to identifying it as a token of the root-position C Major type. Rather, it is the individual occurrence of a chord event, fashioned into an auditory object by the auditory system, which determines the parameter for token individuation.

Consequently, a given chord token may be identified on the basis of its acoustic qualitative appearance and independently of the musical context in which it occurs. To use the previous example of a C Major triad in root position, upon hearing the chord event one is able to have epistemic access to the nature of the chord's components – i.e. the pitch classes C, E and G – as well as their octave positions. Granted, this would require a well-trained ear; yet, it is nonetheless epistemically possible for one to identify chord tokens based on how they sound. However, sonicism should not be taken as a substitute for the conditions identified earlier. Rather, as conditions that they are, they presuppose the perceptual ability to identify chord events as tokens of certain chord types/subtypes. But in order to substantiate this view further, we must consider how the auditory experience of musical sounds at large may contribute to the phenomenology of chord tokens.

---

[18] This is to the exclusion of cases of polyphony which, as I argued in the previous chapter, would not qualify as instantiating chords as Gestalt-based auditory objects.

[19] For any given scale of major or minor tonality, one is able to construct a diatonic scale of triads starting from the tonic through to the leading note. Each triad thus occupies a certain degree of the scale, in relation to the tonic, which conforms to certain patterns; for instance, the first, fourth and fifth degrees of any given major diatonic triad type will always be major chords.

2. The phenomenology of chord tokens

In Chapter Four, I characterised the nature of chords *qua* auditory objects as resulting from both linear and non-linear aspects of auditory grouping. But, in order to account for the nature of chords *qua* musical individuals, we must consider how pitched sounds come to be heard as distinctive pitch classes within a specific octave. In this section, I aim to examine the perception of a chord token as a coalescence of its constituent tones *vis-à-vis* their acoustic qualitative appearance as typically musical sounds. To that end, I will consider the phenomenology of chord tokens in light of Scruton's twofold account of the experience of pitch and the acousmatic experience of musical sounds at large. As indicated by my analysis of chords *qua* sound events and that of chords *qua* Gestalt-based auditory objects, I will subsequently characterise the coalescent nature of chord tokens as equally fundamental to conceiving of their identity as musical individuals. Specifically, I will describe a chord token as a temporal Gestalt occurring within a given temporal window, which also conforms to my earlier account of sounds as perduring entities.

Chords *qua* musical individuals comprise typically musical sounds as their constitutive parts, where the latter are designated as specific pitch classes corresponding to specific complex tones. Hence, in order to account for the perceptual experience of chord tokens, it is essential to consider the phenomenology of pitch perception first – a task which has previously been attempted by Scruton. In his view, we experience pitch as a continuum of higher and lower pitched sounds, which is determined by variations in their frequencies. Scruton sees it as analogous to the experience of colour – i.e. as akin to a spectrum: "between any two colours or pitches", he states, "there lies a third, even if its character is not, to us, perceivably different from its neighbours".[20]

In addition to the vertical dimension of the experience of the pitch continuum (i.e. in terms of high and low 'pitchedness'), the phenomenology of pitch perception is, according to Scruton, marked by two other characteristics. First, pitches are perceived in terms of certain distances between them, which he takes to underpin the perception of musical intervals. Second, Scruton sees this distance between tones as the basis for a pitch continuum that offers a distinctive experience of 'same again' – which has been framed with reference to octaves under the standard of Western musical notation. Octaves, in particular, have eventually come to be divided into twelve different segment-units, which are assigned to specific pitch classes that are fixed (e.g. C, F♯, G, A♭, etc.)

[20] Scruton, 1997, p. 20.

Most importantly, Scruton notes that these aspects of pitch perception mean that the continuum is experienced in two different ways: one that is more closely associated with the physical properties of sound production – i.e. a continuum of frequencies – or, alternatively, another form of pitch perception where an order is "imposed upon" this continuum. It is in the latter case, he argues, that pitched sounds become tones that are organised in patterns of discrete intervals, which are replicated in every octave – thus forging the domain of tonality.[21] Similarly, as Tan et al. describe it, musical pitch is a "multi-dimensional percept": the experience of the frequency continuum is the basis for the perception of pitch height, and this is, in turn, correlated with certain pitch classes (i.e. in terms of discrete intervallic relationships between them, such as C– F♯or G– A♭). Hence, the correlation between these dimensions of pitch perception means that changes in pitch height correspond to changes in pitch class.[22] It is in view of this notion of pitch height and its correspondence with specific pitch classes that (ii) and (iii) above must be understood.

The distinction between these two forms of pitch organisation constitutes the basis for one of Scruton's main thesis in his account of the musical experience. As stated in previous chapters, Scruton takes sounds to be pure events, "which are identified in themselves, and not through other things".[23] As such, sounds may be heard as completely detached from their physical causes – or, as he puts it, they may be heard *acousmatically*. Specifically, Scruton argues that this is precisely the case with musical sounds. In keeping with his account of pitch perception, he distinguishes between two ways in which sounds may be experienced: "an acoustical description", he writes, "refers to pitched sounds and their secondary qualities", whereas a musical description "refers to the tones that we hear in those sounds, and to their audible relations in musical space".[24] Hence, Scruton's acousmatic thesis dovetails with his view of sounds as pure events, in that it points to their independence or apparent 'detachability' from their sources.

The acousmatic view of musical experience has nonetheless attracted some criticism. In *Aesthetics & Music*, Andy Hamilton challenges Scruton's acousmatic thesis by claiming that it "neglects the importance of the human production of musical sounds to which appreciation of music makes essential reference".[25] For Hamilton, the ways in which we experience music are fundamentally associated with the unique nature of a given musical instrument as a physical object that produces a distinctive musical sound. In particular, one

[21] Ibid., p. 22.
[22] Tan et al., p. 74.
[23] Scruton, 1997, p. 106.
[24] Ibid., p. 402.
[25] Hamilton, A., *Aesthetics & Music*, London: Continuum, 2007, pp. 95-96.

of the properties of musical sounds that illustrates this objection is the perceived quality of the sound produced, i.e. the experience of timbre. Specifically, Hamilton argues that timbral experience cannot be detached from the instrument that produces it. For instance, sounding A4 on a harpsichord has a distinctive quality to it when compared to sounding it on a piano. Being able to perceive the aesthetic contrast between the two tones is, in Hamilton's view, an essential aspect of the musical experience.

Another objection to Scruton's acousmatic thesis has been raised by Aaron Ridley, who differentiates between the possibility of sounds being detachable from their sources – which he concedes to a certain extent – and that of sounds being "actually (or necessarily) detached", which he rejects.[26] Similarly, Nudds and O'Callaghan take Scruton's view as tantamount to depriving sounds of "a constitutive ontological connection with the vibrations or activities of objects we ordinarily count as sound sources".[27] However, they acknowledge that, on a charitable reading of Scruton's thesis, it may be said to require "only the capacity to experience and attend to sounds *as* independent from their sources, rather than the capacity to experience sounds without experiencing their sources".[28]

This wider debate around the nature of the musical experience is particularly relevant for my account of the phenomenology of chord tokens. While I agree with Hamilton's objection that the nature and circumstances of sound production may be perceived as essential aspects of the phenomenology of the musical experience, the more charitable interpretation of Scruton's acousmatic thesis offered by Nudds and O'Callaghan seems to point to the heart of the relationship between pitch and tonality. In distinguishing between the acoustical and musical dimensions, Scruton is equally considering the difference between the psychoacoustic experience of the pitch continuum and the distinctively musical experience of tonality. On this view, it is by means of the possibility of experiencing sounds acousmatically that pitched sounds are phenomenologically converted into tones – and these, in turn, into specific pitch classes in intervallic relationships. Hence, it is under that charitable interpretation that I support the acousmatic thesis as part of my account of chord-token individuation, seeing as chord tokens result from coalescences of complex tones that acquire certain musical properties independently of their sources.

[26] Ridley notes that certain sounds (e.g. unknown noises, the sources of which may not be so easily detected) may be treated as free-standing individuals and thus seemingly independent from their sources. Yet, he argues that this does not mean that they do not stand in any special relation to their sources. Ridley, A., *The Philosophy of Music: Theme and Variations*, Edinburgh: Edinburgh University Press, 2004, pp. 53-54.
[27] Nudds, O'Callaghan, p. 6.
[28] Ibid., p. 15. Similarly, Hamilton concedes that both the acoustic and the acousmatic dimensions are important for the appreciation of music, which is to say that he acknowledges the possibility of attending to musical sounds acousmatically. Hamilton, p. 98.

As indicated by (iv) above, the simultaneous instantiation of three or more pitch classes is necessary for a chord token to exist. Perceiving a chord as a 'vertical' coalescence of different complex tones is, as seen in Chapter Four, the hallmark of the experience of harmony – in contrast to the horizontality of the experience of polyphony as independent melodies (even when harmonic intervals are formed). Hence, to conceive of chord tokens as musical individuals is fundamentally to consider their nature as a fusion of different components that is experienced as a single entity. As seen in previous chapters, I sought to characterise this coalescent nature of chords both in terms of the superposition of pressure waves, on a physical level, and of the Gestalt-like structure of chord percepts, from a phenomenological perspective. As would be expected, my account of the nature of chords *qua* musical individuals both conforms to that proposal and further illuminates it.

The coalescent nature of chords is a defining aspect of their nature both *qua* sound events and *qua* auditory objects. On the one hand, chord events are grounded by the superposition of complex waves, which is a form of coalescence. On the other hand, the chord percepts are best understood in terms of a temporal Gestalt resulting from the concomitant attack of its constituent tones within a given temporal window. As seen in Chapter Three, the perception of their coalescent nature depends upon an initial attack, which is fundamental for the simultaneity requirement to be met. In particular, the role of 'starting transients', which James Beament describes as "complex patterns of rapidly changing sounds at the beginning of notes" is significant to the perception of chord tokens *qua* Gestalt-based auditory objects – and, consequently, *qua* musical individuals as well.[29] Hence, simultaneous onset times are crucial for the phenomenology of chord tokens, in that the perception of starting transients contributes to forging the temporal Gestalt – as previously indicated by my analysis of auditory grouping mechanisms in Chapter Four.

In addition to a concomitant attack, the phenomenology of chord tokens must also be informed by how they are subsequently experienced within a given temporal window. Specifically, this concerns the role of decay effects on chord perception. While 'attack' refers to the manner in which its constituent tones are produced and thus points to their onset time, 'decay' is a longer-lasting phenomenon under which their frequency components become increasingly weaker until they cease to be perceived at all. The phenomenon of decay consists in variations in the amplitude and intensity of partials over time. This generates a certain fading effect on the perception of complex tones and, consequently, that of chords.

[29] Beament, J., *How We Hear Music*, Rochester, NY: Boydell Press, 2001, p. 41.

Most importantly, as Brian Kane notes, since the 'lifespan' of these frequency components can be quantified under a Fourier analysis[30] of the relevant waveform, "each partial in the spectrum can be tracked over the course of the sound's existence".[31] As seen in Chapter Four, the auditory system's ability to track, select and group different components of a chord event at their onset and over a period of time constitutes the basis for the perceptual experience of chord tokens within a temporal window. Hence, the earlier notion of a temporal Gestalt is equally suited to describing the coalescence of complex tones into chord tokens, given the distinctively temporal nature of chord events.

It is in view of the coalescent nature of chord tokens as temporal Gestalts – and within their respective temporal windows – that we can best account for their persistence over time. In Chapter Two, I argued that sounds *qua* auditory objects are perduring entities; in other words, they are sound streams that comprise different sounds as temporal parts. Once again, the case of chords could not be more fitting an example: a chord token is a temporally extended stream, perceived within a given temporal window that is determined by simultaneous onset times – i.e. the concomitant attack of its components – and one which persists as a single entity despite qualitative changes to it – i.e. the decay effects described above. The duration of the chord token can thus be framed between the initial attack and the cessation of audible sound that follows decay. Furthermore, given the numerous changes to the intensity of the frequency components and their varying decay patterns, it is not only feasible in principle but empirically justifiable to think of these variations over time as constituting a number of temporal parts of the auditory object that is the chord token (i.e. stream). As I argued in Chapter Two, it is the latter that we hear, once the auditory system selects, groups and also adds new components to the coalescing sound that we perceive within a given temporal window.

Hence, the notions of temporal Gestalt and temporal window are essential to my account of the phenomenology of chord tokens, and they warrant conceiving of chords from a perdurantist perspective. However, Scruton takes the experience of harmony to be primarily of a spatial Gestalt, rather than a temporal one. Specifically, in his characterisation of harmonic organisation, he construes it as dependent on spatial metaphors – e.g. chords are described as "spaced, open, filled, or hollow" – and metaphors of movement, whereby each is a unity that "crosses distances, and which can be grasped all at once".[32]

---

[30] For a recapitulation of the Fourier analysis, please refer to Chapter Three (Section 3).
[31] Kane, B., 'The Elusive "Elementary Atom of Music"', *Qui Parle*, 14, 2, 2004, p. 132.
[32] Scruton, 1997, p. 71.

Granted, while I have previously endorsed some elements of Scruton's account of the musical experience and that of chords as *Gestalten*, I do not accept his argument that certain linguistic descriptions of aesthetic effects warrant the view that chords are to be taken as spatial Gestalten. Were that to be case, the claim that a chord constitutes 'a unity that is grasped all at once' is more suitable to conceiving of it as a temporal Gestalt instead – which is not a metaphorical description but, rather, points to the coalescent nature of chords, as I have argued above. Yet, unlike Scruton's view, I do not intend my account to hinge upon linguistic descriptions – and even less so on figurative ones at that.

That said, the notion of spatiality is better deployed when considering the vertical aspect created by the experience of pitch height. As seen earlier, the pitch continuum is perceived in terms of distances, which are divided into discrete intervals. In particular, it is the concept of intervals that, I want to argue, unlocks the understanding of chords types/subtypes as the blueprint for the domain of tonal harmony. Although the coalescent nature of a chord token is that of a temporal Gestalt, intervals are not of a temporal order. Rather, they are distances in pitch, or pitch *lengths*, which section tonal space horizontally – in the case of melodic intervals – or vertically – in the case of harmonic ones. It is, therefore, in terms of vertical intervals that chord types are to be understood, since it is intervals that are the most fundamental elements of chords *qua* musical individuals.

3. Chord-type identity

While my account of chord tokens is closely related to my proposal that chords are perceived as auditory objects and experienced as temporal Gestalts, the same cannot be said of that of chord types. Rather than frequency components or tones, it is intervals that are the constitutive elements of chord types. In order to substantiate this claim, I aim to give an account of the identity of chord types as determined by intervallic relationships. Specifically, I will show that chord patterns may be decomposed into interrelated yet discrete intervals, which are themselves patterns that are instantiated within chord tokens. Hence, I will maintain that the type/token distinction is able to account for the multiple instantiation not only of chords but also of the intervals which their individual patterns comprise. Although intervals are essentially relationships between tones, I will defend the view that the part-whole relation between patterns and intervals warrants conceiving of chord types in terms of mereological composition. Most importantly, I will argue that chord types must be able to be instantiated if they are to be treated as *musical* individuals.

In the preceding chapters, I characterised the nature of chord events in terms of the mereological amalgamation of component frequencies, and that of chords *qua* auditory objects as Gestalt-based wholes that are not reducible to a mere aggregate of complex tones as their parts. But, rather than component frequencies or complex tones, the constitutive elements of chords *qua* musical types are intervals, instead – a view which is not unfamiliar to the music-theoretical literature. As Paul Hindemith put it, a chord is "by no means an agglomeration of intervals. It is a new unit which, although dependent on the formative power of the single interval, is felt as being self-existent and as giving to the constituent intervals meanings and functions which they otherwise would not have".[33] Similarly, in his outline of Elliott Carter's rationale behind the composition of his *Double Concerto*, Kane notes that he used a different iteration of the atomic metaphor outlined earlier:[34] "intervals are 'atomic' in the sense that they have a clear identity and can be combined with one another to produce chords, harmonic fields, spectra, and melodies".[35]

In tandem with Hindemith and Carter, I want to argue that intervals are not only fundamental to pitch relationships in chord tokens but also essential components which confer structure to chord types. In the former case, it is in virtue of the perceived distances between pitched sounds on the frequency continuum that we are able to distinguish the pitch heights to which we assign pitch classes such as C, E, G, for example – as well as all other remaining pitch classes in any given octave. Indeed, as shall be outlined in Chapter Ten, the manifold historical attempts to tune fixed-string instruments to a particular standard were aimed at preserving the 'purity' of certain intervals as much as possible. In particular, they have tended to favour the interval of the octave – at the ratio of 2:1 – which is the harmonic interval heard when two tones of equal pitch class are simultaneously sounded (e.g. C1–C2, or C2–C3, etc.).

Yet, over and above the significance of intervals for understanding how we experience distances on the pitch continuum, they are equally essential to establishing the identity of chord types. As stated earlier, the tonal spectrum may be sectioned both horizontally and vertically, such that the nature of intervals is determined by a criterion of temporality. In other words, while horizontal distances represent tones that are sounded successively and thus express melodic intervals, vertical distances indicate simultaneously occurring tones which constitute harmonic intervals, instead. Specifically, it is the latter that are taken to be the most fundamental 'building blocks' of tonal harmony.

---

[33] As quoted by Scruton. Ibid., p. 65.
[34] Please refer to Chapter Three (Section 2) for Scruton's view of tones as the atoms of musical ontology.
[35] Kane, p. 120.

Hence, it is in virtue of harmonic intervals that the very structure of a chord pattern is determined. Between any two given tones, there may be the simple interval of a second, a third, a fourth, and so on, but they may also straddle over an octave and form compound intervals such as ninths, elevenths, etc. For example, C–E and F–A are intervals of a third – or simply 'thirds'; E–B is a fifth, G–F a seventh, and the interval between C and the D of the subsequent octave is a ninth. The ordinal number associated with the interval thus represents the distance between any two given tones on the pitch spectrum.

In addition to an ordinal number, harmonic intervals are also identified by a modifier, which represents a certain property that is ascribed to a given interval depending on their extent. This property is specifiable by considering the number of semitones between any two given notes.[36] For instance, C–E is a *major* third, but if E is 'flattened' – i.e. made one semitone smaller – the new interval C–E♭ is a *minor* third. Other intervals have traditionally been described as *perfect*, as is the case of certain fourths (e.g. C–F) and fifths (e.g. A–E). Intervals may also be *augmented* or *diminished*. While augmented intervals result from extending a perfect or major interval by one semitone (e.g. C–F♯, F–B), in diminished ones a perfect or a minor interval has been made shorter by one semitone (e.g. C–G♭, B–D♭). It is therefore of intervals such as these that chord patterns are composed; put simply, their structure is determined by both the number of intervals present in them and the type of interval at stake.

Most importantly, intervals such as these may be instantiated in different chord types. Hence, just as I have adopted the type/token distinction to account for the multiple instantiation of chords, I will similarly take intervals to be best understood in terms of the type/token theory. For instance, intervals such as a minor third, a perfect fifth and a diminished seventh – amongst others – are all types. A minor third may be instantiated by the interval C–E♭ but also as D–F or E–G, just as a perfect fifth may be instantiated either by F–C or G–D and a diminished seventh by either A♯–G♭ or E–D, so these are best treated as tokens of their respective interval types. Indeed, it is the specific patterns of interval types – along with their octave-indexed pitch classes and the specific order in which they appear (i.e. whether in root position[37] or inversions thereof) – that confer chord subtypes their internal structure and their respective tokens a distinctive sound quality.

---

[36] Instead of 'tone', I use the term 'note' here to preclude any possible ambiguity, since 'tone' may also refer to the interval equivalent to two semitones.

[37] As indicated in footnote no. 14 on page 60, the constituent tones of chords in root position may be found in more than one particular order – provided that the fundamental remains at the lowest position. Yet, I will designate the order in which they most commonly appear (i.e. C–E–G for a C Major triad) as their 'typical root position'.

For instance, the typical structure of the G Major triad in root position is G–B–D, and that is the case in any given octave (e.g. G3–B3–D3). Similarly, its first and second inversions – i.e. B–D–G and D–G–B, respectively – have also fixed structures, regardless of the octave in which they are found. It is in this sense that, I believe, we can speak of chord patterns as the "modally and temporally inflexible" entities that Dodd takes types to be.[38] In other words, the properties of chord patterns are necessarily intrinsic to them in virtue of the intervallic relationships they comprise, and there can be no change to these properties through time. However, contrary to Dodd's claim that types are structureless, I take the view that it is precisely the structural nature of the intervallic relationships within a given chord pattern that enables us to conceive of it as both modally fixed and atemporal. This is because, in any possible tonal world and at any given time, the second inversion of the G Major triad (as above) necessarily consists in the D–G–B pattern, all else being equal.

Hence, it is in light of these remarks that (i) and (ii) above must be understood. To be clear, chord patterns conform to a certain structure, which is determined by the intervallic relationships between certain pitch classes. The nature of those relationships entails a particular arrangement of pitch classes that renders chord patterns modally fixed and atemporal types. Consequently, it is the resulting structure that determines the identity of a given chord type, such as that of a major triad or that of a major-minor seventh. Carter's atomic metaphor mentioned earlier thus offers a most pertinent analogy: intervals are the atoms in the ontology of chords *qua types* of musical individuals; they have a distinct identity and can be combined with one another to instantiate chord tokens. Analogously, chord patterns are like molecules, i.e. whole structures comprising tones in particular intervallic relationships, much in the same manner as several atoms are joined within a given molecule by covalent bonds. However, this once again poses a question regarding the mereology of chords, albeit this time from a music-theoretical perspective.

Harmonic intervals are, as per the earlier definition, vertical distances between any two concomitant musical tones. In turn, chord types comprise two or more harmonic intervals, such that the component tones of a chord become inexorably interrelated. This can be clearly seen in triads, which are the simplest expressions of chord types. Consider, once again, the G Major triad in root position. As with any other major triad, it consists of 'stacking' a minor third over a major third: G–B is a major third, and B–D is a minor third. In light of this example, two important observations must be made at this point, which have a bearing on the role of intervallic relationships in the mereology of chord types.

[38] Dodd, p. 37.

First, when harmonic intervals are superimposed on to one another to form chord patterns, the pitch classes that constitute them – such as G, B and D above – come to perform more than one intervallic function. The most obvious example concerns the middle note (B), which is not only in an intervallic relationship to G – in forming a major third – but is equally related to D – as part of the minor third interval (B–D). In that sense, if discrete intervals are, as it were, the atoms of chord patterns, then it is the intervallic *relationships* between pitch classes that enable us to conceive of chord types as akin to molecules. However, the second observation points to a departure from the molecular metaphor. For a given triad in a typical root position, the combination of two thirds forms yet another interval, namely, that of a fifth (the interval between G and D, in the example above). This is the case with all triad types in a typical root position, regardless of the way thirds are superimposed. In other words, any given triad type in a typical root position may be described as a pattern consisting of a fifth built of two superposed thirds. Within it, each pitch class is thus an element that necessarily appears in two intervallic relationships.[39]

It is in light of these observations that I want to consider the mereology of chord types. Briefly, although the constitutive elements of chord types are intervallic relationships, the part-whole relation between intervals and patterns nonetheless warrants conceiving of them in terms of mereological composition. This is for three main reasons. First, chord patterns may be decomposed into different parts (i.e. specific intervals) which together form a single whole, but one where the whole is still more than just a sum of its constituent parts. Rather than a mere grouping of intervals, a chord pattern is an abstract entity in its own right, bearing properties that are not shared by its parts when taken individually.

Secondly, unlike the case of chords *qua* auditory objects – in which both physical and perceptual elements are present – chord patterns observe what Peter Simons has described as one of the strongest intuitions about mereological composition; namely, that "composition is always intracategorial".[40] Under this general rule, composite objects cannot result from straddling the divide across different categories, such as the intermingling of concrete individuals and abstract entities, or of events and substances, for example. Unlike these, chord structures satisfactorily conform to the intracategorial requirement, in that they are built of intervallic relationships between pitch classes representing different complex tones, and their patterns are solely determined by the intervals they comprise.

---

[39] The number of intervallic relationships concerning each constituent tone depends upon the number of tones comprised within a given chord type. For instance, each pitch class within a tetrad is necessarily found in three different intervallic relationships with the other components.

[40] Simons, P., 'Real Wholes, Real Parts: Mereology without Algebra', *The Journal of Philosophy*, 103, 12, 2006, p. 605.

Thirdly, since each pitch class is a common element to at least two intervallic relationships – on the basis that three-tone chords are the simplest of all chord patterns – each of them necessarily appears interrelated in at least two particular intervals. Hence, the chord structure is held together by the intervallic relationships between its constituent tones. Specifically, I want to describe the constituent tones as 'welded' together in virtue of these intervallic relationships. It is in view of this 'welding relation' – to use Simons's terminology – that I want to consider the mereological composition of chord types. In his example of a symphony, Simons states that any of its performances is a whole that may be segmented into several different parts that are welded together as the symphony.[41] Analogously, I take chord types to be whole entities comprising pitch classes in specific intervallic relationships, which is precisely what welds their corresponding complex tones together towards forging the chord patterns that we take to be distinctive musical entities in their own right. Were that not to be the case, they would fail to constitute the chord types that composers and musicians alike routinely encounter in their practice.

An important consequence of the mereology of chord types – unlike that of chord events – is that, since patterns are modally fixed and atemporal entities, chord types may be given a formal expression. To that extent, my account here dovetails with Ingolf Max's proposal of a logical system of chords based on their internal structure. Echoing the line of argument I have pursued above, Max takes intervals to be the starting point of a logical theory of harmony, rather than tones/notes. "Chords", he writes, "are not pure sets of tones or notes ... A chord itself is a 'fusion' of at least two intervals which leads necessarily to at least a third interval within this chord".[42] To that extent, his main thesis is that the logic of chords must be understood with reference to matrices, whereby a chord matrix consists in the pattern of all intervallic lengths within a given chord. Underscored by a formal analysis of the interrelations betweens intervals within chord matrices, Max argues that chords are well-structured invariant entities with an inner logical form, but ones that can only be grasped as such within the logical space provided by the chromatic scale.[43]

While it is beyond the scope of this chapter to consider the broader metaphysical implications of Max's logic of chords, two key contributions from his account are particularly relevant as they are aligned with the analysis I have conducted in this section. First, Max formalises the inner structure of chords *qua* individual patterns; in logical terms, he states, a chord is "neither a sequence of three or more tones, nor a sequence of a tone

[41] Ibid., p. 611.
[42] Max, I., 'A Molecular Logic of Chords and Their Internal Harmony', *Logica Universalis*, 12, 2018, p. 240.
[43] Ibid., p. 240.

and an interval sound in any order, nor a sequence of two or more interval sounds. Each chord is context-freely identifiable by its characteristic *inner structure* ... [such that] the whole structure is the pattern of one sound".[44] Second, echoing the atomic metaphor, Max's molecular logic of chords provides the basis for individuating chord patterns in terms of matrices of their interval lengths, rather than tone components. These claims not only conform to (i) and (ii), as identified above, but may be taken to endorse my attempt to establish the identity of chord patterns as types.

The defining aspect of chord-type identity is encapsulated by Max's postulate that the whole structure of a given chord is the pattern of one sound. In speaking of a molecular logic of chords, Max implicitly accepts that chords can be treated as abstract, composite entities; yet, he is equally concerned with placing them within the domain of tonality. But the broader questions regarding which chord matrices turn out to be formed – and whether they are phenomenally identifiable or even audible at all – do not seem to be a concern under Max's logical system. "The chromatic scale", he writes, "is to be thought as open in both directions and, therefore, infinite. Independently from our hearing capacities we have – from our logical point of view – an infinite number of tones, intervals and chord patterns".[45] On this view, the identity and individuation of chord types do not depend upon whether they may ever be instantiated or reidentified as tokens at all. In other words, Max is solely interested in providing a formal account of any possible interval combination within a matrix, with little care for the musical context in which they may occur – or whether or not they may occur at all. Put simply, Max treats 'chord' as a purely formal concept, as he himself acknowledges it.

Granted, the numerous possibilities of intervallic combinations in a given matrix indicate that the identity of chord types is primarily determined by the intervals that constitute them, whether or not they may be capable of instantiation. For instance, the famous Tristan chord, originally consisting of the intervals between F, B, D♯ and G♯, has framed this chord pattern in terms of a combination of an augmented fourth, an augmented sixth and an augmented ninth above a bass note. Sounding this chord will thus generate an instance of the Tristan chord; but this chord has not always had an instantiation, and it is conceivable that it never would have had to be exemplified in order to be classed as a musical entity. In short, a logical pattern of interval combination is sufficient for a chord type to exist.

---

[44] Ibid., p. 242.
[45] Ibid., p. 241.

That said, it would not seem very satisfying to speak of chords *qua* musical individuals that may never become an audible particular. It is for that reason that my defence of the type/token theory to account for the multiple instantiation of chords hinges upon my earlier characterisation of chord events and that of chords *qua* auditory objects, as outlined in earlier sections of this chapter. Given that the musical experience is, by definition, phenomenal, to conceive of chords in terms of infinitely possible logical constructs that are independent from our hearing ability – as Max puts it – seems tantamount to conceiving of possible musical works in some obscure notation that precludes them from ever being instantiated, so they might never be *heard* as such.

As seen in Section 1, Scruton has a related concern, when he seeks to reconcile the identity of musical sounds as unrepeatable (pure) events with the multiple instantiation of tones. However, given his earlier distinction between the acoustical and musical dimensions of experience, Scruton warns against an acoustic-based criterion of individuation that fails to capture the identity of a musical individual within the 'phenomenal world' of tones. To the extent that his acousmatic thesis warrants the differentiation between musical and non-musical sounds, Scruton's concern seems valid; but his reticence is unjustified with regard to the specific case of chord types, given their inner structure – to which Max also alludes. Hence, while Scruton rejects an ontology of musical sounds that ties them to their acoustic base, Max offers an account of the logic of chords that is not in the least concerned with the possibility of instantiation. Despite the merits of both views, it seems that a third way is not only desirable but also possible – which is what has been attempted in this chapter under my account of the identity of chord types and the phenomenology of chord tokens.

Granted, chords are types of musical individuals that may be instantiated at multiple times. Yet, chords are also the auditory objects that we hear when three or more tones are sounded simultaneously, whether they are heard as isolated objects or within a specific musical context that gives them distinctive musical properties. In them we hear more than just the audible counterparts of the acoustic components of chord events; indeed, the auditory system is the ultimate judge of what we hear, with the cochlea as its *de facto* sound analyser – and enhancer, too. The chord percept – which is how we hear those sound events that are grounded by complex waves – is a temporally extended stream, captured within a given temporal window, and experienced as a temporal Gestalt. Chords are thus the perfect example to illustrate why sounds may be conceived both as events and objects, and in terms of waves. But they are nonetheless best appreciated for their musical properties, as the fundamental elements of tonal harmony that they most certainly are.

The main purpose of the first part of this study has been to give an account of the ontology of chords. Given that chords are sounds, in the first two chapters I examined the ontology of sound, where I sought to defend a proposal that reconciles different strands of the philosophical debate. Most importantly, my thesis of sonic reconciliation is what has made my account of the ontology of chords possible in the subsequent chapters. After accounting for the nature of chords *qua* sound events, I argued that chords must also be conceived as auditory objects in their own right. To that extent, my analysis of the case of chords not only conforms to my thesis of sonic reconciliation but should also serve as an illuminating case study. In addition, given that chords are taken to be the musical entities that are the basis of tonal harmony, the present chapter carries particular importance in meeting the first aim of this study fully. Having completed my analysis of the ontology of chords, I subsequently intend to give an account of the aesthetics of consonance and dissonance – two properties commonly ascribed to chords that are essential for understanding the metaphysics of tonal harmony.

Part II

Aesthetics

Introduction

Some chords are taken to be agreeable sounds, whilst others are said to be jarring to the ears. Even relatively inexperienced listeners are able to qualify their experience of chord tokens in aesthetic terms. Just as they may enjoy hearing a progression of chords played in the background, they may suddenly wince at a 'rough-sounding' chord. It is, therefore, an empirical fact that chords are objects of aesthetic appreciation. Although they may be described under a range of aesthetic terms, the overarching question for the aesthetics of tonal harmony seems to revolve around how consonant or dissonant chords are perceived to be. It is with these aesthetic properties that the second part of my thesis is concerned.

Accordingly, my account begins with what I will characterise as the 'problem of consonance'. This concerns not only conceptual difficulties but a broader metaphysical question that is the task of subsequent chapters to address. Amidst the disputes that have historically beset the aesthetics of consonance and dissonance, it seems that an enquiry into their nature *qua* properties is all too often neglected. In order to reverse this, my proposal will draw on an existing conceptual framework in order to account for what I will describe as the dual phenomenology of consonance (and dissonance), as pertaining to two different domains: that of 'sensory consonance/dissonance' (i.e. the psychoacoustic component) and that of 'harmony' (i.e. the distinctively musical experience of chords). Most importantly, whilst acknowledging the different ways in which consonance and dissonance may be experienced, I will show that a reconciliation between the psychoacoustic component and the aesthetic experience of it in a musical context is not only possible but desirable.

Having previously proposed an account of the ontology of chords that encompasses their psychophysical attributes as well as the phenomenological aspects of chord perception, my strategy in what follows is not only similar but fundamentally dependent upon the analysis of chords that I offered in Part I. Hence, after outlining the problem of consonance in Chapter Six, I will examine the nature of the acoustic base underlying the perception of consonance and dissonance in Chapter Seven. Subsequently, I shall turn to an assessment of the aesthetics of chord perception both as isolated chord tokens and within a distinctively musical context in Chapter Eight. As a corollary of my arguments in these chapters, I will provide a unifying treatment of the metaphysics of tonal harmony in Chapter Nine. This will be followed by a consideration of the instability of chord tokens *vis-à-vis* some important tuning and temperament differences in Chapter Ten – which have historically proven to be more intractable than the problem of consonance itself.

Chapter Six: The Problem of Consonance

There is a long-standing problem concerning the nature of consonance and what it is for a chord to be consonant. Beyond specific historic-cultural and musicological contexts, most theoretical efforts to account for the perception of consonance – and that of dissonance[1] – have largely been focused on conceptual issues arising from different attempts at understanding those phenomena. Perhaps unsurprisingly, this has led some theorists to interpret the question of consonance as resulting from semantic ambiguity. Yet, despite some conceptual difficulties and variations in the ways that consonance and dissonance came to be understood, I am going to argue in this chapter that the problem is of a metaphysical nature, instead. To that end, I will characterise it as a consequence of the inability to distinguish between the two ways in which the properties of consonance and dissonance may be experienced – the strictly sensory, on the one hand, and the distinctively musical, on the other – and how they may be interrelated. Hence, I will argue that the aesthetics of chords must be understood in terms of a 'dual phenomenology', whilst maintaining that the two components may be reconciled under the solution I propose here.

1. Statement of the problem

The metaphysics of consonance has been overshadowed by some broader conceptual issues that have detracted from two central questions around its nature *qua* property. These questions constitute the two horns of what I will refer to as the 'problem of consonance'. The first and most fundamental one concerns what kind of property consonance is: on the one hand, consonance is treated primarily as a psychoacoustic property; on the other hand, it is taken to be a musical property. The second question, in turn, concerns the relationship between those putative properties and the grounding base underlying chord perception, as characterised in Chapter Three. Taken together, these two questions constitute the two horns of a problem which, I will argue, is of a metaphysical nature. This is because the problem of consonance concerns the very nature of consonance *qua* property, and that of its relationship to the psychoacoustic base of chords.

---

[1] Although this chapter will most frequently include specific references to 'consonance', in most cases they may equally be said of 'dissonance'. For ease of exposition, I shall describe the problem in terms of consonance by default, but will occasionally reiterate that those references are also applicable to dissonance.

The problem of consonance is at the heart of what has come to be seen as an almost intractable conflict between psychoacoustics and music. On the one hand, there is the strictly sensory view under which consonance must be understood as a psychoacoustic property of certain isolated chord tokens. On the other hand, there is a context-based approach to tonal harmony, under which consonance is not a psychoacoustic property but an aesthetic attribute more akin to secondary qualities, instead. Correspondingly, the first horn of the problem may be stated thus: does 'consonance' refer to two distinct properties, or is consonance a single property that is experienced in two different ways? These questions are, in turn, the basis for the second horn of the problem, which concerns the relationship between consonance and the grounding base of chords. Under the psychoacoustic view, consonance is taken to be closely related to – if not determined by – the physical base of chords. Conversely, from a distinctively musical perspective, consonance is understood in terms of context-based factors and treated as virtually independent from the physical properties at work in sound production.

Whether one takes consonance to consist in two distinct properties or, as I shall argue, as a single property with a dual phenomenology, the problem is not simply solved by justifying either of those two views solely on a conceptual/semantic level. Rather, the question of the relationship between consonance and the psychoacoustic base, I want to argue, must also be addressed. To that end, after outlining the specificities of the debate around consonance/dissonance in Section 2, I will adopt a twofold conception that preserves the distinctiveness of the aesthetic experience of chords in the psychoacoustic and musical domains whilst providing a unifying treatment of the two. The solution I propose in Section 3 is nonetheless but a framework to be substantiated in the forthcoming chapters, as part of my account of the metaphysics of tonal harmony.

2. The consonance/dissonance debate

In the previous section, I indicated that the problem of consonance is of a metaphysical nature, rather than merely a conceptual one. The purpose of the present section is thus to consider why the latter view has come to dominate the music-theoretical debate around the nature of consonance – as well as that of dissonance. Notwithstanding the conceptual issues arising from musicological developments along the history of Western tonal harmony, I will maintain that the nature of the problem is metaphysical in essence, in that it concerns what kind of properties consonance and dissonance are.

The debate around the nature of consonance has often been characterised under a series of conceptual difficulties, starting from the very attempt at a definition. Beyond the claim that some chords are agreeable, 'consonance' is arguably one of the most equivocal terms in the aesthetics of music. For instance, in his entry on 'Acoustics' for the *Encyclopedia of Aesthetics*, Alexander Rehding identifies three different ways in which consonance (and dissonance) may be understood. First, the acoustic approach, which has as its main focus the intervallic relationships that have traditionally been favoured in the context of Western tonal harmony – despite historical variations within that same tradition. Second, and closely related to the previous one, Rehding points to the physiological/psychological understanding of harmony. This approach is aimed at identifying potential explanations for variations in consonance perception *vis-à-vis* the auditory sensation of roughness that is associated with dissonance. Accounting for those variations is equally a desideratum in the third approach, which is contextualist in character. On this view, consonance and dissonance are not described in relation to the psychoacoustic base of chords but under a context-dependent approach based on the notions of tension and relaxation, respectively. These three differing characterisations, Rehding points out, stem from the fact that there is fundamentally no universally agreed-upon definition of 'consonance'.[2]

One way to assess the problem of consonance is by considering how these different approaches have emerged as a result of certain musicological developments in Western tonal harmony. One such attempt at grappling with those conceptual differences whilst finding a 'common thread' has been made by James Tenney in his historical survey of the concepts of consonance and dissonance. In his view, the problem of consonance is first and foremost a semantic one, and one that has resulted from specific and complex historical variations in the understanding of harmony. Tenney places particular importance on addressing this problem as a precondition for accounting for the nature of both consonance and dissonance – without which any theoretical effort to do so would be, under his view, doomed to failure from the outset. Hence, his work is focused on the historical evolution of the 'consonance/dissonance-concept', as he calls it. Specifically, Tenney maintains that it is only by considering the five different historical iterations of this concept, going as far back as Antiquity, that it becomes possible to bring clarity to the "tangled network of meanings and interpretations which so confuse the issue today".[3]

---

[2] Rehding, A., 'Acoustics', In Kelly, M. (ed.), *Encyclopedia of Aesthetics*, OUP, 2014, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199747108.001.0001/acref-9780199747108-e-6.

[3] Tenney, J., *A History of 'Consonance' and 'Dissonance'*, New York, NY: Excelsior Music Publishing, 1988, p. 3.

As a preliminary observation to that proposal, Tenney establishes two important distinctions that have a bearing on the metaphysics of consonance. First, he points to the difference between aesthetic attitudes, explanatory theories and practical uses of consonance and dissonance, on the one hand, and conceptions of the terms ‘consonance’ and ‘dissonance’, on the other. As he indicates in his work, it is specifically with those conceptions that his survey is concerned. Second, and most importantly, he distinguishes between what he calls the two ‘grammatical senses’ in which ‘consonance’ and ‘dissonance’ are used, namely, the qualitive and the entitive senses. While he takes ‘qualitive’ to refer to “the property, attribute, or *quality* associated with a sound or aggregate of sounds”, the ‘entitive’ sense would point to “the sound or aggregate itself which manifests that quality”.[4]

Tenney notes that the qualitive and entitive senses may sometimes be discerned by the context in which the terms ‘consonance’ and ‘dissonance’ are used, or under what he calls ‘certain grammatical markers’ (such as the pluralised terms ‘consonances’ and ‘dissonances’, when they are meant in an entitive sense).[5] However, he stops short of providing an account that substantiates that distinction further. Rather, as a consequence of his attempt at identifying successive variations in the understanding of ‘consonance’ and ‘dissonance’, Tenney’s account eventually reduces the problem of consonance to a matter of semantics, instead. Specifically, he takes the historical variations of the consonance/dissonance-concept to be part of a broader semantic problem resulting from conceptual issues, some of which would seem to have emerged from an inability to differentiate between the two senses. In other words, those issues would arguably be the consequence of a failure to establish a conception of consonance that specifies how it may be ascribed to certain chords, or whether they may have it essentially.

Conversely, I want to argue that the conceptual difficulties which he identifies are not the source of the problem. In proposing the distinction between the two senses, Tenney seems to suggest that consonance is some form of dispositional property that certain auditory objects have either essentially or not. Yet, while he specifies that his consonance/dissonance concept is solely intended as a tool to keep track of the historical variations in the usage of the two terms, the problem of consonance seems to become subsumed under the ‘tangled web’ of semantic variances which his survey is aimed at tracking. By stating the problem as stemming from grammatical usage, Tenney prioritises conceptual issues over what underlies them, thus losing sight of what I take to be a more fundamental question at its core over the nature of consonance as a property.

[4] Ibid., pp. 4-5.
[5] Ibid., p. 5.

To be clear, what the distinction between qualitive and entitive senses seems to reveal is that the problem of consonance goes beyond the conceptual dimension. Indeed, it should be interpreted as indicative that a metaphysical question precedes the semantic problem. On the one hand, both consonance and dissonance are taken to be properties ascribed to chords, which would seem to match the qualitive sense under which their corresponding concepts are used. On the other hand, 'consonance' and 'dissonance' may refer to certain chord types that are inherently consonant or dissonant – which came to be known as 'concords' and 'discords', respectively. In this case, rather than just 'qualifying' certain sounds, these conceptions of 'consonance' and 'dissonance' seem to point to an *a priori* essence of chord types. This is arguably what the entitive sense is meant to denote – i.e. as referring to a sound that *itself* manifests either of those qualities as its very essence.

Hence, unlike Tenney, I hold the view that the problem of consonance is fundamentally a metaphysical one. Briefly, it concerns the nature of consonance *qua* property of some harmonic intervals (i.e. dyads) and chords, and whether it is an essential psychoacoustic property they have *qua* auditory objects or a context-dependent aesthetic attribute they come to possess *qua* musical individuals. In the former case, 'consonance' may be understood in an entitive sense, whereas the latter points to a qualitive sense, instead. Most importantly, these two different ways of conceiving of the nature of consonance – as well as dissonance – have epistemological consequences for how we understand them in relation to the psychoacoustic base underlying chord perception.

As indicated earlier, the relationship between consonance/dissonance perception and the psychoacoustic base is not only central to the problem of consonance but is at the heart of what seems to be a conflict between psychoacoustics and music. On the one hand, psychoacoustic research is aimed at investigating to what extent consonance perception may be explained objectively, on the basis of the acoustic nature of dyads and chords. As such, it consists in an empirical assessment as to whether we can truthfully speak of certain sounds as concords or discords – i.e. in an entitive sense. If such is the case, it is then aimed at identifying the psychophysical causes of consonance perception and establishing some parameters for predicting the corresponding aesthetic responses and judgements.

On the other hand, in a distinctively musical context, 'consonance' and 'dissonance' are taken to be context-dependent properties – e.g. based on the experiences of tension and repose. On this view, both consonance and dissonance are then seen as properties that can only be ascribed to harmonic intervals and chords in the qualitive sense, i.e. as and when they acquire those properties in relation to other sounds in the musical context in which they occur. For instance, a concord may sound discordant in a given harmonic progression, while a dissonant chord may be resolved by another (less) dissonant chord, instead. Hence, the question of objectivity in consonance and dissonance perception is replaced by that of the relationship between the two within a musical context.

This disparity between the music-theoretical and psychoacoustic approaches reveals the extent of the problem, in that it goes beyond the conceptual level on to the question around the ontology of consonance *qua* property – which, I believe, is at the very root of it. Briefly, the terms 'consonance' and 'dissonance' are used to refer *both* to psychoacoustic properties, on the one hand, *and* musical properties, on the other. Given the complexities of the historical developments around those two concepts, it is perhaps unsurprising that conceptual issues have resulted from different attempts at characterising them as such. While they may at times designate the acoustic properties associated with concords and discords, they may otherwise point to a purely aesthetic quality ascribed to harmonic intervals and chords in relation to other neighbouring musical sounds in a given context.

Yet, the tension between music and psychoacoustics cannot simply be reduced to those conceptual difficulties. Were this to be the case, the problem could be most straightforwardly solved by establishing a distinction between those two kinds of concepts, thus differentiating between psychoacoustic consonance/dissonance and a distinctively musical conception of them. But the problem is that these two conceptual approaches to consonance and dissonance perception are not strictly independent from one another. Rather, the main difficulty concerns what has historically come to be known as traditional tonal harmony, which has been conceived as such from the contrapuntal and figured-bass period (ca. 1300-1700).[6] As Tenney points out, one of the key developments in that period is the coexistence of pre-existing acoustic criteria for consonance and dissonance classification, on the one hand, and changes to polyphonic standards introduced by newly developed rules of counterpoint,[7] on the other, such that 'consonance' and 'dissonance' began to be understood in more functional terms.[8]

---

[6] Ibid., p. 39. I am following Tenney's chronology here, which may comprise a longer historical period than other theorists may give it when characterising the advent and systematisation of traditional tonal harmony.
[7] As indicated earlier, 'counterpoint' is frequently used interchangeably with 'polyphony', where different voices (i.e. melodic lines) may be sounded concomitantly whilst preserving a relative perceptual independence from one another. In the sense deployed here, however, 'counterpoint' specifically refers to the system of rules that underlies the organisation of those simultaneous voices. Kennedy, J., Kennedy, M., Rutherford-Johnson, T. (eds.), 'Counterpoint', *The Oxford Dictionary of Music*, 2013, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199578108.001.0001/acref-9780199578108-e-2218. For further detail on counterpoint, please refer to footnote no. 32 on page 77.
[8] Tenney's usage of 'functional' here is not to be interpreted in light of the more modern notion of a 'functional harmony', which lies beyond the scope of the present study.

This particular development in the understanding of tonal harmony can illuminate the problem of consonance *vis-à-vis* the relationship between psychoacoustics and music – a thesis which I will endeavour to substantiate more fully over the ensuing chapters. At this stage, it suffices to point out that, while the aforementioned criteria for consonance/dissonance classification stem from an understanding of consonance as a psychoacoustic property, under the rules of counterpoint they came to be considered from a more operational perspective, based on the relationships between harmonic intervals in the syntax of polyphony. To that extent, they veer towards a treatment of consonance and dissonance as aesthetic properties of musical individuals instead, as described above.

Yet, as Tenney notes, the developments in tonal harmony during that historical period revolved around a "precise, one-to-one correspondence between the rules of counterpoint and the consonance/dissonance categories referred to by those rules".[9] This observation leads to two important questions concerning the metaphysics of consonance. If a clear correspondence between consonance *qua* psychoacoustic property and consonance *qua* aesthetic property is empirically observed, then a question arises as to whether they should be conceived as distinct properties. If not, then another question may be asked as to whether it is possible to provide a unifying treatment of consonance whilst respecting any phenomenological differences between the two domains. Indeed, it is this very possibility of a correlation between psychoacoustics and music that, I believe, prevents them from being treated as completely independent domains. This is particularly relevant for considering what has come to be seen as the fundamentals of tonal harmony.

One of the most important contributions that are normally associated with traditional tonal harmony is that of eighteenth-century theorist Jean-Philippe Rameau. Rameau wrote extensively on what he took to be the fundamental principles of composition – especially those concerning harmony – which came to predominate in Western music until the late nineteenth century. Specifically, he identified two chord types that are, in his own words, "as it were the only Chords in Harmony": namely, the "perfect Chord" and the "Chord of the Seventh". Both of these, he maintained, "are only affected to a Progression of the Bass [...]; and if we are going to alter that Progression, we shall not thereby alter their Chords, but only the Disposition".[10] In doing so, Rameau firmly established the notion of a 'fundamental harmony' by emphasising the importance of the bass note in chords (i.e. the root) for shaping the harmonies generated by them.

---

[9] Ibid., p. 39.
[10] Rameau, J-P., *A Treatise of Music, containing the principles of composition*, Dublin: Printed for Luke White, 1779, p. 28.

The cornerstone of Rameau's theory is the perfect chord, which he describes as comprising the bass note, the third, the fifth and the octave. Since the octave is a repetition of the bass note but at a higher-pitch, the chord comprising the bass, the third and the fifth came be to described as the triad in its typical root position. Alongside of its inversions, they constitute the foundation of triadic harmony. The importance of the perfect chord in Rameau's treatment of intervallic relationships could thus not be overstated; he took the internal composition of the perfect chord (i.e. its intervals) to be so fundamental that "all their harmony is included and comprehended" within it.[11] In addition to the third and the fifth, the fourth and the sixth were also seen as consonant, in that the fourth is the interval between the fifth and the octave, and the sixth that between the third and the octave.

Just as the perfect chord was taken to be the basis of all consonance, the seventh chord – comprising the triad plus a third interval upon the fifth – is described by Rameau as "the origin of all the Discords".[12] As a dissonance, the chord of the seventh – along with its three inversions – must be prepared and resolved by a consonant interval, he states; and that should be undertaken in conformity with the rules that he sets out in his treatises. Specifically, they include principles concerning the use both of concords and discords in harmonic progressions, for example, amongst several other explanations that made his systematisation of harmony a reference for nearly two centuries. In particular, one of its most relevant principles concerns the distinctive role of the third, in that he saw it as the source and basis of the only two modes in tonal harmony. Under Rameau's theory, the difference between the major and minor modes is determined by the relationship of the third to the fundamental: "as the Third can be but either Major or Minor, or Sharp or Flat", he states, "so likewise the Mode is distinguished by those two Sorts".[13]

Furthermore, given that both triads and seventh chords are built of thirds that are superposed on to one another, the importance of that interval is markedly observed in Rameau's texts. In his famous *Treatise on Harmony* (1722), he identifies only three primary consonances, namely, the fifth, the major third and the minor third, in that "from these is constructed a chord called *natural* or *perfect*";[14] yet, he eventually grants that thirds are the most fundamental elements of all chords, in that fifths themselves consist of two superposed thirds. Most importantly, Rameau unequivocally states that "we should attribute to them all the power of harmony, if we reduce harmony to its simplest terms".[15]

[11] Ibid., p. 11.
[12] Ibid., p. 30.
[13] Ibid., p. 27.
[14] Rameau, J-P., *Treatise on Harmony* (1722), New York: Dover, 1971, p. 16.
[15] Ibid., p. 39.

As stated earlier, the fundamentals of traditional tonal harmony enshrined in Rameau's theory came to predominate in Western music until the late nineteenth century. But with the emergence of experimental psychoacoustics, the tension between music and psychoacoustics became more prominent. Specifically, some empirical studies have ever since indicated a few discrepancies between some of the core beliefs that underpin traditional tonal harmony, on the one hand, and some of their findings on the nature of consonance and dissonance perception, on the other. The problem of consonance thus rises to the fore in the clash between the two. This is made all the more complex with the changing attitudes to tonal harmony from the late nineteenth-century onwards, which paved the way for a more contextualist approach to consonance and dissonance altogether.

Hence, it is in view of the complexities arising from these developments that Rehding's three approaches to consonance and dissonance mentioned at the beginning of this section are best understood. However, his account does not seem to shed any new light on the problem of consonance itself. That said, in a more recent contribution to the consonance debate, Rehding takes a more pronounced position in favouring a subjectivist view and rejecting the acoustic approach that has predominated in psychoacoustic research to the present day.[16] This kind of suspicion around the musical value of the import from research on the psychoacoustics of consonance is, nonetheless, not an isolated phenomenon. Hermann Helmholtz's comprehensive study on musical acoustics – which will feature more prominently in the following chapter – has over time become the recipient of scathing criticism. Thus, it is perhaps unsurprising that the metaphysics of consonance has been overshadowed by this apparent conflict, which has been made even more difficult to unravel given the growing chasm between music and psychoacoustics.

Yet, contrary to what Rehding has suggested, it is the very feasibility of that earlier correspondence between psychoacoustic criteria for consonance (as well as dissonance) and the aesthetics of a distinctively musical experience of harmony that, I believe, points to a solution to the problem of consonance. Rather than separate, independent domains – and despite the popularity of the contextualist view amongst musicians and music-theoretical circles at large[17] – I want to support the view that the psychoacoustic approach and the distinctively musical experience of chords are, indeed, frequently enough intertwined in traditional tonal harmony. It is towards this resolution that I turn to in the next section.

---

[16] Rehding, A., 'Consonance and Dissonance', In Rehding, A., Rings, S. (eds.), *The Oxford Handbook of Critical Concepts in Music Theory*, New York: OUP, 2018, URL = https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780190454746.001.0001/oxfordhb-9780190454746-e-14, p. 441. Rehding's subjectivist approach will be outlined in Chapter Eight (Section 3.3).
[17] The contextualist approach will be considered in greater detail in Chapter Eight (Section 3.2).

3. Outline of the solution

So far in this chapter, I have sought to introduce what I take to be the problem of consonance (Section 1) and set it against the music-theoretical debate around it (Section 2). In this section, I want to indicate that a thesis of reconciliation between music and psychoacoustics is possible. Specifically, I will propose that the conflict between the two is only apparent, for it is mired in the 'tangled web' of concepts that Tenney has described. I will argue that it is possible to address some of those conceptual issues by tackling the underlying metaphysical questions instead, which is what enables the unifying treatment of consonance I shall offer in the subsequent chapters. To that end, I will argue that my thesis of the dual phenomenology of consonance/dissonance warrants the possibility of reconciliation, whilst preserving the distinctiveness of the sensory and musical experiences.

The framework of the solution to the problem of consonance that I want to advance in this chapter has three specific aims. First, it is targeted at establishing that consonance is best understood as a single property with a dual phenomenological nature. This is because the particular case of traditional tonal harmony points to an element of convergence of the psychoacoustic understanding and the musical view, rather than the complete dissociation advocated by those who favour a strictly context-dependent approach made popular by further developments in Western tonal harmony. Indeed, it is an empirical fact that consonance and dissonance are experienced in more than just one way, depending on whether dyads and chords are perceived as isolated auditory objects or within a specific musical context. Hence, the difference between the two cannot simply be addressed in terms of semantics (Tenney) or different approaches (Rehding), but must be accounted for in view of the distinctiveness of the sensory and of the musical experiences, respectively. It is, therefore, under the distinction between the two ways of experiencing consonance that I am going to argue that consonance itself has a dual phenomenological nature.

Secondly, and as a consequence of the first aim, I will propose an account of the metaphysics of consonance that is compatible with that thesis of dual phenomenology, whilst also considering its relationship to psychophysical factors underlying consonance perception. In that same vein, the third aim of my proposal is to show how a reconciliation between psychoacoustics and music may be made possible on those grounds, with a view to providing a unifying treatment of the metaphysics of tonal harmony. It is worth remembering that, as pointed out earlier, although I shall frequently refer solely to 'consonance', the same aims are applicable to the case of dissonance.

Granted, there is an element of tension between certain orthodox parameters of harmonic analysis in traditional tonal harmony and some of the conclusions that have been drawn from specific findings in experimental psychoacoustics.[18] On the one hand, those parameters are based on a series of systematic rules whereby certain intervals are favoured whilst others are avoided, unless these may be resolved within a given harmonic context. Although these rules may at times be overridden so as to produce a certain aesthetic effect, such rules have otherwise generally been adopted.[19] On the other hand, psychoacoustic research of late indicates that the complexity of the problem is greater than expected. While many of the assumptions underlying traditional tonal harmony may frequently be confirmed, some discrepancies have been observed between a few rules and experimental evidence, along with a certain degree of ambiguity in some cases which prevents any firm conclusions of a thoroughly established correlation between psychoacoustics and harmony from being drawn. Arguably, this would be sufficient evidence to prevent any attempt at reducing the aesthetics of consonance to nothing more than a psychoacoustic assessment.

Whether or not a one-to-one correspondence between acoustic criteria and the fundamentals of tonal harmony is observed in *all* cases of consonance and dissonance perception, the fact that such correspondence may be empirically confirmed not only for some but several cases is in itself evidence of a convergence between the psychoacoustic and the musical view. Hence, to establish an ontological distinction between two kinds of property to which 'consonance' may refer is tantamount to dismissing the empirical fact that psychoacoustic and musical aspects may converge under a single property in chord perception. This is best exemplified by considering the intervals that are the components of Rameau's perfect chord – perfect fifths being one of the prime examples of it.[20] Yet, the psychoacoustic view and the musical experience of consonance may also diverge within a distinctively musical context in which chords are perceived in relation to other neighbouring sounds. Hence, there must be a way to account for these cases of convergence and divergence that does not entail two independent notions of consonance (and dissonance) as pertaining to the sensory and musical domains, respectively. Although this is a case I shall argue more fully in Chapters Eight and Nine, the remainder of this section constitutes an outline of my overall proposal in Part II.

[18] This will be considered in greater detail in Chapter Seven as well as Chapter Eight (Section 2).
[19] I am offering here a general characterisation of the standard 'textbook version' of tonal harmony, which does not comprise any changes associated with the movement towards chromatic music and atonality from the late nineteenth-century onwards. Further detail on these musicological developments as well as the notion of 'resolution' will be considered in Chapter Eight (Section 3) as well as Chapter Nine (Section 3).
[20] Further detail on the traditional view of the consonance or dissonance of certain harmonic intervals will be considered in Chapter Eight (Section 2).

It is with a view to reconciling these differences that I will uphold a twofold conception of consonance that is not only aimed at solving the problem but is also able to address the broader question of the aesthetics of harmony from a contextualist perspective. My attempt at doing so is not entirely unprecedented; rather, my proposal is indebted to the conceptual framework originally proposed by Ernst Terhardt in the 1970s – which, in turn, echoes the one offered by Helmholtz a century earlier. Perhaps controversially, I will argue that their contributions are still the best candidates to illuminate the psychoacoustic principles that may be correlated with the fundamentals of traditional tonal harmony. Most importantly, I shall adopt Terhardt's terminology for my overall proposal in Part II, since his was also an attempt at grappling with the problem of consonance *vis-à-vis* the relationship between music and psychoacoustics. However, by adopting his terminology I do not mean to offer unhindered support for his theory of consonance; to the contrary, my account of the metaphysics of consonance is a significant departure from his view.[21]

Terhardt does not initially specify what he takes the problem of consonance to be. Yet, his remarks in 'Pitch, Consonance and Harmony' point to the claim that findings from psychoacoustic research on consonance perception appear "significant and consistent, and thus provide a solid basis of a certain kind of consonance", but not "the kind of consonance which plays a basic role in tonal music".[22] The former – which he then referred to as 'psychoacoustic consonance' – concerns "the undisturbed simultaneous sounding of pure tones", since "the disturbing element which destroys consonance is roughness".[23] However, while this may account for what he describes as the pleasantness of a sound in a very general sense,[24] Terhardt notes that psychoacoustic consonance alone does not explain the importance of harmonic intervals which form the basis of traditional tonal harmony.

Hence, under Terhardt's account the problem lies in reconciling the sensory experience of the psychoacoustic base of chords with the phenomenology of music perception – or, more specifically, the experience of harmony itself. On the one hand, psychoacoustics may be able to account both for the auditory perception of roughness as pertaining to certain dyads and chords and why listeners are most likely to favour those that are not affected by the acoustic phenomenon of interference. On the other hand, those explanations alone are nonetheless insufficient to account for consonance and dissonance perception within a distinctively musical context.

---

[21] The ways in which my account departs from Terhardt's will be detailed in Chapter Nine (Section 1).
[22] Terhardt, E., 'Pitch, Consonance and Harmony', *The Journal of the Acoustical Society of America*, 55, 1974, p. 1062.
[23] Ibid., p. 1062.
[24] Ibid., p. 1068.

Yet, Terhardt later came to refine his conceptual framework for understanding consonance. In 'The Concept of Musical Consonance: a Link between Music and Psychoacoustics', he set out to establish the boundaries between music and psychoacoustics whilst at the same time unifying them as two components of a broader conception of harmony. Specifically, his new attempt stems from the pertinent observation that "there is still considerable divergence of opinion and confusion about basic questions such as: how the term musical consonance can be appropriately defined, how musical consonance may be phenomenologically observed, which psychophysical principles are involved in it, and to what extent it essentially provides a suitable basis for musical theory".[25] This, however, comes with an acknowledgement of the apparent conflict with which the concept of consonance is fraught, as described in the previous section.

In what seems to be an attempt at addressing these questions, Terhardt proposes a conceptual distinction that is aimed at reconciling the psychoacoustics of consonance – which he came to refer to as 'sensory consonance' – with the experience of it within Western tonal music – which he designated as 'harmony' – under a twofold concept of 'musical consonance'.[26] Correspondingly, it is in terms of this distinction that I will be drawing from Terhardt's conceptual framework in my analysis of the aesthetics of consonance and dissonance in Chapter Eight, which will subsequently inform my account of the metaphysics of 'musical consonance' in Chapter Nine.

Under Terhardt's account, the sensory component – which is applicable to both consonance and dissonance – is described in terms of certain aspects of the psychoacoustic base of dyads and chords (such as those concerning the power spectra of complex tones and the presence of interference, for example) as well as their perceptual attributes (such as the auditory sensations of smoothness and roughness). Conversely, 'harmony' concerns the more systematic dimension of the auditory experience of dyads/chords within a musical context. Specifically, the 'harmony' component involves principles which conform to the rules of traditional tonal harmony, such as: the affinity between tones in intervallic relationships traditionally deemed consonant; the inner compatibility of component tones that enables chord inversions; the possibility of harmonising melodies; and the importance of fundamental-note relations within a chord.[27]

[25] Terhardt, E., 'The Concept of Musical Consonance: a Link between Music and Psychoacoustics', *Music Perception*, 1, 3, 1984, p. 277.
[26] Throughout Part II, I will refer to 'harmony' (between scare quotes) in order to designate one of the two components of 'musical consonance' (also between scare quotes) as proposed by Terhardt, and in order to distinguish them from the general use of the terms (in which case no scare quotes will be added).
[27] Ibid., p. 279.

One of the advantages of Terhardt's twofold conception of musical consonance is that it seems capable of preserving the distinctiveness of the perception of consonance (as well as dissonance) within the domains of music and psychoacoustics, respectively – and that without 'blurring' the boundaries between the two. Specifically, I will argue that this distinction is suitable for dealing with the consequences of the dual phenomenological nature of consonance and dissonance – a view which will be substantiated further in Chapter Nine. However, there are two important ways in which my thesis is a significant departure from his proposal. Specifically, Terhardt's account is one of reducibility, and his conception of 'harmony' is particularly limited in its scope – thus unable to account for the distinctively musical experience of chords adequately.

The first and most controversial aspect of Terhardt's account concerns his attempt to reduce 'musical consonance' to the psychoacoustic base of chords. In distinguishing between sensory consonance and 'harmony', Terhardt initially seems intent on conveying an account of the different ways in which chords may be perceived and aesthetically experienced. While the sensory component concerns the psychoacoustic dimension of the perception of consonance and dissonance, 'harmony' refers to the distinctively musical context in which they occur. But his overall conclusion is particularly problematic: Terhardt argues that empirical evidence from psychoacoustic research has not only provided the foundations of musical consonance but that "we can now say that each of its two components ... has been reduced to solid psychoacoustic facts: sensory consonance to the perception of sound fluctuations (roughness) and sharpness of timbre; and harmony to the perception of the pitch of complex signals [i.e. virtual pitch]".[28] Hence, Terhardt's thesis seems to be one of metaphysical reducibility.[29]

In response, I will reject this approach as not only misguided but at risk of undermining Terhardt's conceptual framework at its foundations. Conversely, the overall proposal I will seek to substantiate throughout Part II is that, although the perception both of sensory consonance and 'harmony' is ontologically dependent upon the psychoacoustic base, given the dual phenomenology of consonance the metaphysics of 'musical consonance' must be accounted for in two different ways. The key criterion that distinguishes between the two, I shall argue, is the nature of that relationship: while the sensory component is inexorably tied to the psychoacoustic base, the context-dependence of 'harmony' affords it a degree of flexibility which the former does not seem to allow.

---

[28] Terhardt, 1984, p. 292. I have provided an outline of the notion of virtual pitch in Chapter Four (Section 1) but, given its importance in Terhardt's thesis, it will also be considered further in Chapter Nine (Section 1).
[29] As indicated earlier, I will argue against this view in Chapter Nine (Section 1).

Secondly, although Terhardt seeks to reduce 'harmony' to psychoacoustic solid facts, his characterisation of it is nonetheless based on distinctively *musical* principles. Specifically, the principles to which he alludes when describing the domain of 'harmony' – e.g. tonal affinity, compatibility and fundamental-note relation, as mentioned earlier – indicate that his account encompasses elements that are normally taken to belong to the musical domain, but ones that do not exhaust its breadth. Arguably, by restricting the scope of his conception of 'harmony' to those principles, Terhardt would seem to be in a better position to justify advocating the thesis of metaphysical reducibility, in that a broader understanding of 'harmony' that includes the contextualist view would not have conformed to this approach. Yet, by dismissing other important dimensions of the aesthetics of consonance and dissonance, Terhardt's conception of 'harmony' is only partial.

As a result, Terhardt's conception of 'musical consonance' lacks the context-dependence aspect of consonance and dissonance perception. To use Rehding's earlier characterisation, his proposal is strictly limited to a combination of the acoustic and physiological/psychological approaches, seeing as his account does not go far enough in terms of the experience of chords within a distinctively musical context. Conversely, in accounting for the metaphysics of 'harmony', I shall treat the latter as pertaining to a broader phenomenon which encompasses that distinctively musical aspect. In particular, I will seek to include references to specific examples of context-dependence in my assessment of the aesthetics of tonal harmony in Chapter Eight (Section 3).

The purpose of the present chapter has been to outline the problem of consonance in light of the most relevant contributions to the conceptual debate around it. Having outlined the framework of my solution in this section, I take Terhardt's concept of 'musical consonance' – which, under my proposal, refers to tonal harmony as broadly conceived[30] – to be particularly useful to conveying my thesis of reconciliation between music and psychoacoustics. Rather than two distinct properties, I am going to argue that consonance is a single property with a dual phenomenological nature – a notion which is best captured by a two-component concept which enables the possibility of providing a unifying treatment of the metaphysics of tonal harmony. Prior to that, however, it is essential to consider the psychoacoustic base of chords first, to which we subsequently turn.

[30] For the purpose of disambiguation, I will take 'musical consonance' to be interchangeable with 'tonal harmony' insofar as it comprises both the psychoacoustic and the musical dimensions of consonance and dissonance. Yet, as indicated earlier, I shall use the term 'traditional tonal harmony' to be restricted to the foundational principles of the Western music tradition, such as those enshrined in Rameau's treatises.

Chapter Seven: The Psychoacoustics of Tonal Harmony

In the previous chapter, I indicated that consonance must be understood in terms of a dual phenomenology, as pertaining to the psychoacoustic and musical domains, respectively. It is the task of this chapter to account for the former. Drawing from my analysis of chords *qua* sound events in Chapter Three, I aim to outline certain aspects of their acoustic base that have come to be associated with the perception of consonance and dissonance. To that end, I shall focus on two acoustic properties in particular – namely, harmonicity and interference – which have been identified as the most likely physical causes for the perception of those phenomena as such. They will be considered in light of Helmholtz's approach to harmony as well as other empirically-based attempts to assess the roles of harmonicity and interference in consonance and dissonance perception. This will thus provide an overview of the fundamentals of the physical base of chords that are relevant for my account of the aesthetics of tonal harmony in the next chapter.

1. Introduction

The emergence of experimental psychoacoustics in the nineteenth century and its subsequent developments have been significant in providing a fertile ground for a range of explanations for the perceptual phenomena of consonance and dissonance. Yet, the question around the possibility of definitive psychoacoustic explanations for consonance and dissonance perception has not been fully settled to this day. In particular, in the last few decades there has been a flurry of research-based activity furthered by a technological advancement that has enabled a variety of computational models of consonance to be tested under experimental conditions. As a result, the psychoacoustics of tonal harmony benefits from a vast literature that continues to generate interest and further research. Given the sheer volume of publications in that field, two preliminary observations seem necessary in order to delineate the scope of the task ahead.

First, several of these studies cover an array of experiments on the perception of harmony that straddles across different musical traditions and beyond cultural boundaries. While they offer some interesting insights about the experience of consonance from an ethnological perspective, the remit of the present study is strictly limited to research on Western tonal harmony, instead. Hence, the cross-cultural element will not feature in the account I aim to give here. Second, given that several hypotheses have been put forward and tested by means of a range of experiments over the last two centuries – and with varying degrees of acceptance in the scientific community, as one would expect – it is not my intention to provide a survey of the historical developments in the psychoacoustics of tonal harmony. Rather, I shall confine the scope of this chapter to outlining the relevance of the psychoacoustic base in accounting for the perception of consonance and dissonance.

2. A physiological basis for harmony

In this first section, I will identify some of the most relevant aspects of the psychoacoustic base of chords for consonance and dissonance perception. This will be conducted in light of Helmholtz's contribution, which was aimed at providing a physiological and experimental basis for the perception of tonal harmony. Specifically, I will consider his views on consonance in parallel with Terhardt's twofold conception of musical consonance. Subsequently, I will seek to indicate how Helmholtz sought to establish a correlation between consonance and dissonance perception with the role of two acoustic properties in particular: namely, harmonicity and interference.

To this day, Helmholtz's *On the Sensations of Tone* (1863) is still regarded by many as the most comprehensive and detailed manual on musical acoustics. The full title of his work is particularly telling; his proposal is to consider the sensations of tone "as a physiological basis for the theory of music". With the developments in experimental research that took place during his time, Helmholtz was able to capitalise on new instrumental technology which afforded a scientific basis to his theses. He thus set out to "connect the boundaries of ... *physical and physiological acoustics* on the one side, and of *musical science and aesthetics* on the other".[1] To that end, his work covers substantial ground and is significant for its level of detail – albeit limited in terms of philosophical aesthetics.

Helmholtz distinguished between two kinds of consonance, namely, '*Konsonanz*' and '*Klangverwandtschaft*'. In Terhardt's view, these can be correlated with his twofold conception of 'musical consonance': while the former may be said to correspond to his own concept of sensory consonance, the latter is akin to his usage of 'harmony'. Indeed, Terhardt goes as far as to say that Helmholtz was "the founder of the two-component concept of musical consonance".[2] However, Terhardt attaches two caveats to this correspondence.

---

[1] Helmholtz, H., *On the Sensations of Tone*, New York: Dover, 1954, p. 1.
[2] Terhardt, 1984, p. 283.

First, Terhardt notes that Helmholtz's peculiar choice of terminology in his work – coupled with a somewhat unrefined distinction between the two kinds of consonance – has led to the perception that his account is fraught with obscurity. For Terhardt, it is precisely because of the apparent confusion engendered by his terminology that his work eventually became the target of substantial criticism, "rather than valid arguments against the ideas and facts standing behind those terms".[3] In addition to the issue of terminology, some translation difficulties have contributed further to the problem. For instance, Rehding notes that what came to be translated into English as 'compound tone' had originally been referred to as '*Klang*' in the original German. Yet, the latter "can indicate either a pitch presented as a complex timbre or the fused perception of a major triad with particular reference to its root (fundamental)".[4] Hence, the translation of '*Klang*' as 'compound tone' prevents the distinctive experience of fused perception from being conveyed as such.

Secondly, and more problematically still, Helmholtz's description of '*Konsonanz*' as resulting from the absence of the auditory sensation of roughness has been perceived by many as an attempt to account for 'musical consonance' *as a whole*. This, Terhardt argues, has led to a widespread misinterpretation of his thesis. Specifically, in the context of new developments in Western music from the late nineteenth-century onwards – such as the emergence of chromatic music and the movement towards atonality – Helmholtz's views on consonance came to be seen as an effort to 'naturalise' harmony, instead. It is for that reason, according to Terhardt, that his work became the recipient of scathing criticism and has consequently been strongly resisted against in music circles. But, although Helmholtz's theory has long been anathema to many, Terhardt professed himself to be indebted to his conceptual approach in describing it as "most appropriate, successful, and promising".[5]

Hence, a closer examination of Helmholtz's notion of '*Konsonanz*', in particular, is not only pertinent but essential to my account of the sensory component of 'musical consonance'. Specifically, Helmholtz understood it in terms of perfect consonances, which he took to result from certain intervallic relationships between component tones. Most importantly, he observed that these are related to one another in certain numerical ratios. It is in virtue of establishing this correlation as causation, instead, that his work was seen as a defence of a naturalistic harmony – and thus interpreted as an attempt at reviving the Pythagorean tradition of accounting for consonance in terms of mathematical relations.

---

[3] Ibid., p. 282.
[4] Rehding, A., 'Acoustics', In Kelly, M. (ed.), *Encyclopedia of Aesthetics*, OUP, 2014, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199747108.001.0001/acref-9780199747108-e-6. Rehding favours 'sonority' as a more suitable translation for '*Klang*', instead.
[5] Terhardt, 1984, p. 277.

Indeed, Helmholtz maintained that "when strings of different lengths but of the same make, and subjected to the same tension, were used to give the perfect consonances of the Octave, Fifth, or Fourth, their lengths must be in the ratios of 1 to 2, 2 to 3, or 3 to 4, respectively".[6] Hence, these perfect consonances should, in his view, be understood via an analysis of the acoustic base of the complex tones that we hear in music. In addition to the Pythagorean perfect consonances, other intervals such as major and minor thirds (at the ratios of 5:4 and 6:5, respectively) later came to be considered as consonances as well. Most importantly, Helmholtz indicated that what is distinctive about these intervals is that they "cause an agreeable kind of gentle and uniform excitement to the ear which is distinguished by its greater variety from that produced by a single compound tone".[7]

In his experiments, Helmholtz observed that certain harmonic relations between complex tones are not solely dependent upon their fundamental frequencies but, crucially, on their overtone series. Specifically, he noted that the different partials of individual tones will either 'blend' with no auditory sensation of roughness or, instead, 'beat' amongst themselves and produce that sensation. Helmholtz then sought to establish a causal relationship between the presence of intervals in certain numerical ratios and consonance, on the one hand, and the relationship between beating[8] and dissonance, on the other. This constitutes one of Helmholtz's most significant contributions to understanding the sensory component of 'musical consonance', which points to the two acoustic properties that are the main focus of this chapter – namely, harmonicity and interference.

Helmholtz's experiments led him to establish a correspondence between the presence of what later came to be known as 'harmonicity' and the perception of consonance. This resulted from the observation of certain harmonic relationships resulting from low-number integer ratios between overtones, which meant that their frequencies would not beat when combined. Conversely, an auditory sensation of roughness appeared to result from the combination of frequency components that "mutually disturb each other" – which is perceived as "distressing and exhausting".[9] Consequently, Helmholtz associated dissonance with the auditory sensation of roughness and attributed it to the acoustic phenomenon of beating, which is a type of destructive interference.

---

6 Helmholtz, p. 1.
7 Ibid., p. 330.
8 Beating tends to be perceived as an auditory sensation of roughness or a 'waxing and waning' quality to the sound heard, as it were. Technically speaking, this sensation mirrors the periodic changes to the amplitude of the resultant waveform, which consist in alternating patterns of reinforcement and interference between waves of different frequencies, amplitudes and phases. Speaks, p. 294.
9 Helmholtz, p. 330.

Helmholtz's account of interference, in particular, is of pioneering importance to the psychoacoustics of tonal harmony. Under his theory, the combination of two complex tones would be at its most dissonant when the beat frequency (i.e. the difference between the two frequencies) ranges between 25 and 30 Hz, after which the difference between them increases to a larger interval (such as a minor third) and the beats become indistinct; consequently, the auditory sensation of roughness disappears.[10] On that basis, sensory dissonance is taken to be causally related to the perception of the beat frequency, which would determine the degree to which the resulting sound will be perceived with a more or less intense sensation of roughness. This becomes an increasingly complex psychoacoustic phenomenon when one considers the case of chords, in which three or more complex tones – each comprising its own fundamental and overtones – may beat amongst themselves. These consequences will be considered in greater detail in Section 4.

3. Harmonicity

As seen in the previous section, one of Helmholtz's key contributions to the understanding of the psychoacoustics of consonance concerns the relationship between the fundamental frequency of a given complex tone and its overtones. By emphasising the importance of the overtone series in consonance perception and endorsing the view that perfect consonances are those harmonic intervals in specific whole-number relations, Helmholtz pointed to what later came to be referred to as *harmonicity*. In this section, I aim to provide an account of the nature of this acoustic property and its role in consonance perception, which has been a main focus in psychoacoustic research of late.

To that end, it is essential to consider, once again, the acoustic nature of complex tones. In Chapter Three, I identified three necessary (albeit insufficient) conditions that a given sound event must meet in order to be classed as a chord. The first condition – namely, periodicity – must be the case not only for chords to exist but musical sounds in general, which normally consist in complex tones (rather than pure ones). One of the properties of complex tones, as described in that chapter, is that the fundamental tone – which is what is associated with each individual note, such as A4 – is accompanied by a series of overtones that are multiples of its lowest frequency. When these partials are integer multiples, in particular, they are referred to as 'harmonics', and together they constitute the harmonic series of a given fundamental tone.

[10] Ibid., p. 179.

This recapitulation of the acoustic base of chord events is important for a specific reason: one of the factors that are taken to have a causal role in consonance perception is their harmonicity. Essentially, this refers to the acoustic property concerning the numerical relationship between different overtones and the fundamental. In the case of harmonics, their frequencies are in a whole-number relationship to the lowest frequency (i.e. the fundamental or first harmonic). Yet, in cases of non-harmonic partials, their frequencies are not integer multiples of the fundamental. In other words, while harmonic overtones indicate the presence of harmonicity, non-harmonic partials may lead to the opposite acoustic phenomenon – i.e. that of 'inharmonicity'. To that extent, two observations on the nature and role of harmonicity – and the absence thereof – must be considered.

First, the *perceived* harmonic series does not always comprise exact whole-number frequency multiples. As Benade notes, both the human voice and most musical sounds tend to involve frequency components that are related as such, but that is not always the case with musical sounds. That said, he points out that "nothing drastic happens to the perceived sound as long as the string has *nearly* integer frequency relations".[11] But instruments with fixed strings – such as the piano – present further difficulties. For instance, the stiffness of a piano's strings affects the harmonicity of the tones it produces. Most importantly, Rigden notes that the resultant inharmonicities become increasingly more pronounced for the highest frequencies of the overtone series of any individual tone, such that the higher the frequency of the component, the more it departs from its would-be harmonic counterpart.[12] Nonetheless, Giordano has argued that, although the spectra of real piano tones may contain a large number of partials (which are not strictly 'harmonic'), that "does not necessarily mean that they are all important for how the tone is perceived".[13]

Secondly, the acoustic property of harmonicity also concerns the difference between resolved and unresolved harmonics, which is essentially a perceptual/phenomenological distinction. As Andrew Oxenham describes it, resolved harmonics are those that can be recognised in the auditory experience of complex tones under certain circumstances – as in Helmholtz's experiments, for example, although they may also be heard by trained musicians. Contrastingly, unresolved harmonics tend to be those higher component frequencies that cannot generally be distinguished by the auditory system[14] – which are arguably those to which Giordano refers in the aforementioned claim.

---

[11] Benade, p. 62.
[12] Rigden, p. 166.
[13] Giordano, p. 17.
[14] Oxenham, A. J., 'The Perception of Musical Tones', in Deutsch, D. (ed.), *The Psychology of Music*, San Diego, London: Academic Press, 2012, p. 14.

The subtleties of harmonicity perception identified above will become more relevant in view of psychoacoustic findings on consonance perception, which I shall consider in Chapter Eight (Section 2). What is most important for the present section, however, is the link between the acoustic concepts of periodicity and harmonicity. Specifically, the repetition rates of oscillatory wave patterns that characterise the nature of periodicity are individual to each and every multiple frequency component of a complex tone; and it is the numerical relationships between them and the fundamental that entails their harmonicity. Both Helmholtz and Terhardt take these relationships to be essential for understanding the perception of consonance, although they approach it differently. While Terhardt set out to establish that the identification of a virtual pitch – as a result of the process of fundamental-tracking, as described in Chapter Three – provides a psychoacoustic basis for consonance perception,[15] Helmholtz was mostly concerned with the detrimental effects of beating resulting from the absence of harmonicity, as indicated in the previous section.

In spite of their views, the current psychoacoustic evidence supporting the 'harmonicity hypothesis' – i.e. the claim that harmonicity provides a causal explanation for the perception of consonance – is somewhat ambiguous. As William Forde Thompson notes, harmonicity is widely acknowledged as an important factor in pitch perception, but its role in consonance perception is "less clear".[16] Arguably, one reason for this lack of firm evidence in empirical findings is that it is difficult to dissociate the perception of harmonicity from that of the absence of beating, which is associated with inharmonicity.

Indeed, Terhardt himself points out that inharmonicity is not only explained in terms of the given frequencies of the different partials but also on the power spectra of the complex tones that compose the chord. As detailed in Chapter Four, the power (or sound) spectrum of a complex tone reveals the relative amplitudes of the partials (i.e. how much acoustic power their frequencies individually possess), thus determining the formant region of the tone and the distinctive timbre of the resultant sound. To that extent, Terhardt notes that the sensations of roughness/sharpness could instead be explained in terms of a sound spectrum in which more power is concentrated in the higher harmonics,[17] rather than by the absence of low-number ratios. Yet, although harmonicity is technically an acoustic property involving frequency components, its role in consonance perception would only be adequately assessed in psychoacoustic terms, i.e. in view of the workings of the auditory system in processing their power spectra.

---

[15] The role of virtual pitch in consonance perception will be considered in greater detail in Chapter Nine.
[16] Thompson, W. F., 'Intervals and Scales', In Deutsch, D. (ed.), *The Psychology of Music*, San Diego, London: Academic Press, 2012, p. 113.
[17] Terhardt, 1984, pp. 282, 285.

That said, several psychoacoustic studies have up until recently supported the view that harmonicity does, indeed, play a causal role in consonance perception. In particular, Peter Harrison and Marcus Pearce have recently endorsed that view. In 'Simultaneous Consonance in Music Perception and Composition', they provide a comprehensive survey of psychoacoustic findings on the perception of harmony as well as their own assessment of the current evidence towards understanding consonance *vis-à-vis* the acoustic base of chords. Amidst the range of possible hypotheses for explaining consonance perception causally (such as amplitude fluctuations, masking of neighbouring partials, vocal similarity and spectral evenness, amongst others),[18] Harrison and Pearce indicate that there is growing consensus that the presence of harmonicity is best placed as a causal explanation – a consensus which they both share.

Similarly to what Helmholtz and Terhardt previously attempted, Harrison and Pearce provide a conceptual framework to accompany the psychoacoustic data they analyse. For instance, they refer to 'periodicity/harmonicity' as a joint mechanism instead, which features prominently in their own account of consonance perception. Given that I have provided different definitions for those two terms, it seems important to consider their rationale in combining them. As stated earlier, the concepts of periodicity and harmonicity are, indeed, intrinsically related. Yet, Harrison and Pearce's characterisation of their relationship is particularly illuminating: on the one hand, they describe 'periodicity' as a representation of sounds as repetitive waveform patterns in the time domain; on the other hand, 'harmonicity' consists in its expression in the frequency domain, i.e. manifested in terms of the overtone series that results from multiples of the fundamental frequency.

This being the case, Harrison and Pearce maintain that, since "each periodic sound constitutes a (possibly incomplete) harmonic series rooted on its fundamental frequency" and given that "every harmonic (incomplete or complete) is periodic in its fundamental frequency", then harmonicity and periodicity may be treated as essentially equivalent phenomena.[19] It is worth noting, however, that while there *is* a distinction between the two, Harrison and Pearce's conceptual move does not raise any immediate worries for my assessment of the import from experimental psychoacoustic evidence to my account of the metaphysics of 'musical consonance' – so I shall henceforth replicate their terminology in relaying aspects of their analysis.

---

[18] Harrison, P. M. C., Pearce, M. T., 'Simultaneous Consonance in Music Perception and Composition', *Psychological Review*, 127, 2, 2020, p. 216. The authors attach a series of references to these possible explanations, which have been omitted for ease of exposition.
[19] Ibid., p. 217.

In spite of the growing consensus in recent psychoacoustic studies that periodicity/harmonicity is a determinant factor in consonance perception, Harrison and Pearce also acknowledge that the overall evidence presented is somewhat ambiguous. In their survey of a broad range of studies, they note that periodicity/harmonicity plays a decisive role in auditory scene analysis[20] as well as speech processing – and yet, its association with consonance is once again less clear. Nonetheless, Harrison and Pearce consider two explanatory routes, based on the latest evidence. One possibility is that long-term exposure to the human voice as well as Western music may induce familiarity with periodicity/harmonicity, such that participants may favour musical sounds that have high harmonicity. Alternatively, they suggest that "the ecological importance of interpreting human vocalisations creates a selective pressure to perceive these vocalisations as attractive".[21] Both of these involve what I shall refer to as the 'vocal similarity' hypothesis – which, as mentioned earlier, is one of the hypotheses behind consonance perception.

In fact, Terhardt himself had argued for a version of that hypothesis. In the context of his defence of the explanatory power of the concept of virtual pitch in accounting for the perception of musical intervals, Terhardt proposed that virtual-pitch cues (i.e. those involved in fundamental tracking) can only be generated subject to a previous perceptual learning process, which he takes to happen concomitantly with the cognitive development required for acquiring speech-processing abilities. "In that process", he states, "the correlations between the spectral-pitch cues of voiced speech sounds (i.e. of harmonic complex tones) are recognised and stored. [Thus] The knowledge about harmonic pitch relations that is acquired in this way is employed by the system to generate virtual pitch".[22]

By associating the relevance of harmonicity for consonance perception with the spectral pitch of speech sounds, Terhardt once again echoes Helmholtz's physiological approach to the phenomenon of consonance *qua* '*Konsonanz*'. He noted that Helmholtz had already considered that the notion of musical intervals may be grasped by means of an unconscious processing of the harmonics present in vocal sounds, as one does with the singing voice.[23] For instance, Helmholtz argued that the power spectra of those harmonics play an important role in vowel sounds, in that they may become "proportionally too powerful" in loud vocalic sounds as compared with those of a weakly spoken vowel.[24]

---

[20] The process of auditory scene analysis was introduced in Chapter Two, with reference to Bregman's widely known specialism in that area.
[21] Ibid., p. 218.
[22] Terhardt, 1974, p. 1063.
[23] Terhardt, 1984, p. 286.
[24] Helmholtz, p. 112.

Although the vocal similarity hypothesis has gained some support as an explanation for the importance of harmonicity, Harrison and Pearce indicate that no studies have yet demonstrated that it contributes to consonance other than via the two acoustic properties being examined here, so they evaluate it solely as part of the periodicity/harmonicity and interference hypotheses. Nonetheless, what is particularly relevant about the vocal similarity thesis is that it is rooted in the empirical evidence that the human voice involves frequencies in harmonic relationships,[25] so it 'scores high' in harmonicity.

Furthermore, there have been other approaches to assessing the role of periodicity/harmonicity in consonance perception based on physiological processes. One such approach can be seen in recent experimental evidence from neuroscientific research, which indicates a correlation between 'neural periodicity' and consonance. For instance, a study undertaken by Gavin Bidelman and Ananthanarayan Krishnan indicates that brainstem frequency-following responses from participants can be indexed to the perceived consonance of dyads and chords in terms of 'neural pitch salience', and they generate ratings that point to the harmonic relations between isolated intervals. As a result, Bidelman and Krishnan suggest that "brainstem neural mechanisms mediating pitch processing show preferential encoding of consonant musical relationships", such that high levels of neural pitch salience values can be used to predict both consonance and dissonance ratings given by participants.[26] Hence, this may be taken to support the view that periodicity/harmonicity is a determinant factor in consonance perception.

4. Interference

Besides harmonicity, another acoustic property that is seen as having a fundamental role in the psychoacoustics of tonal harmony – or, more specifically, in *dissonance* perception – is that of interference. In particular, understanding the effects of beating was one of Helmholtz's key contributions to musical acoustics. Similarly, in their assessment of the latest empirical findings, Harrison and Pearce point to interference effects as being correlated with dissonance ratings across several different studies.[27] In this section, I aim to offer a brief overview of the nature of interference, with particular focus on destructive interference as resulting from the acoustic phenomena of masking and beating.

[25] Bowling, D., Purves, D., Gill, K., 'Vocal Similarity Predicts the Relative Attraction of Musical Chords', *Proceedings of the National Academy of Sciences of the United States of America*, 115, 2018, p. 216.
[26] Bidelman, G. M., Krishnan, A., 'Neural Correlates of Consonance, Dissonance, and the Hierarchy of Musical Pitch in the Human Brainstem', *Journal of Neuroscience*, 29, 42, 2009, p. 13165.
[27] Harrison, Pearce, p. 218.

Prior to examining its role in dissonance perception, it seems essential to consider the nature of interference at closer inspection. As a physical phenomenon, the occurrence of interference is due to differences in frequency, amplitude and phase between the relevant component waves, such that the properties of the resultant wave are determined by the pattern of interactions between their frequencies. There are two ways in which interference may occur: it may be constructive, in which case one wave reinforces the other with a maximal increase in intensity; or it may be destructive, which occurs when one wave pattern cancels the other out. This may be observed to varying degrees, to the extent that, in some cases, a total cancellation may ensue and no sound is perceived.[28]

It is with destructive interference, in particular, that Harrison and Pearce are concerned – as tends to be the case in psychoacoustic research at large. Yet, it is worth noting that a range of factors in musical sound production may lead to the occurrence of this type of interference. For instance, Giordano notes that, rather than just being aimed at amplifying the sound of a piano, the nature of its soundboard is fundamental to understanding the different levels of destructive interference that may take place. Given the interaction between sound waves from all different parts of the soundboard at a given time, he observes that "one finds many cancellations and the net sound signal at the listener is much smaller than would be the case if all these sound waves had arrived in phase with each other".[29] Nonetheless, for the purposes of assessments of isolated chords in experiments, these broader instrument-based concerns are arguably secondary.

One of the phenomena that generate destructive interference is known as 'masking', which involves not only the circumstances of musical sound production but certain aspects of the frequency-processing abilities of the auditory system as described in Chapter Four. Briefly, masking consists in the preponderance of one particular frequency such that it obstructs the perception of another. However, masking is not a symmetric phenomenon. As Stanley Gelfand notes, low-frequency tones are better 'maskers' than high-frequency ones: while the latter are "only effective over a relatively narrow frequency range in the vicinity of the masker frequency, ... low frequencies tend to be effective maskers over a very wide range of frequencies".[30] Furthermore, although masking is a complex phenomenon, Harrison and Pearce point out that it is a long-established principle that masking increases if the frequency difference is small and at a higher sound pressure level.[31]

---

[28] Speaks, pp. 268-269.
[29] Giordano, p. 101.
[30] Gelfand, S. A., *Hearing: An Introduction to Psychological and Physiological Acoustics*, New York: Marcel Dekker, 1981, p. 241.
[31] Harrison, Pearce, pp. 218-219.

A more prominent source of destructive interference, however, is the acoustic phenomenon of beating. As noted earlier, the impact of beating on dissonance perception was already a concern for Helmholtz, who set up experiments to measure the presence of beats as and when they occurred between certain overtones. But, in another instance of the terminology issue alluded to earlier, he arbitrarily distinguished between 'interference' as resulting from the disturbance caused by two "perfectly equal simple tones", on the one hand, and 'beats' as referring to the fusion of two "nearly equal" simple tones, on the other.[32] That aside, what he sought to emphasise is that the main issue concerns the phase of each wave pattern as they merge. By shifting in and out of phase, the resultant waveform oscillates between constructive and destructive interference – thus generating an auditory sensation of roughness.

There are two different kinds of beating that may affect dissonance perception. Specifically, they consist in beats of 'imperfect unisons' or beats of 'mistuned consonances'. As Olson describes them, imperfect unisons refer to very marginal differences in frequencies (e.g. 440 Hz and 442 Hz) which would not prevent the auditory system from perceiving the two tones as being the same tone (which is, in that particular case, A4). Contrastingly, mistuned consonances arise from a combination of pure tones that would normally constitute a consonant interval (e.g. 440 Hz and 880 Hz) but one of them departs slightly from the set frequency (e.g. 440 Hz and 883 Hz). Hence, when the frequency of one of the tones varies, the number of beats heard per second will also vary, which may result in an auditory sensation of roughness.[33]

Yet, with regard to complex tones, Helmholtz maintained that they will have both their pitch and timbre affected by the presence of beats. This is because the interaction between partials of mistuned frequencies is particularly problematic. Specifically, Helmholtz discovered that beating amongst overtones plays a greater role in dissonance perception because, for each beat of the fundamental, there will be two of the second harmonic, three of the third harmonic, and so forth.[34] By considering the formant region of individual complex tones, one would be able to assess whether the beating partials carry sufficient power in order to make a difference – in which case the sound generated is perceived as dissonant. It is thus unsurprising that psychoacoustic researchers have tended to measure dissonance ratings in terms of beats under the 'interference hypothesis' – i.e. the view that destructive interference is a causal explanation for dissonance perception.

[32] Helmholtz, p. 160.
[33] Olson, p. 259.
[34] Helmholtz, p. 165.

Generally, psychoacoustic assessments of the role of interference tend to be focused on the beat frequency between different component waves. As seen earlier, this approach was taken by Helmholtz in his quest for a physiological explanation for dissonance perception, but other factors are also seen as contributing to that phenomenon. For instance, Harrison and Pearce note that slow amplitude fluctuations between approximately 0.1 to 5 Hz are not perceived as unpleasant variations in loudness – i.e. the sensation of roughness caused by that range of beat frequency is relatively minor. Yet, fast amplitude fluctuations of about 20 to 30 Hz are distinctively perceived with a sensation of roughness. Hence, the correlation between beat frequency and dissonance became commonplace in the psychoacoustic literature on harmony. Specifically, it is associated with what is known as the 'critical band' (or 'critical bandwidth').

The notion of a critical band is closely related to those of beating and masking, albeit at a more granular level of the physiology of auditory perception. In particular, it is associated with the frequency-analysing role performed at inner ear level, as described in Chapter Four. The applications of the notion of critical band are complex, but it fundamentally establishes the degree of separation that must be observed between two different frequencies such that they may be heard as individual tones merging, rather than a single fused-tone pitch that beats rapidly. Thus, the critical band is a frequency region within which simultaneously sounding tones *will* beat. The parameters of this region are worked out within the cochlea, along the basilar membrane, as a function of its ability to operate on the basis of tonotopical mapping.[35] In other words, the cochlea is responsible for discriminating between the component frequencies of the incident waveform; in doing so, a certain limit of frequency discrimination is established, which consists in the required difference between any two frequencies such that they may be heard separately.

As White points out, "if the beat frequency ... is smaller than a certain amount, the resonance regions on the basilar membrane overlap, and we hear but one tone of intermediate frequency, but beating in loudness. These are the beats known to every musician and are used by them in tuning their instruments".[36] Perhaps a simpler way of explaining the relationship between the notion of critical band and the auditory sensation of roughness is in terms of frequency regions. Rigden identifies four in total: the first region, which he calls the 'beat region', covers a beat frequency of 0 to 15 Hz, in which case only a single fused-tone pitch is perceived; a second region, comprising a frequency

[35] The relevance of tonotopical mapping for the perception of complex tones was considered in Chapter Four (Section 1).
[36] White, 1980, p. 117.

difference between 15 Hz and the limit of frequency discrimination mentioned above, is marked by the perception of roughness of the fused-tone pitch; a third region, in which the frequency difference is beyond the limit of frequency discrimination but falls within the critical band, thus generating beats and, consequently, roughness; and a fourth region, beyond the boundary of the critical band, in which case the roughness sensation fades out and two smooth tones are heard.[37] The extent of the critical band is thus limited by the value of the frequency difference at which the auditory sensation of roughness subsides.

Albeit fraught with technical descriptions, this brief characterisation of the concept of critical band will be relevant for my account of the metaphysics of sensory dissonance in Chapter Nine (Section 2.1.3). For the present purposes, the importance of that notion could not be overstated, in that it further confirms the correlation between interference and roughness and thus with dissonance perception. In addition, the fact that the width of the critical band differs across the audible pitch range arguably makes it an essential conceptual tool for the assessment of experimental psychoacoustic evidence. As Thompson notes, "for pitches below about 400 Hz, the width of a critical band varies in a manner that is roughly intermediate between a linear frequency scale (Hertz) and a logarithmic frequency scale (semitones). For pitches above 400 Hz, the width of a critical band varies in a manner that is close to logarithmic".[38] The implications of these variations go beyond the scope of this study; yet, as a consequence of them, Thompson points to the fact that sensory dissonance is more pronounced for low-register pitches as compared to high pitches. This is an example of the relevance of the psychoacoustic base for understanding the sensory component of 'musical consonance', which will inform my account in Chapter Nine.

The purpose of this chapter has been to provide an account of the psychoacoustics of tonal harmony. After considering Helmholtz's contribution to an understanding of the acoustic base of chords, I examined the nature of two acoustic properties, in particular. Granted, both harmonicity and interference are acoustic phenomena. But, although it is essential to consider the nature of the acoustic base of chords in order to account for the metaphysics of 'musical consonance', the question remains as to what extent the two properties perform a causal role in assessments of the *aesthetics* of consonance and dissonance. Indeed, my earlier thesis that each of these has a dual phenomenology calls for an appraisal of the aesthetic experience of consonance and dissonance both *qua* sensory properties and *qua* musical properties. It is the task of the next chapter to do just that.

---

[37] Rigden, pp. 68-70.
[38] Thompson, p. 110.

Chapter Eight: The Aesthetics of Tonal Harmony

In Chapter Six, I proposed that consonance is a single property with a dual phenomenology. It may be perceived as a psychoacoustic property, but it is also experienced as a distinctively musical property. Given that consonance is a property of chords, I subsequently sought to convey an account of their acoustic base in Chapter Seven. The purpose of the present chapter is thus to consider the aesthetics of tonal harmony in view of those earlier considerations. In line with the distinction between the sensory component and 'harmony', I aim to offer an analysis of consonance and dissonance as pertaining to chords *qua* isolated auditory objects and *qua* individuals within a musical context. I will show that the two-component conception of 'musical consonance' enables us to account for the dual phenomenology of the consonance/dissonance of chord tokens both individually and within a musical context, whilst also providing a conceptual framework for a reconciliation between music and psychoacoustics.

1. Introduction

Having previously adopted Terhardt's distinction between sensory consonance and 'harmony', the aim of this chapter is to account for the aesthetics of each component of 'musical consonance' in view of their acoustic base, as outlined in Chapter Seven. Given the psychoacoustic nature of the sensory component, my account of the aesthetic experience of chords *qua* isolated auditory objects will be closely related to the discussion in that chapter, so it will be conducted in light of the latest psychoacoustic evidence available. Conversely, my treatment of the 'harmony' component will be more closely aligned with an understanding of musical aesthetics as such.

Granted, an account of the aesthetic experience of isolated chord tokens does not need to be confined to psychoacoustic experiments, in that most people are able to qualify them as either 'pleasant' or 'jarring', for instance. Yet, for the purposes of an assessment of aesthetic responses to the sensory kind of consonance/dissonance, the scope of my account in Section 2 will be restricted to an appraisal of the psychoacoustic evidence for the causal role of interference and harmonicity in consonance/dissonance perception – which, I will argue, echoes some of the conclusions contained in Helmholtz's work.

Despite a certain degree of disparity in some empirical findings, my argument will follow the growing consensus that both harmonicity and interference are fundamental acoustic properties to the perception of sensory consonance/dissonance. However, it is not my intention to assess the merits and flaws of individual studies but, rather, focus on certain methodological and philosophical issues for the aesthetics of the sensory component of 'musical consonance' in terms of aesthetic judgements and responses.

Conversely, my account in Section 3 concerns the distinctively musical experience of chord tokens. Given the breadth that is inherent to tonal harmony as an academic subject, my ambition in that section will be confined to the contrast between the psychoacoustic view of consonance/dissonance and contextualist/subjectivist approaches to harmony. While the contextualist view has been most clearly articulated by Scruton, the subjectivist approach has been defended by Rehding in his account of consonance and dissonance. In line with Scruton's view, I will argue that 'harmony' is context-sensitive. However, I will reject their dismissal of the importance of the acoustic base for a musical experience. My account will also include an assessment of aesthetic responses to consonance/dissonance in a distinctively musical context. This will be considered in light of my analysis of aesthetic responses to, as well as judgements of, the sensory consonance/dissonance of isolated dyads and chords. I will show that the contrast between them is what warrants my thesis of the dual phenomenology of consonance and dissonance, as indicated in Chapter Six.

2. The aesthetics of sensory consonance/dissonance

Since chords are objects of aesthetic appreciation in their own right – whether or not they are perceived as strictly musical – this first section is concerned with the aesthetics of the sensory component of 'musical consonance'. Having outlined some key aspects of the psychoacoustics of tonal harmony in Chapter Seven, my account here will be focused on the relevance of the acoustic properties of harmonicity and interference for aesthetic responses to isolated dyad/chord tokens. More specifically, I aim to consider the nature of those responses themselves and corresponding aesthetic judgements. I will argue that the sensory component must be understood *both* in terms of auditory sensations *and* aesthetic responses – a claim that is particularly relevant for my account of the metaphysics of sensory consonance/dissonance in the next chapter. Given the experimental character of the research on the perception of dyads and chords, both 'consonance' and 'dissonance' will be treated in this section with reference to participant ratings.

Prior to considering the implications of psychoacoustic research for the aesthetics of sensory consonance/dissonance, it seems important to identify the harmonic intervals that typically form the basis for psychoacoustic assessments. The standard description of the kind of harmony that this section is concerned with is generally given in terms of intervallic relationships within certain chords. For instance, in a recent manual aimed at providing an 'engineering view' of harmony, a chord is said to be consonant when it includes only consonant intervals, such as the octave, the perfect fifth, the perfect fourth, as well as thirds and sixths. Conversely, a given chord is deemed dissonant if it comprises "at least one dissonant interval", which would include seconds, sevenths, as well as augmented and diminished intervals.[1] This is the received account of tonal harmony from acoustics.

Although the characterisation above may seem useful for an engineering module on the general principles underlying tonal harmony, music specialists these days would arguably take these descriptions to be at best oversimplifications or remnants of orthodoxy. The reason why they would most likely reject it will be more appropriately considered in Section 3, but their dismissal of such prescriptive criteria points to the heart of the distinction between sensory consonance/dissonance and 'harmony'. This is because music specialists and musicians alike have come to treat the latter as fundamentally context-dependent, such that consonance and dissonance may not be taken as categorical properties of chords at all. This is in direct opposition to the understanding of harmony most famously associated with Helmholtz, as detailed in Chapter Seven.

Helmholtz's own taxonomy is somewhat different from the one outlined above, however. Specifically, he distinguished between absolute, perfect, medial and imperfect consonances. Under his classification, the octave, the twelfth and the double octave are absolute consonances, followed by the perfect intervals of the fifth and the fourth, the medial consonances of the major sixth and the major third, and the imperfect minor third and minor sixth.[2] As seen in the previous chapter, the rationale for Helmholtz's view is based on the low-number integer ratios between frequencies in those intervallic relationships. From the highest to the lowest "degree of harmoniousness of consonance" (with respective ratios in brackets), he takes the octave (2:1) to be the most consonant of all, followed by the twelfth (3:1), the fifth (3:2), the fourth (4:3), the major sixth (5:3), the major third (5:4), the minor third (6:5) and the minor sixth (8:5).[3]

[1] Barbancho, A. M., Barbancho, I., Tardón, L. J., Molina, E., *Database of Piano Chords: an Engineering View of Harmony*, Springer, 2013, pp. 13-14.
[2] Helmholtz, p. 194.
[3] Ibid., p. 183.

In particular, the second most consonant interval featuring in the hierarchy above – i.e. the absolute consonance of the twelfth – may come as a surprise even to a familiar reader, in that it does not normally appear as such in most manuals on musical acoustics. Aside from the fact that this type of compound interval is not the most commonly found in traditional tonal harmony, what is particularly distinctive about Helmholtz's assessment is that it is purely based on the empirical observation of the presence of harmonicity or, conversely, the absence of beating between different partials in a twelfth – which stems from his observation of "the undisturbed or disturbed coexistence of sounds".[4]

As indicated in the previous chapter, the view that both harmonicity and destructive interference in the form of beating perform a causal role in the perception of sensory consonance/dissonance has gained support in psychoacoustic circles ever since. For instance, Harrison and Pearce have collated extensive evidence concerning the possible mechanisms underlying consonance perception, amongst which both harmonicity and interference feature prominently. Prior to considering the relevance of their conclusions for the aesthetics of the sensory component of 'musical consonance', a brief note on their methodology and taxonomy seems important.

On the basis of their comprehensive survey, Harrison and Pearce identify three main mechanisms for which there is a greater or lesser degree of consensus in the psychoacoustic literature, namely: interference, periodicity/harmonicity and culture-specific factors related to the perception of consonance.[5] They subsequently classify the evidence under the following threefold criteria: (i) stimulus effects, based on which they assess how consonance perception varies as a function of the type of stimulus and/or circumstances affecting it; (ii) listener effects, under which the perception of harmony may also vary due to the nature of the participants' responses; and (iii) composition effects – i.e. the only distinctively musical category of evidence – which focuses on how compositional practice may provide evidence for assessing consonance perception.[6] Given the scope of the present study, these effects will be considered in terms of Western tonal harmony.

[4] Ibid., p. 184.

[5] Harrison and Pearce also consider the 'culture hypothesis', under which certain culture-specific factors are taken to contribute to consonance perception. However, this hypothesis will not generally be assessed here, since the present section is concerned with the sensory component of 'musical consonance', as opposed to a culture-specific experience of chords *qua* musical individuals. That said, in the specific case of Western music, Harrison and Pearce point out that some researchers have suggested that "Western listeners internalise codified conventions of Western harmony ... whereas others argue that Westerners learn aesthetic preferences for periodicity/harmonicity ... [but] these competing explanations have yet to be tested". Hence, they conclude that the effects of culture-specific factors on consonance perception in the specific case of Western harmony "remain unclear". For that reason, my account will be focused on the other two mechanisms, for which the psychoacoustic evidence available is significantly more robust. Harrison, Pearce, p. 224.

[6] Ibid., pp. 221-223.

2.1. *Psychoacoustic findings*

The purpose of this section is to consider the import from psychoacoustic findings into an account of consonance/dissonance perception. In particular, I shall focus on the empirical relevance both of periodicity/harmonicity and interference for the sensory component of 'musical consonance', which will be restricted to assessments of dyad and chord tokens. This will be conducted in light of the latest psychoacoustic evidence as collated by Harrison and Pearce, which has been particularly abundant in this century.

In their survey of psychoacoustic findings on harmony, Harrison and Pearce identify both periodicity/harmonicity and interference as performing causal roles in consonance and dissonance perception. Given my account in Chapter Seven, this may seem unsurprising. However, the most telling aspect of their assessment is that, on a granular level, there are some relevant contrasts between the evidence presented for interference and for periodicity/harmonicity as causal explanations. For instance, under stimulus effects, both the spectral content of a given chord's component tones and differences in pitch height were observed as evidence for the role of interference in consonance perception, but neither of them was identified in the case of periodicity/harmonicity.

In the case of the spectral content of chords, Harrison and Pearce point out that their perceived consonance depends on the power spectrum of each of its constituent complex tones, since this will determine whether the beating partials carry sufficient acoustic power to be heard as such. Most importantly, this applies whether the partials are harmonic or not – so harmonicity does not itself seem to be a contributing factor. To that extent, they note that "interference theories clearly predict these effects of tone spectra on consonance" whereas periodicity/harmonicity does not.[7] Similarly, in the case of pitch height, the perceptual difference between certain intervals on the lower register of a musical instrument and the equivalent at higher frequencies can be explained on the basis of the notion of critical band described in Chapter Seven (Section 4). As Thompson points out, the widths of the critical band between the two registers vary – and this variation may be explained by interference theories.[8] For Harrison and Pearce, this testifies to the explanatory power of interference theories, in that they are able not only to explain why intervals that are traditionally deemed consonant do not sound as consonant on the lower register but also *predict* that phenomenon, in view of the extent of the critical band.

---

[7] Ibid., p. 221.

[8] Specifically, Thompson notes that the logarithmic nature of pitch distance means that the critical band region will be smaller when the interval ratio appears at lower frequencies. Thompson, p. 110.

That said, the evidence collated by Harrison and Pearce indicates that interference might not play an important role with regard to listener effects, which are particularly relevant for a consideration of the aesthetics of sensory consonance/dissonance. This was observed in experiments with participants who suffer from congenital amusia – a condition that affects the perception of musical sounds at neurological level. Harrison and Pearce note that those experiments seem to indicate that 'amusics' do not tend to exhibit aversion to traditionally dissonant chords but, all else being equal, normal aversion to interference. Consequently, the conclusion drawn in a related study by Cousineau et al. was that interference is not relevant to consonance perception.[9] However, this finding is not shared in another study with amusics by Marin et al., in which they suggest that "small but reliable preferences for consonance ... were driven by interference".[10] Hence, the conclusion that interference has no role with regard to listener effects has been disputed.

Another example of the purported lack of evidence for the role of interference concerning listener effects may be found in research with Western listeners[11] undertaken by McDermott et al. Specifically, their study indicates that participant ratings could be correlated with periodicity/harmonicity preferences but not so with interference.[12] However, although this could be seen as detrimental to the interference hypothesis, Harrison and Pearce warn that findings from this study in particular must be regarded as only preliminary, given the limitations of the psychometric measures involved.[13]

Furthermore, Harrison and Pearce point to the case of composition effects, where the evidence for interference concerning consonance/dissonance ratings is more robust. With the exception of chord prevalences in Western compositional practice – for which interference *alone* would not be able to account – other music-specific aspects were observed as confirming the role of interference in dissonance perception, in particular.[14] These include experiments with musical scales as well as chord spacing, for instance. Hence, although the role of interference has not been thoroughly confirmed, it nonetheless features as a mechanism that seems to underlie the kind of compositional practice that has traditionally aimed at minimising the occurrence of dissonant dyads and chords.

---

[9] Cousineau, M., McDermott, J., Peretz, I., 'The Basis of Musical Consonance as Revealed by Congenital Amusia', *Proceedings of the National Academy of Sciences of the United States of America*, 109, 2012, pp. 19858-19863.
[10] Harrison, Pearce, p. 222. This is in reference to: Marin, M., Thompson, W. F., Gingras, B., Stewart, L., 'Affective Evaluation of Simultaneous Tone Combinations in Congenital Amusia', *Neuropsychologia*, 78, 2015, pp. 207-220.
[11] As Harrison and Pearce define it, 'Western listeners' refers to listeners from the musical traditions that they broadly identify with the musical style historically developed in Europe. Harrison, Pearce, p. 216 (footnote).
[12] McDermott, J., Lehr, A., Oxenham, A., 'Individual Differences Reveal the Basis of Consonance', *Current Biology*, 20, 2010, pp. 1035-1041.
[13] Harrison, Pearce, p. 222.
[14] Ibid., pp. 223-224.

In spite of the lack of conclusive empirical evidence for the role of interference, Harrison and Pearce maintain that its effects consist in one of the best candidates to provide a causal explanation for consonance/dissonance perception. Yet, they acknowledge that their assessment of the evidence is at odds with other recent studies that challenge the view that interference is relevant for consonance perception. Notwithstanding that objection, Harrison and Pearce's survey of the evidence points to the overall conclusion that "consonance perception in Western listeners is likely to be driven by multiple psychological mechanisms, including interference, periodicity/harmonicity and cultural background".[15] But, as indicated earlier,[16] the evidence for the relevance of 'cultural background' is unclear. Indeed, Harrison and Pearce recognise that familiarity with certain musical sounds (which cultural factors might involve) may nonetheless be due to their periodicity, harmonicity or the internalisation of certain Western tonal structures, instead.[17]

In addition, one particularly relevant observation from Harrison and Pearce's survey concerns the nature of those 'Western tonal structures', in that the evidence concerning them is distinctly controversial. This is because recent psychoacoustic studies have provided conflicting explanations for the consonance of certain dyad/chord patterns. Specifically, Harrison and Pearce identify several studies that are implicated, which attempt to provide the following explanations for consonance on the basis of chord-structure effects, namely: interference (Hutchinson and Knopoff, 1978); interference and "additional unknown factors" (Vassilakis, 2001); interference and cultural knowledge (Johnson-Laird et al., 2012); periodicity/harmonicity (Stolzenburg, 2015); periodicity/harmonicity and interference (Marin et al., 2015); interference and sharpness (Lahdelma and Eerola, 2016); and vocal similarity (Bowling et al., 2018).[18] Although there is a certain degree of overlap in some of these cases, the most controversial aspect of their competing explanations is that these studies are often based on computational models that test the ability of a given hypothesis to predict both consonance and dissonance ratings amongst participants.

---

[15] Ibid., p. 224.
[16] Harrison and Pearce's conclusion regarding the potential effects of cultural background on consonance perception falls under their assessment of the culture hypothesis, which I have briefly outlined in footnote no. 5 on page 133.
[17] Ibid., p. 219. The issue of familiarity will be considered under my account in Sections 2.3 and 2.4, instead.
[18] Hutchinson, W., Knopoff, L., 'The Acoustic Component of Western Consonance', *Journal of New Music Research*, 7, 1978, pp. 1-29; Vassilakis, P., *Perceptual and Physical Properties of Amplitude Fluctuation and Their Musical Significance* (PhD Thesis), Los Angeles: UCLA, 2001; Johnson-Laird, P., Kang, O., Leong, Y., 'On Musical Dissonance', *Music Perception*, 30, 2012, pp. 19-35; Stolzenburg, F., 'Harmony Perception by Periodicity Detection', *Journal of Mathematics and Music*, 9, 2015, pp. 215-238; Marin et al., 2012 (as per footnote no. 10 on page 135); Lahdelma, I., Eerola, T., 'Mild Dissonance Preferred Over Consonance in Single Chord Perception', *I-Perception*, 2016, pp. 1-21; Bowling, D., Purves, D., Gill, K., 'Vocal Similarity Predicts the Relative Attraction of Musical Chords', *Proceedings of the National Academy of Sciences of the United States of America*, 115, 2018, pp. 216-221. The 'vocal similarity' hypothesis was considered in Chapter Seven (Section 3).

These discrepancies in findings concerning the specific case of chord structures have an important bearing for the aesthetics of sensory consonance. This is because the chord structures that psychoacoustic research tends to be concerned with comprise those intervals that have traditionally been seen as the most fundamental elements of tonal harmony, i.e. the intervals found in triads and their inversions. For instance, Harrison and Pearce note that Western listeners tend to treat major triads as prime examples of consonance.[19] These, like minor triads, comprise the consonant intervals identified at the beginning of Section 2 (i.e. major/minor thirds, fifths, fourths and major/minor sixths). What is particularly distinctive about them, in psychoacoustic terms, is that those intervals have high harmonicity, whereas others that are regarded as dissonant – e.g. the augmented fourth, also known as the tritone – have relatively low harmonicity. This, as Harrison and Pearce point out, could be attributed to the fact that the tritone "cannot be easily approximated by a simple frequency ratio".[20]

The question of the observance of whole-number frequency ratios – as noted in previous chapters – resurfaces again in the assessment of harmonicity as one of the key mechanisms underlying consonance perception. As seen in Chapter Seven, Helmholtz had already taken the ratios of certain intervals to be essential to determining whether the overtones would beat amongst themselves and thus result in an auditory sensation of roughness. If that is the case, as the evidence seems to support, it is not unreasonable to expect that psychoacoustic studies that are focused on the consonance of specific chord patterns such as triads would point to conclusions that involve both harmonicity and interference – as Harrison and Pearce's own assessment does.

For that reason, Harrison and Pearce are not alarmed by the above findings on chord structures. Instead, it is the different computational models underscoring those studies that Harrison and Pearce are especially concerned with, since they take them to have an important role in psychoacoustic research by enabling consonance and dissonance ratings to be predicted. In their view, those discrepancies may have resulted from some methodological difficulties, amongst which they identify: the extent and scope of testing methods, such that individual studies test their own theories but rarely any others; and the size of stimulus sets, which tends to be relatively small – arguably in order to support reliable conclusions.[21] It is with these limitations, along with further methodological issues that make relevant findings difficult to generalise, that the next section is concerned.

[19] Harrison, Pearce, p. 221.
[20] Ibid., p. 218.
[21] Ibid., p. 222.

2.2. *Methodological issues*

The predictive power of psychoacoustic research is particularly relevant for an account of the sensory component of 'musical consonance', in that it points to the nature of the relationship between consonance/dissonance and the psychoacoustic base of chords. Specifically, if consonance/dissonance ratings may be predicted on the basis of the presence or absence of certain acoustic properties, then the latter could be said to explain why we perceive chords (as well as dyads) as either consonant or dissonant. However, psychoacoustic studies are frequently beset by methodological difficulties. In this section, I will argue that the use of specific measurements as well as computational models for predicting consonance and dissonance judgements raises some important questions for the aesthetics of the sensory component of 'musical consonance'.

One of the methodological difficulties that affect psychoacoustic research on consonance/dissonance perception concerns the use of certain parameters that are unique to a particular study and thus not easily comparable to other external criteria. For instance, in one of the studies featuring in Harrison and Pearce's survey (namely: Parncutt, 1989),[22] the model used is aimed at measuring consonance in terms of pattern-matching pitch perception tasks, in order to assess the role of periodicity/harmonicity. Yet, the parameters for this measurement are unique to Parncutt's study. Specifically, he uses the criteria of 'pure tonalness' and 'complex tonalness' for predicting consonance. As Harrison and Pearce point out, while pure tonalness describes "the extent to which the input spectral components are audible, after accounting for hearing thresholds and masking", complex tonalness refers to "the audibility of the strongest virtual pitch percept".[23] But they also indicate that these measurements reflect interference and periodicity/harmonicity models respectively, so Parncutt's findings may not be so easily differentiated and, consequently, less likely to be able to be extrapolated.

As the example above illustrates, the specificities of individual studies and the methodological complexities that ensue may make it difficult for findings to be generalised. Similarly, another difficulty involving psychoacoustic findings concerns the conceptual framework under which they are presented. Researchers may put forward their own set of definitions as well as different measuring parameters – as Parncutt's notions of pure and complex tonalness indicate, for instance – which colour their findings in a particular way. As a result, this may render the possibility of drawing firm conclusions less promising.

---

[22] Parncutt, R., *Harmony: a Psychoacoustical Approach*, Berlin: Springer-Verlag, 1989.
[23] Harrison, Pearce, p. 225.

Beyond the methodological issues identified above, there are other important issues concerning the conceptual dimension of sensory consonance/dissonance that remain unaddressed in psychoacoustic studies. Specifically, the very definitions of 'consonance' and 'dissonance' are at the centre of a methodological difficulty. As part of experimental procedures, participants are frequently asked to evaluate chords in terms of their degree of consonance – even though this may often be conveyed as "whatever that means to them".[24] For instance, in a renowned psychoacoustic study, Kameoka and Kuriyagawa state that the concept of consonance is "rather vague" but, for the purposes of their experiments, they define it as a sensation of "clearness", whereas dissonance is defined as a sensation of "turbidity".[25] To add to the problem, these are the English-equivalent terms given for the original Japanese terms, so there may also be an issue of translation at times.

In addition to vagueness and cultural-linguistic specificities, the terminological issue may be coupled with the methodological difficulty of the lack of comparability mentioned earlier. Kameoka and Kuriyagawa's study is, once again, illustrative of this problem. After defining consonance and dissonance as stated above, they point out that their concept of consonance is "basically equivalent" to the concept of 'tonal consonance' in an earlier paper by Plomp and Levelt. However, the meaning that these two researchers attribute to it is in stark contrast with theirs: in Plomp and Levelt's words, consonance refers to "the peculiar sensorial experience associated to isolated tone pairs with simple frequency ratios", and tonal consonance "indicates this characteristic experience".[26]

This, however, is not an isolated instance of a terminological issue. Similar cases include definitions of 'consonance' and 'dissonance' that may be equally beset by the issue of equivocation. Amongst the studies cited earlier, consonance is frequently identified with 'pleasantness', whereas dissonance is taken to mean 'unpleasantness'. Granted, some contributions do offer alternative terms: for instance, 'consonant' is also described as synonymous with 'stable',[27] 'smooth'[28] or 'relatively attractive',[29] while 'dissonant' is defined with reference to the antonyms of those terms, respectively. Other recent attempts at definitions are particularly unclear, such as the rendering of 'dissonant' as 'out of place'.[30]

---

[24] Terhardt, 1984, p. 279.
[25] Kameoka, A., Kuriyagawa, M., 'Consonance Theory Part I: Consonance of Dyads', *The Journal of the Acoustical Society of America*, 45, 1969, p. 1452.
[26] Plomp, R., Levelt, W. J. M, 'Tonal Consonance and Critical Bandwidth', *The Journal of the Acoustical Society of America*, 38, 1965, p. 548.
[27] Johnson-Laird et al., p. 19.
[28] Lahdelma, Eerola, p. 6.
[29] Bowling, D., Purves, D., 'A Biological Rationale for Musical Consonance', *Proceedings of the National Academy of Sciences*, 112, 36, 2015, p. 11155.
[30] Cousineau et al., p. 19858.

Most strikingly, 'consonance' and 'dissonance' are at times simply given single-word definitions such as those, seemingly without any further thought as to how they may be construed by different participants in an experiment.[31] Although some studies may indicate that neither consonance nor dissonance are easily defined, they often settle for one or two of the terms mentioned above – or similar ones – as part of the methodology around the rating criteria in a given experiment. In addition, participants may be asked to rate the consonance or dissonance of dyads and/or chords in terms of their own understanding of the corresponding definitions given or, alternatively, with reference to numerical scales between pairs of opposites – such as that between 'rough' and 'smooth', for example.[32]

That said, one might argue that the putative definitions identified above are not meant to be taken literally but are, instead, given to participants solely as 'semantic glosses' under which consonance and dissonance may be grasped. Alternatively, they may be intended to be taken in metaphorical terms – which may be the case when dissonance is rendered as 'out of place'. On this view, the parameter for definition would not be strictly semantic but one that points to the comparative element in the perception of different sounds, thus enabling participants to judge them accordingly. Furthermore, it could be argued that single-word definitions are not cases of oversimplification but, rather, attempts at eliminating any potential misapprehension that may result from wordier ones.

However, this charitable interpretation of the choice of terminology in psychoacoustic studies on consonance/dissonance perception raises an important issue for an assessment of the sensory component of 'musical consonance'. While single-word definitions may be taken by some as easier to grasp, it could instead be argued that they allow not only the possibility of misinterpretation due to the vagueness of certain terms – such as 'clearness' and 'turbidity', for instance – but may also generate a strong element of subjectivity from participants – which may be evoked by defining consonance and dissonance as 'relatively attractive' or 'relatively unattractive', for example. This is particularly problematic when the main purpose of psychoacoustic studies is to establish to what extent consonance/dissonance perception may be approached objectively. Hence, this seems to point to a distinction between individual tastes – i.e. a subjective element – and preferences that may be objectively predicted across experiments. These questions are of an aesthetic character, as they concern the nature of participants' responses and the judgements they make – which is the task of the next two sections to address.

---

[31] An instance of it may be found in McDermott et al., p. 1035.

[32] This is the case in Lahdelma and Eerola's study in particular, where the numerical scale ranges between '1' ('*rough*') and '7' ('*smooth*'). The researchers attempt to justify their measurement criteria by dint of convention, seeing as the two terms have been "used extensively in research literature". Lahdelma, Eerola, p. 6.

2.3. *Philosophical issues*

The conceptual issues besetting psychoacoustic studies, as described in the previous section, indicate that adopting a theoretical framework is central to accounting for the sensory component of 'musical consonance'. Yet, such a framework is generally lacking from empirical studies – or at least not substantiated as such. In particular, those terminological difficulties are symptomatic of a wider issue regarding the kind of aesthetic responses and judgements under which consonance and dissonance should be assessed. In this section, I will consider some of the philosophical issues posed by psychoacoustic studies on consonance/dissonance perception. Specifically, I want to argue that they are of a phenomenological, epistemological and methodological nature – some of which are also relevant for my account of 'harmony' in Section 3.

First, the differing characterisations of consonance and dissonance offered in psychoacoustic studies invariably raise questions of a phenomenological nature. In particular, the most prominent one concerns what those putative definitions are meant to qualify. While some are focused on the aesthetic response of listeners – as is the case with 'pleasantness' and 'unpleasantness' – others seem to point to the quality of the sound itself – such as 'smooth' and 'stable', and respective counterparts. Depending on the approach taken in a given study, participants may be led to focus on either of the two, which may thus require that their attention be drawn *either* to the auditory experience they are having, in the case of the quality of the sound, *or* how they feel about that experience, in the case of their response to it. Instead, I want to argue that an account of the aesthetics of sensory consonance/dissonance must consider *both* the auditory sensations that listeners have – i.e. as pertaining to the quality of the sound heard – *and* the aesthetic responses that they evoke – i.e. those of pleasantness or unpleasantness, for example. Prior to substantiating this view, however, it seems important to consider the other difficulties arising from the nature of psychoacoustic research on the sensory component of 'musical consonance'.

Secondly, the aesthetics of sensory consonance/dissonance is susceptible to some epistemological difficulties that may result from the general parameters adopted in many psychoacoustic studies. Such difficulties may be observed on two levels. On a stricter reading, when participants are presented with single-word definitions of 'consonance' and 'dissonance' they would arguably limit their assessment of the stimuli with reference to their understanding of those specific terms – which, unless they are further specified, may only generate a partial assessment of consonance and dissonance.

Furthermore, on a broader reading, listeners may come to interpret those definitions in substantively different ways. This is particularly the case with the more obscure ones, such as those of 'turbidity' or 'out-of-place' for dissonance; but it is equally applicable to the more recurrent definitions that focus on the expected aesthetic responses, such as those of 'pleasantness' and 'unpleasantness'. For instance, participants may associate pleasantness with familiarity, in which case consonance and dissonance ratings may reflect individual tastes, instead.[33] Were that to be the case with other participants, it would then become questionable to what degree similar ratings as those could amount to empirical evidence for an objective assessment of consonance perception. Perhaps the greatest difficulty here lies in assessing whether those participants taking part in a given experiment have sufficient epistemic access to the concepts of 'consonance' and 'dissonance' and, in case they do, to what extent they have restricted their assessment to that specific understanding of those concepts. It is precisely for that reason that single-word definitions are problematic, in that they are arguably more likely to be construed under some form of individual taste.

Thirdly, these phenomenological and epistemological difficulties may be coupled with those arising from semantic variability in the definitions of 'consonance' and 'dissonance' as outlined in Section 2.2, thus generating further methodological issues. If different experiments provide participants with epistemic access to different conceptions of consonance and dissonance, then one may question whether findings from different studies are in any way comparable even if the data set provided – i.e. the dyads and chords used in the experiments – is the same. This difficulty is then compounded by the phenomenological distinction mentioned earlier, seeing as studies that place emphasis on the aesthetic responses of 'pleasantness' and 'unpleasantness' may not, in principle, be comparable to those that define 'consonance' and 'dissonance' with reference to the auditory sensations of 'smoothness' and 'roughness'. Furthermore, one may also question to what extent it is possible to provide a summative assessment of the psychoacoustic evidence available – such as that of Harrison and Pearce's, for example – when the parameters upon which participants have offered their ratings are so disparate. Over and above that, since the issue of equivocation persists even in the most recent attempts at accounting for sensory consonance/dissonance, psychoacoustic theories are more likely to become an easy target for those who dispute the view that those phenomena may be in any way sufficiently accounted for in terms of the acoustic base of dyads/chords.

---

[33] An exception in this particular case is found in Lahdelma and Eerola's study, in which 'consonance' and 'preference' were distinguished. Specifically, a numerical scale for participant ratings was provided for each of them individually: as indicated earlier, while 'consonance' was rated between 1 for 'rough' and 7 for 'smooth', under 'preference' 1 stood for 'low' and 7 for 'high'. Ibid, p. 6.

Indeed, these difficulties are not solely relevant for a psychoacoustic assessment of consonance/dissonance perception. Rather, they are particularly important for an account of the *aesthetic* dimension of the sensory component of 'musical consonance', in that they point to the nature of the participants' responses and of the ratings they provide when presented with dyad/chord tokens. In other words, the aesthetic dimension concerns the very nature of the core evidence upon which psychoacoustic findings are based, which is a combination of the *responses* that participants have when experiencing the sounds of certain dyads/chords and the ratings they offer on them, which express certain *judgements*.

To be clear, both of these aspects pertaining to the assessment of isolated dyads and chords are essentially of an aesthetic character. This is because the stimuli are not being assessed in terms of their acoustic or psychoacoustic properties. Rather, when participants are asked whether they have enjoyed hearing the sounds of certain dyads/chords, or whether they appreciate a distinctive quality to those sounds, they are being asked to evaluate their own aesthetic experiences. This, however, goes beyond the presence of certain physical or psychoacoustic properties in dyads/chords. Hence, in the next section I will consider the nature of aesthetic responses *to* and judgements *of* dyad/chord tokens.

2.4. Responses and judgements

Having identified some important issues above, my attempt to address them in this section is twofold. First, I want to establish a fundamental distinction in the aesthetics of sensory consonance/dissonance, which concerns the nature of *judgements* that participants make when they assign specific ratings to certain auditory sensations, on the one hand, and the kind of *responses* evoked by these sensations, on the other. Subsequently, I will seek to account for those aesthetic responses in light of the notion of aesthetic pleasure. I will suggest that pleasantness is a distinctive aspect of the aesthetic experience of sensory consonance – while the contrary is the case for sensory dissonance.

First, it seems essential that sensory consonance/dissonance be assessed *both* in terms of the quality of the sound of dyads/chords as they are heard *and* aesthetic responses to them. In the case of the sensory component of 'musical consonance', I will take 'smooth' or 'rough' to refer to auditory sensations, whereas 'pleasant' or 'unpleasant' pertain to responses, instead. While some psychoacoustic studies may include references to both types of qualification – e.g. as a combination of 'smooth' and 'pleasant' – the question of what kind of experience participants are being asked to qualify does not generally arise.

Instead, I want to argue that there is an important distinction between the auditory experience that participants have and how they feel about such an experience. While the latter points to the nature of the aesthetic response, it is the former that, I believe, is the proper object of aesthetic judgement. Although this distinction may not seem to be of interest to those who undertake psychoacoustic research, the considerations made in the previous sections seem to justify the view that it should be. By not making that distinction clear to participants – or by tacitly assuming they should be aware of it – one may reasonably question to what degree consonance and dissonance ratings can be adequately correlated with the parameters stipulated for a given experiment.

For that reason, I propose that such parameters must distinguish between the two kinds of experience. The criteria given to participants should enable them to consider *both* the auditory sensations of specific sounds they are presented with *and* the aesthetic responses that are evoked by them. However, this does not mean that the two should be taken as independent of one another. Rather, the main reason for providing a clear distinction between them is so that participants are aware that one concerns the quality of the sound as it is perceived, whereas the other points to their aesthetic responses to it. This methodological solution is also relevant for my account of the metaphysics of the sensory component, so it will be incorporated into the analysis I offer in Chapter Nine (Section 2).

The second part of my argument concerns the very nature of those aesthetic responses and judgements with regard to consonance/dissonance perception – which also has a bearing on the account I shall give in the next section on the aesthetics of 'harmony' (i.e. the distinctively musical component of 'musical consonance'). First, in the case of aesthetic judgements of dyads and chords, there seems to be a conceptual correspondence between judgements and the auditory sensations that constitute their objects – in keeping with my earlier claim as to how they should be understood. Specifically, when participants rate the sounds of dyads/chords as either 'smooth' or 'rough', their ratings are taken to match the auditory sensations of smoothness and roughness. In turn, as seen earlier, smoothness and roughness tend to be attributed to the presence of harmonicity or destructive interference, respectively. It is, therefore, under this neat correspondence that psychoacoustic researchers are able to predict judgements of consonance and dissonance *vis-à-vis* the non-aesthetic base of dyads/chords. Were it not to be the case, there would be little use for computational models of consonance/dissonance perception. Most importantly, this correspondence is what is taken to enable experimental data to be converted into psychoacoustic findings on the nature of sensory consonance/dissonance.

To that extent, it seems that the explanatory power of psychoacoustic theories is indeed observed in predicting ratings of isolated dyad/chord tokens. Yet, I want to argue that the distinction between judgements of auditory sensations, on the one hand, and participant responses to them based on how they feel about the experience, on the other, should nonetheless be observed. This is so that judgements – i.e. the descriptive component – and responses – i.e. the evaluative component – may be correlated as such. Some of the difficulties identified earlier may therefore be addressed by making a clearer specification of the content and purpose of aesthetic 'verdicts' of sensory consonance or dissonance, which I will take to comprise *both* judgements *and* responses. But the case of aesthetic responses to consonance and dissonance requires further analysis, since I take them to be rooted in experiences of aesthetic pleasure – or the lack thereof.

As pointed out earlier, aesthetic responses to consonance and dissonance are most frequently characterised in terms of pleasantness and unpleasantness, respectively. While pleasantness is taken to result from an auditory sensation of smoothness, unpleasantness is frequently associated with that of roughness. But there does not seem to be the same kind of conceptual correspondence between pleasantness and smoothness (or between unpleasantness and roughness) as the one observed in aesthetic judgements. This may also be gleaned from considering other descriptions of auditory sensations associated with consonance and dissonance – e.g. in terms of 'stability' or 'instability' – which do not seem to elicit any obvious correspondence with pleasantness or unpleasantness, either. Hence, one may question whether a given consonance or dissonance judgement *must* be accompanied by aesthetic experiences of pleasantness or unpleasantness, respectively.

In order to address this question, I want to consider the nature of aesthetic responses to consonance/dissonance in light of the notion of aesthetic pleasure. While it is not my intention to enter into a broader debate regarding this notion, it seems important to consider some views which I take to be illuminating for my account in this section. In her attempt at characterising aesthetic appreciation, Anne Sheppard describes aesthetic pleasure as its emotional element, which can vary from mild expressions of pleasure to those of rapturous enthusiasm.[34] In particular, the most typical aspects of that experience seem to be that aesthetic pleasure manifests itself as a desire to continue or repeat the experience, and that doing so is taken to be an end in itself. Conversely, Sheppard notes that there are also negative aesthetic responses, which can arouse the reverse of pleasure (albeit to varying degrees) and which may be expressed as a desire to repel the offending object.

[34] Sheppard points out that this emotional element must be distinguished from those emotions that are not of a detached (i.e. typically aesthetic) nature. Sheppard, A., *Aesthetics*, Oxford: OUP, 1987, pp. 64-65.

One way of accounting for the experiences of consonance and dissonance is in terms of a distinction between positive and negative aesthetic responses in the manner of Sheppard's view. On the one hand, consonance may be taken to evoke a positive response, which is best described in terms of pleasantness, whereas the reverse applies in the case of dissonance. Underscoring this distinction is the notion of aesthetic pleasure, such that consonance may be heard as having an attractive quality to it, whereas dissonance could be said to have a repellent nature, instead. But, although I take this to be the appropriate characterisation in the specific case of the sensory component of 'musical consonance', it seems that an account of aesthetic responses to consonance/dissonance must also accommodate the possibility that auditory sensations taken to be consonant or dissonant may *not* evoke the responses of pleasantness and unpleasantness, respectively – which is particularly relevant for my account of the aesthetics of 'harmony' in the next section.

Granted, I take aesthetic pleasure to be central to what it means to have a positive aesthetic experience. In supporting this view, my account is aligned with Rafael De Clercq's analysis of aesthetic pleasure, which places it as essential to the experience of beauty.[35] Yet, De Clercq distinguishes between what he calls general and specific cases of such an experience, thus offering a more nuanced view. As he describes them, while the specific sense pertains solely to cases in which non-disturbing forms of beauty are pleasurable to perceive, the general sense refers to those in which both disturbing and non-disturbing forms of beauty are experienced as pleasurable.[36] This would enable us to address the question posed earlier since, under the general sense, a dissonance judgement would not in principle have to be accompanied by an aesthetic experience of unpleasantness.

Borrowing De Clercq's distinction, I take the specific sense to apply in cases where participants experience pleasure in perceiving consonance, whereas the general sense involves the possibility that they may enjoy hearing dissonance (i.e. the disturbing form). Given that there is no intrinsic connexion between dissonance and unpleasantness, it is then possible to judge the sound of a given chord to be dissonant without the unpleasantness that is often associated with it. But, although there does not seem to be anything paradoxical about that possibility in the case of 'harmony', I take this to be counterintuitive with regard to the sensory component. This is because dissociating sensory dissonance from a response of unpleasantness is tantamount to challenging psychoacoustic evidence at its very core – a stance which I do not support, as indicated in Chapter Seven.

[35] De Clercq, R., 'Aesthetic Pleasure Explained', *The Journal of Aesthetics and Art Criticism*, 77, 2, 2019, p. 129. However, De Clercq defends the Identity Thesis, under which the experience of aesthetic pleasure is identical to the experience that something appears to be beautiful – a commitment I find unnecessary. Ibid., p. 124.

[36] Ibid., p. 122.

In view of the above, I will take aesthetic responses to sensory consonance to be best characterised under the specific sense of aesthetic pleasure. This is because pleasantness seems to be a distinctive aspect of the aesthetic experience of sensory consonance,[37] without which it would be questionable whether the auditory experience of dyads/chords can be said to be one of consonance if it is restricted to describing it based solely on the quality of the sound heard (e.g. its smoothness). In the context of psychoacoustic experiments, if participants are not given the opportunity to rate their experiences in terms of how they feel about those auditory sensations, it also becomes questionable whether the respective findings should be taken as indicative of how consonance and dissonance are experienced, rather than merely of how dyads and chords may be described, instead.

Yet, as pointed out earlier, although I have distinguished between aesthetic responses and judgements with regard to consonance/dissonance perception, I do not take them to be independent of one another in the sensory case. Hence, I take *both* judgements of dyads/chords *and* responses to their sensory consonance/dissonance to be essential to the response-dependence biconditionals which will be formulated in Chapter Nine. The same, however, cannot be said of the case of 'harmony', to which we turn in the next section.

3. The aesthetics of 'harmony'

Having considered the main aesthetic concerns regarding sensory consonance and dissonance *vis-à-vis* the psychoacoustic base of dyads/chords, the focus of the present section is on how those properties may be perceived within a distinctively musical context. In order to account for the distinction between the two components of 'musical consonance', I will first outline the different approaches to consonance and dissonance identified by Rehding. These include the contextualist view of harmony, which has been defended by Scruton. After dismissing what I take to be an apparent conflict between music and psychoacoustics, I will argue that the contrast between experiences of consonance and dissonance in a distinctively musical context cannot be assessed on the same terms as participant responses to isolated dyad/chord tokens, as described in the previous section. Instead, I will maintain that both consonance and dissonance have a dual phenomenology, which is what warrants the distinction between the sensory component and 'harmony'. It is under this proposal that I will argue that 'harmony' is context-sensitive.

[37] Similarly, I will take unpleasantness to be distinctive to aesthetic responses to sensory dissonance.

3.1. Three approaches

As seen in Chapter Six, Rehding identifies three ways in which consonance and dissonance may be understood, i.e. acoustically, physiologically/psychologically and as context-dependent. Yet, Rehding has more recently offered a different conceptual overview of those two phenomena, which is the purpose of this section to consider. Specifically, he characterises the three approaches to consonance/dissonance in terms of events, voice-leading and tension/relaxation, respectively.[38]

First, Rehding describes the 'event approach' as pertaining to consonance and dissonance *qua* isolated acoustic phenomena – an approach which he deems a useful fiction. In particular, Rehding rejects the harmonicity hypothesis, which he takes to be a helpful shortcut but dismisses as dubious. Instead, he argues that the foundational role of the harmonic series as a basis for consonance perception is arbitrary. To illustrate this claim, Rehding points to the fact that the seventh harmonic in a given overtone series tends to be left out of consideration, whilst also questioning the ability of the harmonicity hypothesis to account for historical and cultural differences in consonance perception.[39] For Rehding, there is thus an inherent tension between psychoacoustic accounts and music-theoretical approaches, which he takes to result from irreconcilable assumptions.

Secondly, Rehding describes consonance/dissonance in terms of voice-leading rule.[40] This approach is based on the rules of counterpoint, under which dissonances are expected to be resolved into consonances. For instance, Rehding argues that the classification of certain intervals as imperfect consonances is indicative of the context-based nature of consonance and dissonance. He notes that although thirds, sixths, and tenths were traditionally deemed dissonant when considered as isolated intervals, they were not avoided in certain voice-leading contexts. In particular, Rehding makes reference to a thirteenth-century music theorist who "didn't mince his words when he called the major sixth in itself a *vilis discordantia*, a vile dissonance, which immediately becomes an *optima concordantia*, the best consonance, if it precedes the octave".[41] Rehding takes this to illustrate the fact that the event and voice-leading approaches depart from one another and are thus not fully congruent. In his view, the upshot of the voice-leading approach is that consonance and dissonance are interrelated musical phenomena, rather than isolated acoustic events.

---

[38] Rehding, 2018, pp. 438-440.
[39] Ibid., pp. 441-442.
[40] Karp defines 'voice leading' (or 'part writing') as "the progression of each of the several voices of a work from one tone to the next with proper regard for the relationship between the various voices". Karp, p. 431.
[41] Rehding, p. 448.

The third approach identified by Rehding is that of consonance and dissonance conceived in terms of the notions of tension and relaxation. To understand consonance and dissonance in this way is to take the contextualist view of harmony even further. This is because chords are not said to be consonant or dissonant in virtue of their acoustic properties; rather, it is the dynamics of tension and relaxation in a given musical passage that 'colours' how they are perceived. This has been a more popular approach to consonance/dissonance, especially in view of certain developments in the understanding of harmony that culminated with the movement towards atonality in the twentieth century.

As Daniel Harrison notes, this approach originated with the emergence of chromatic music in the late nineteenth-century. Specifically, it stems from an understanding that the "fundamental sensations of harmonic tonality could be separated from the sounding entities that traditionally produced those sensations", which meant that the correspondence between the psychoacoustics of consonance/dissonance and tonal harmony was found to be "withered and unnecessary".[42] Instead, Harrison points out that chords came to be described in terms of their harmonic function, such as that of the 'tonicness' of any given tonic triad.[43] This dissociation between the psychoacoustic and the musical views of consonance/dissonance is what underpins the tension/relaxation approach to harmony.

Traditionally, chords that were deemed dissonant were used to create tension, while consonant chords were subsequently deployed in order to generate an auditory sensation of repose. But under those changing attitudes to harmony, harmonic structures became less 'law-like', thus enabling chord progressions to gain new forms of expression. As Arnold Whittall points out, a clear consequence of the third approach is the notion of the 'emancipated dissonance', under which both dyads and chords that were traditionally seen as dissonant came to be treated as "relatively stable harmonic entities, functioning in effect as 'higher' or more remote consonances".[44] Hence, on this view, harmony is essentially functional, to the extent that any categorical distinction between consonance and dissonance becomes arbitrary. This firmly points to the view that harmony is context-sensitive, to the detriment of the several psychoacoustic attempts at establishing causal explanations for consonance/dissonance perception. In particular, a version of this contextualist view was developed by Scruton, in his account of the aesthetics of music.

---

[42] Harrison, D., *Harmonic Function in Chromatic Music*, Chicago: University of Chicago Press, 1994, p. 11.
[43] Harrison's example here is drawn from Hugo Riemann's harmonic theory. Ibid., p. 11.
[44] Whittall, A., 'Consonance and Dissonance', In Latham, A. (ed.), *The Oxford Companion to Music*, OUP, 2014, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199579037.001.0001/acref-9780199579037-e-1581. As Tenney notes, the 'emancipation of the dissonance' was first heralded by Schoenberg. A similar view was echoed by Stravinsky, who resisted the need of resolution when he claimed that "nothing forces us to be looking constantly for satisfaction that resides only in repose". Tenney, p. 3.

3.2. The context-dependence of harmony

In this section, I will briefly characterise the contextualist approach to harmony proposed by Scruton, whose account of the ontology of chords was considered in Chapter Four. In line with the Gestalt-based view of chord perception outlined in Section 2 of that chapter, Scruton describes the perception of harmony as hearing not just two or more sounds occurring simultaneously but also hearing "a relation between them, and sometimes a new entity that is formed by their conjunction".[45] Yet, he argues that this distinctive unity of tones must be considered in tandem with its relations to other sounds, in view of the context in which they occur (e.g. a given musical passage of a given composition). For that reason, Scruton argues that consonance is context-dependent, and therefore not in any way determined by the acoustic base of chords. In particular, he states that the same goes for the musical experience of dissonance – thus casting it under a whole new light.

Underlying Scruton's thesis of the context-dependence of harmony is the belief that, in strictly musical terms, chords are not intrinsically consonance or dissonant. Similarly to what Tenney indicated with his 'consonance/dissonance concept', Scruton points out that the history of Western music is filled with examples that testify to that view. For instance, he notes that fourths were largely deemed consonant in mediaeval harmony, only to be subsequently considered dissonant in the Renaissance period. Conversely, major and minor thirds, which were originally taken to be dissonant by early mediaeval musicians, became imperfect consonances around the twelfth century, and were then later seen as paradigms of consonance in the classical period.[46] For Scruton, these examples indicate that consonance and dissonance are not intrinsic properties of certain harmonic intervals.

Similarly, chord patterns used by different composers may sound either consonant or dissonant, depending on the harmonic context within a given piece. Indeed, it may specifically be their intention to use dissonant chords, for instance, in order to generate tension and then subsequently bring about resolution with a chord that would be perceived as consonant in relation to it, thus enabling the dissonance to 'vanish'.[47] This being the case, Scruton argues that this phenomenon demands an explanation, as do those variances in chord perception identified above. In his view, such an explanation can only be provided by considering the specific musical context in which they occur – to the detriment of the psychoacoustic view on consonance/dissonance perception.

---

[45] Scruton, 1997, p. 64.
[46] Ibid., p. 241.
[47] Please see footnote no. 33 on page 77 for a definition of resolution.

3.3. Music versus *psychoacoustics?*

Having considered the contextualist approach to harmony above, the question of the relationship between music and psychoacoustics thus comes to the fore. Under the accounts put forward by Rehding and Scruton, the distinctively musical experience of consonance/dissonance is taken to be at odds with the psychoacoustic approach to tonal harmony. Yet, contrary to this view, I will propose in this section that it is possible to bridge the divide between the psychoacoustic and the distinctively musical views on consonance/dissonance, whilst simultaneously preserving them as independent domains. Specifically, I will argue that it is the foundational principles of traditional tonal harmony that enable them to be reconciled as the two components of 'musical consonance'.

The contextualist view of harmony is normally taken to be irreconcilable with the psychoacoustic approach to consonance/dissonance perception. For instance, just as Rehding dismisses the harmonicity hypothesis, Scruton equally rejects Helmholtz's theory of musical acoustics. As seen in Chapter Seven, the latter came to be associated with a naturalistic view of harmony. In describing Helmholtz's approach as such, Scruton points to the apparent importance of the triad as "the 'natural' harmony *par excellence*, since its existence is implied in the tone itself".[48] This is because the three component tones of any given triad appear as the fourth, fifth and sixth harmonics of the overtone series of the fundamental. But, by way of contrast, Scruton argues that "a major triad in the bass will generate more conflicting overtones than a minor ninth in the upper register; nevertheless, we hear the first as consonant and the second as dissonant".[49] Hence, Scruton rejects the claim that any laws of harmony may be derived from psychoacoustics – as well as any strict observance of precise mathematical ratios for certain intervals, for that matter.

Although Scruton acknowledges the familiarity of the psychoacoustic views of consonance and dissonance, he maintains that this type of explanation does not account for the experience of harmony itself. His view is based upon the distinction between concords and discords, on the one hand, and consonance and dissonance, on the other. For Scruton, while the former concerns solely the acoustic realm, consonance and dissonance are specifically musical phenomena.[50] Perhaps unsurprisingly, this dichotomy is in direct correlation with his acousmatic thesis, as outlined in Chapter Five (Section 2).

---

48 Scruton, 1997, p. 243.

49 Although Scruton acknowledges that this phenomenon "remains unclear", he suggests that it is in terms of the context-dependence of dissonance that these discrepancies are best understood. Ibid., p. 244.

50 Ibid., p. 64.

In Scruton's view, the musical experience is distinctively acousmatic, in that musical sounds can be perceived as "emancipated from their causes".[51] On this view, chords *qua* musical individuals are heard acousmatically, which may be observed when they are perceived as having a particular harmonic function within a musical context. On that basis, Scruton takes the distinctively musical experience of chords to be independent from their acoustic base. His acousmatic thesis can thus be said to enable the emancipation of harmony from any deterministic psychoacoustic explanations. Yet, although I will argue in the next section that 'harmony' *is* context-sensitive, my argument in the present section runs counter to the view that consonance and dissonance are emancipated from psychoacoustics in the way that the contextualist approach seems to entail. Specifically, two main issues seem to arise from the accounts given by Rehding and Scruton, respectively, which are particularly relevant for my proposal in the next section.

First, in dismissing the importance of the harmonicity hypothesis for consonance perception as useful fiction, Rehding is effectively challenging the growing consensus in psychoacoustic circles that harmonicity does indeed play a causal role in the aesthetic experience of chords *qua* isolated auditory objects, as detailed in Section 2. It seems unlikely that a plethora of psychoacoustic studies would have continued to be undertaken to this day had the harmonicity hypothesis been confirmed to be nothing more than a useful fiction. This is particularly striking when one considers that Harrison and Pearce's study, for instance, is just as recent a contribution to the debate on consonance/dissonance perception as Rehding's – and one amongst many. Hence, the fact that psychoacoustic research on the experience of harmony still generates academic interest is, at the very least, a sufficient reason for the psychoacoustic component not to be dismissed as mere fiction.

Secondly, as Ridley correctly points out in his assessment of Scruton's acousmatic thesis, the possibility of hearing sounds as detachable does not entail that they are necessarily detached from their sources.[52] Analogously, the possibility of hearing consonance and dissonance as independent from the acoustic base does not entail that they cannot be ascribed to dyads and chords in virtue of their acoustic properties. Furthermore, Scruton's outright rejection of Helmholtz's theory on the basis of a strictly contextualist approach to harmony is tantamount to a dismissal of the interference hypothesis, in that Helmholtz is perhaps its most famous advocate. However, as seen in Section 2 above, Harrison and Pearce's survey of the latest psychoacoustic evidence points to interference as the 'highest-scoring' causal mechanism underlying dissonance perception.

[51] Scruton, 2009, p. 22.
[52] Ridley, p. 54.

In view of the above, I want to argue that music and psychoacoustics may be reconciled. Specifically, I take the foundational principles of traditional tonal harmony to testify to that claim. As indicated in Chapter Six (Section 2), traditional tonal harmony was originally forged on the basis of a correspondence between an understanding of the acoustic properties of certain intervallic relationships and the rules of counterpoint.[53] Although Rehding points out that the concepts used in music theory cannot be fully 'mapped out' on to psychoacoustic principles,[54] what he seems to ignore is that the very fact that a certain correspondence is indeed observed is of metaphysical significance, whether or not this may be the case for *all* such musical concepts. In other words, even if there may be cases in which musical and psychoacoustic principles may not converge, this does not preclude the fact that the many cases in which they are congruent demands an explanation. Indeed, by stating that "consonance and dissonance are in the beholder's ears",[55] Rehding's subjectivist approach dismisses the underlying nature of the foundations of traditional tonal harmony, which is underpinned by those cases of congruence. Contrary to his belief, I hold the view that the historical changes in the attitudes to harmony cannot be seen to reduce the relationship between music and psychoacoustics to mere fiction.

Similarly, Scruton's rejection of any relationship between psychoacoustic principles and the musical experience of harmony falls prey to the same difficulty identified above. Granted, his distinction between concords/discords and consonance/dissonance may be interpreted in a way that is compatible with the account I have been pursuing in Part II – which will be substantiated further in Chapter Nine. Yet, by taking concords and discords to be strictly acoustic rather than musical sounds, Scruton's view not only detracts from any possible metaphysical ties between music and psychoacoustics but also fails to acknowledge the ways in which the two domains are seen to converge in traditional tonal harmony. While psychoacoustics alone cannot account for the aesthetics of harmony, it seems that the relationship between the sensory component and traditional tonal harmony is not just a token of some form of tentative connexion between psychoacoustics and music. Rather, as indicated by the outline of Rameau's views I offered in Chapter Six, it points to the foundation of traditional tonal harmony at its very core.

53 As indicated by the use of computational models in psychoacoustic research, compositions from the contrapuntal/figured-bass period show evidence that composers consistently sought to avoid instances in which the overlapping voices entailed intervals traditionally heard as dissonant. For instance, Harrison and Pearce note that, in a particular study, the analysis of thirty of J. S. Bach's polyphonic works indicated a dual concern, namely, that of minimising both interference – because of its negative aesthetic valence – and the perception of tonal fusion – in order to preserve the perceptual independence of the different voices. Harrison, Pearce, p. 224.

54 Rehding, 2018, p. 447.

55 Ibid., p. 461.

At the heart of this apparent rivalry between music and psychoacoustics is the problem of consonance, as outlined in that same chapter. On the one hand, those who favour a music-oriented approach seem to treat consonance and dissonance fundamentally as context-dependent properties, in virtue of the harmonic function that dyads and chords perform within a given musical passage. On the other hand, those who accept the psychoacoustic account tend to treat consonance and dissonance as categorical properties instead, which are closely related to the nature of the acoustic base. Accounting for this disparity is thus fundamental to the metaphysics of tonal harmony, which is the focus of the next chapter. But, in order to assess the contrast between the sensory component and 'harmony' satisfactorily, it seems important to consider the question of the nature of aesthetic responses to dyads/chords and respective judgements of consonance and dissonance once again – which is what I shall attempt to do in the next section.

3.4. A dual phenomenology

In Section 2.4, I argued that a distinction must be drawn between the content of aesthetic judgements and the nature of aesthetic responses to consonance and dissonance. While judgements concern the sonorous quality of dyads and chords (e.g. whether they are rough or smooth), the responses are better understood via the notion of aesthetic pleasure – or, more precisely, in terms of pleasantness and unpleasantness. The corollary of my argument in that section was that both the auditory sensations of smoothness/roughness and the aesthetic responses of pleasantness/unpleasantness must be considered in psychoacoustic assessments of consonance/dissonance perception.

Although the proposal summarised above was outlined in the context of the sensory component in particular, the issues identified in Sections 2.2 and 2.3 have similar consequences for the aesthetics of 'harmony'. While it is not my intention to reiterate those earlier concerns individually, it seems important to consider the ways in which they may affect an understanding of the distinctively musical component of 'musical consonance', especially in view of the contextualist approach described in Section 3.2. To that end, there are two main difficulties that similarly arise for the aesthetics of 'harmony', which I will aim to address in this final section. Most importantly, I will substantiate my thesis of the dual phenomenology of consonance and dissonance as warranting the distinction between the sensory component and 'harmony'. It is under my account in this section that I will characterise the perception of 'harmony' as being fundamentally context-sensitive.

First, the issue of equivocation that besets psychoacoustic studies on consonance and dissonance at large can equally be found under the musical component – and arguably to an even greater degree of ambiguity. For instance, in his account of the history of 'dissonance', Tenney points out that the varying descriptions of it do not particularly seem to illuminate our understanding of the concept in any way. He asks: "are we to interpret *dissonance* ... as meaning 'not complete in itself', as implied ... by Stravinsky, or as less 'comprehensible' (Schoenberg), less 'euphonious' (Hindemith), less 'agreeable' or 'pleasant', ... more 'active' or 'unstable' (Kraft and others), etc. – or as some combination of some or all of these meanings?".[56] Furthermore, neither does the notion of the emancipation of dissonance seem helpful in furthering the understanding of 'dissonance', since it is used to describe it in terms of what it is *not*. Hence, Tenney's concern is indeed justified; yet, as I indicated in Chapter Six, this goes beyond the semantics of 'dissonance'. Specifically, as seen in Section 2.3, the issue of equivocation has both phenomenological and epistemological consequences which have a bearing on how listeners respond to an auditory sensation of roughness and make judgements about it.

Dovetailing with that claim, the second difficulty concerns both the nature of aesthetic responses to consonance and dissonance and the kind of judgements that listeners may come to make of the sounds of dyads and chords *qua* musical individuals. However, although the issue of equivocation raises difficulties for both the sensory and the musical components of 'musical consonance', the consequences of it for aesthetic responses and judgements in those respective domains are markedly different. It is on that basis that I want to defend my proposal of the dual phenomenology of consonance/dissonance, which is at the centre of my solution to the problem of consonance.

To be clear, the phenomenology of consonance/dissonance perception as pertaining to the sensory component is in stark contrast to the musical experience of those phenomena in the context of 'harmony' – which I will characterise as context-sensitive. In order to highlight this contrast, it is essential to consider once again the nature of aesthetic judgements and aesthetic responses to consonance/dissonance – both of which, I believe, must be incorporated into an account of the aesthetics of tonal harmony for it to be satisfactory. In the case of judgements of sensory consonance and dissonance in particular, I argued that it is the sonorous quality of a given dyad/chord that is their proper object, which can be correlated with the sensations of smoothness and roughness, respectively.

56 Tenney, p. 3. Perhaps unsurprisingly, the concept of consonance is equally fraught with issues of equivocation. For the sake of an overall balance in my analysis of the two phenomena, my argument here is focused on dissonance, in particular.

Although the characterisation above is best equipped to account for aesthetic judgements in the sensory case, the same cannot be said of its musical counterpart. Indeed, the auditory experience of isolated dyad/chord tokens is not replicated as such when they are perceived within a musical passage. This is more clearly observed in instances of dissonance perception, in particular – both under a conception of dissonance in terms of voice-leading rule and that of the contrast between tension and repose, as outlined earlier. For instance, in the distinctively musical context of polyphony and, more specifically, that of counterpoint, the notion of resolution is fundamental to dissonance perception. This is because a given dyad or chord that is heard as dissonant when taken as an isolated auditory object – such as a major seventh interval or a diminished triad, for instance – tends to be perceived differently when it is resolved by a consonant dyad or chord.

Under a more orthodox notion of resolution in counterpoint, dissonance perception is affected by the perceived relationship between musical sounds, such that the roughness of a discord is (to a greater or lesser degree) attenuated by the smoothness of a concord that follows it. Yet, this may also be observed under a conception of dissonance that is based on the relationship between tension and relaxation. Any given discord that is used to generate an auditory sensation experienced as tension may be resolved either by a concord or a relatively less dissonant chord, which brings about an auditory experience of repose. Hence, unlike in the case of the sensory component, the perception of a dissonant chord in relation to other neighbouring sounds considerably alters the phenomenal quality of that sound. This phenomenological contrast between sensory dissonance and the musical perception of it is particularly relevant for my account. Specifically, it is precisely the disparity between the two that, I believe, warrants any talk of the 'emancipated dissonance'.

As seen earlier, Schoenberg's 'emancipation of the dissonance' was taken to represent a significant blow to the traditional understanding that dissonances must be resolved into consonances. Two important consequences of this shift concerning the notion of resolution, in particular, point to the phenomenological contrast I have sought to characterise above – and this may be understood on two levels. First, dissonances can be said to be emancipated in the sense that they need not be resolved by a consonant chord or interval. On this view, if dissonant dyads/chords may be resolved by less dissonant ones, then resolution is not exclusively a function of consonant dyads or chords. Hence, this development in the conception of dissonance points to a particularly revealing aspect of the phenomenology of 'harmony' that is not shared by its sensory counterpart: namely, that the property of 'resolvability' is distinctively musical.

This claim has important implications for the metaphysics of 'harmony', which will be considered in the next chapter. Most importantly for the present section, however, is the second consequence of the emancipation of dissonance for the notion of resolution which is equally indicative of the phenomenological contrast between sensory dissonance and its musical counterpart. If discords are taken to be emancipated from the need for resolution altogether, then the roughness of the corresponding auditory sensation in a given context may therefore be 'embraced' as such. Specifically, it is the harmonic role of discords as tension-generators that may, instead, be taken as their defining phenomenal quality – as can be observed in a chord progression that ends without any seeming resolution.

Hence, this elimination of what I will refer to as the 'resolution constraint' indicates that the phenomenological contrast between the musical and the sensory case could not be overstated. Rather than being perceived as reverting back to its default (i.e. sensory) position, a dissonant dyad or chord may acquire a specific harmonic role in a given context that is distinctively musical; and yet, its role in creating tension need not be associated with an auditory sensation of roughness or the orthodox expectation that it must be resolved. On this view, therefore, both the notion of resolution and the very nature of the sensory-based correspondence between the auditory sensation of roughness and the aesthetic judgement of dissonance are called into question.

In view of the above, I take these instances of the phenomenological contrast between the sensory and musical components of harmony to be indicative of a substantial case for supporting my proposal for the dual phenomenology thesis, which has been exemplified above in the case of dissonance, in particular.[57] Having illustrated the contrast between aesthetic judgements on sensory dissonance and on a distinctively musical perception of it, I shall henceforth continue to substantiate my proposal by considering the *experience* of consonance and dissonance in order to characterise the contrast between aesthetic responses in the psychoacoustic and musical domains, respectively. In the specific case of aesthetic responses, I indicated in Section 2.4 that the main questions around it concern the nature of the response itself, on the one hand, and how it may be correlated with aesthetic judgements of consonance and dissonance, on the other. In order to highlight the phenomenological differences between the sensory and the musical components, I shall consider how those two questions may be addressed in the musical case – and the ways in which such treatment is different from the sensory one.

[57] As would be expected, the phenomenology of consonance perception is to be conceived in similar terms. As pointed out earlier, once the resolution constraint is removed, consonant dyads/chords are deprived of their function of resolving dissonances. As Rehding notes, with the emancipation of the dissonance, the consonance "no longer has any power to resolve anything". Rehding, 2018, p. 457.

First, one may object that neither pleasantness nor unpleasantness may be the most adequate ways to characterise responses to the phenomena of consonance and dissonance, respectively. Yet, as seen in the previous section, this tends to be the basis for participant ratings in psychoacoustic experiments with isolated dyads and chords. While this may not be problematic for an assessment of the sensory component of 'musical consonance', as I have argued then, the case of 'harmony' raises a few questions of its own. These similarly point to the very notion of aesthetic pleasure in consonance and dissonance perception and, most importantly, how they are experienced when dyads or chords are perceived *qua* musical individuals in a particular context.

As described in the previous section, I take typically aesthetic responses to have an evaluative character and to involve an element of attraction or repulsion owing to the nature of the object being contemplated. If the aesthetic response to consonance in the context of 'harmony' is conceived in terms of pleasantness, as per the sensory case, then one would expect that a musical passage that comprises only consonances would be deemed most pleasant. However, this does not seem to be the case. Even prior to the advent of chromatic music, compositional practice of the formative period of traditional tonal harmony invariably included dissonant elements, which is an essential aspect of the auditory experience of resolution. Thus, although a musical passage filled only with consonances may generate the auditory experience of harmonic progression, it does not evoke that of resolution. This is particularly relevant, given that resolution – whether the traditional form or that based on the relationship between tension and relaxation – can be said to be one of the most pleasurable aspects of the phenomenology of 'harmony'.

Hence, the notion of resolution is once again at the centre of the phenomenological contrast between the sensory and the musical components of 'musical consonance' – and in particular, of dissonance. Rather than shunned as unpleasant, dissonances are thus fundamental to co-manifesting the property of 'resolvability' that is a distinctively musical quality. Whether it is perceived as only attributable to consonances or to any given dyad or chord, there cannot be resolution as such if there is no tension created by a certain degree of dissonance. However, this dissonance is not typically experienced as unpleasant – quite the contrary. Indeed, what the changing attitudes to the understanding of harmony – as epitomised by the emancipation of the dissonance – point to is that the tension generated by certain dissonant sounds in a given musical context is frequently expected to evoke aesthetic pleasure, whether or not they are subsequently resolved.

For instance, a case in point is that of Wagner's *Tristan und Isolde*, where tension is exacerbated and sustained over long periods and resolution is repeatedly delayed until an eventual musical 'climax'. Yet, far from being an experience that repels, the element of tension created by dissonant chords in that musical context evokes a response rooted in aesthetic pleasure – which, in the sensory case, is associated with consonant sounds, instead. Alongside several other examples, this testifies to the claim that the strict association of dissonance with unpleasantness, which is commonplace in psychoacoustic studies, cannot be extended to apply to the phenomenology of 'harmony'.

Furthermore, with regard to the second question raised earlier, the contrast between the phenomenology of the sensory component and that of its musical counterpart can be equally observed by considering the correlation between auditory sensations, aesthetic judgements and aesthetic responses. In the sensory case, I have argued that 'consonant' and 'dissonant' *qua* aesthetic judgements are warranted by the correspondence between the auditory sensations of smoothness and roughness and the sonorous quality of their respective objects. In addition, I have previously proposed that psychoacoustic theories should restrict aesthetic judgements specifically to the rating of the perceived quality of the sounds of dyads or chords, rather than how listeners feel about those auditory experiences. But, if we 'transpose' this understanding to the musical case, some counterintuitive consequences would seem to follow.

On the one hand, if the harmonic function of a dissonance within a musical context is taken to be that of preparation for resolution – as per the voice-leading view identified by Rehding – then the auditory sensation of roughness would arguably be experienced specifically in terms of the anticipation of the auditory experience of resolution. Yet, this raises the question as to whether listeners would even judge the sonorous quality of dyads or chords as rough, since the auditory experience of anticipation is what would seem to prevail in a distinctively musical context.[58] On the other hand, on a more traditional understanding of the role of consonances, their function is precisely that of bringing about an auditory experience of repose; yet, this need not be described in terms of smoothness. Instead, in virtue of the relationship between preparation and resolution, consonances seem 'welded' to the dissonant sounds that precede them, to the extent that they both need that relationship in order to acquire their respective qualities. This is what is normally referred to as the 'need of resolution' – which, like resolution itself, is a distinctive aspect of the phenomenology of 'harmony' that is not shared by that of the sensory component.

[58] A clear example of it, as noted earlier, is that of the major sixth, which was once described as a 'vile dissonance' that immediately becomes the 'best consonance' when it is followed by the octave.

Most importantly, this also points to a peculiar feature of the contrast between the sensory and the musical components, which is more clearly observed when we consider the conceptions of dissonance and consonance in terms of tension and relaxation, respectively. If the harmonic function of dissonances is taken to be essentially that of generating tension, and that of consonances to convey an auditory experience of relaxation, then the correspondence between roughness and smoothness could arguably be maintained. However, if dissonances are experienced as being resolved by other less dissonant dyads or chords, then this correspondence is violated. This is because the less dissonant sound that brings about resolution would not itself be perceived as rough but would, instead, effectively perform the role traditionally attributed to a consonance.

As per the earlier example of Wagner's music-drama, the Tristan chord is not resolved by a consonance but, rather, by a less dissonant dominant-seventh chord.[59] Yet, as Rehding points out, "even though partial resolution is achieved, residual tension remains, which is not fully resolved until much later".[60] Nonetheless, whether or not dissonances can only be said to bring about partial resolution in a given context, it is the dissociation between the experience of resolution and the sensation of roughness that is relevant here, in that it represents another contrast between the phenomenology of sensory dissonance and its musical counterpart. Specifically, while the sensory component is based on the association of the sonorous quality of a given dyad or chord as rough with the auditory sensation of roughness and the aesthetic judgement of dissonance, as I have previously argued, the same correspondence is not experienced as such in the case of 'harmony'.

As a consequence of this disparity, one last important contrast between the phenomenology of sensory consonance/dissonance and that of 'harmony' may be observed, which concerns the relationship between aesthetic responses and aesthetic judgements, in particular. As outlined in Section 2.4, I proposed that consonance and dissonance judgements in the sensory case should not be taken as independent from the aesthetic responses of pleasantness and unpleasantness. However, while the correspondence identified above dovetails the association of the aesthetic response of pleasantness to consonance judgements and that of unpleasantness to dissonance for the sensory component, the same once again cannot be said in the case of 'harmony'.

---

[59] For further detail on the Tristan chord, please refer to footnote no. 6 on page 79 as well as in-text references on page 96.

[60] Rehding, A., 'Acoustics', In Kelly, M.(ed.), *Encyclopedia of Aesthetics*, OUP, 2014, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199747108.001.0001/acref-9780199747108-e-6.

As stated earlier, dissonant sounds can be said to evoke aesthetic responses of pleasantness in the case of 'harmony' precisely for the tension they create, or in virtue of the anticipation of resolution – and not due to an auditory sensation of smoothness. This may be observed in some of the common descriptions offered of tension-generating sounds like 'dramatic' or 'intense', for example, which may stem from an experience of aesthetic pleasure. Conversely, the sole use of consonances may sometimes be described as uneventful or even boring – as smooth as their sounds may be said to be on a sensory level. Hence, what this seems to reveal is that, unlike in the psychoacoustic case, the distinctively musical experience of dyads and chords does not require that consonance/dissonance judgements be correlated with pleasantness/unpleasantness; indeed, it frequently subverts that correspondence.

In view of the above, I take these considerations to mount a substantial case for my thesis of the dual phenomenology of consonance/dissonance that warrants the distinction between the sensory component and 'harmony'. However, as indicated earlier, the dual phenomenology of consonance/dissonance does not entail a complete independence between the psychoacoustic and the musical domains. Rather, I have sought to defend the view that the rivalry between the psychoacoustic approach and the contextualist view is only apparent, when we consider the ways in which they converge in traditional tonal harmony. In particular, its foundational principles testify to the much-derided relationship between psychoacoustics and music. Hence, instead of creating an artificial boundary between the two components, the notion of 'musical consonance' that I have adopted here is a unifying conception that enables a dialogue between the two, whilst recognising that each leads to a different form of enquiry into the aesthetics of tonal harmony. Most importantly, the dual phenomenology thesis enables us to account for the metaphysics of tonal harmony in a way that respects the distinctiveness of each of the two components of 'musical consonance' – which is what I aim to achieve in the next chapter.

Chapter Nine: The Metaphysics of Musical Consonance

Having considered the aesthetics of the sensory component and that of 'harmony' in Chapter Eight, the task of the present chapter is to account for the metaphysics of 'musical consonance' as a whole in view of the dual phenomenology of consonance and dissonance. My analysis in this chapter is thus based on the contrast between the psychoacoustic and the musical components in terms of the phenomenology of consonance/dissonance perception as pertaining to chords *qua* isolated auditory objects and *qua* musical individuals, respectively. This will be conducted in tandem with my outline of the psychoacoustic base of chords as detailed in Chapter Seven. To that end, I am first going to argue that consonance and dissonance are irreducible to the psychoacoustic base of chords. But the ways in which they are ontologically dependent upon it differ. I shall propose that the aesthetics of sensory consonance/dissonance is best construed in terms of response-dependence, whereas that of 'harmony' cannot be so. Instead, it is under the notion of aesthetic supervenience that – I will suggest – 'harmony' is best understood.

1. The irreducibility of consonance/dissonance

The first aim of the present chapter is to provide reasons against the view that consonance and dissonance are reducible to the psychoacoustic base of dyads/chords. Although the objections that I shall raise here should be applicable to any such attempt at reduction, my argument will be conducted as a response to Terhardt's thesis that both components of 'musical consonance' can be reduced to psychoacoustic facts. Specifically, I will argue that neither the sensory component nor 'harmony' is reducible to the psychoacoustic base of dyads/chords. My argument will be focused on Terhardt's claim that the sensory component is reducible to sound fluctuations and sharpness of timbre, whereas 'harmony' is reducible to virtual pitch. After rejecting a few possible interpretations of Terhardt's view, I will argue that the metaphysics of 'musical consonance' must be accounted for in view of the dual phenomenology of consonance and dissonance. Hence, although I will continue to adopt his conceptual framework – as indicated in Chapter Six – my account in this section is a significant departure from his proposal.

Two preliminary observations concerning Terhardt's approach are worth noting. First, although his thesis applies to both kinds of consonance that he identifies (i.e. the sensory and the musical) there is a certain degree of nuance in his view, which seems to stem from his acknowledgement that Helmholtz's theory was misunderstood.[1] Arguably in order to prevent his account from incurring a similar fate – especially given his declared support for Helmholtz's contribution – Terhardt specifies that "one cannot expect that musical consonance as an entity could be reduced to one basic psychoacoustic principle".[2] Instead, he maintains that individual explanations for the perception of the sensory component and that of 'harmony' must be discerned: in the sensory case, he attributes it to the auditory sensations of roughness as well as sharpness of timbre, whereas in the case of 'harmony' he ascribes it to the perception of virtual pitch.[3] But these explanations share a key characteristic: they both point to *psychoacoustic* properties of chords *qua* auditory objects.

Yet, one might question whether Terhardt's view is one of metaphysical reducibility. Under a standard reading of reductionism, the nature of what is reducible (in this case, consonance and dissonance) is given in terms of the reductive base (i.e. certain psychoacoustic properties). Whether or not Terhardt intended his reference to reducibility to be interpreted in metaphysical terms, he *does* claim that "each of the two components of musical consonance ... has been reduced to solid psychoacoustic facts", and he treats the principles of tonal harmony as mere "products of certain features of auditory processing".[4] Hence, I take the corollary of his view to be such that harmony can indeed be accounted for solely in terms of psychoacoustic properties – a view which I will seek to reject here.

Since Terhardt claims that both the sensory component and 'harmony' are reducible to psychoacoustic facts, my argument will be focused on individual issues that arise for each of them. First, in the case of the sensory component, Terhardt reduces it to two psychoacoustic facts, in particular. Despite acknowledging that the frequency components of a dyad or a chord's constituent tones and their respective power spectra are both relevant for consonance perception, Terhardt argues that it is the perception of (i) sound fluctuations – i.e. the auditory sensation of roughness – and (ii) sharpness of timbre that are the psychoacoustic facts that account for the perception of the sensory component of 'musical consonance'.[5]

---

[1] As seen in Chapter Seven (Section 2), Helmholtz's theory of musical acoustics came to be seen as a full-throated attempt to provide a naturalistic explanation for the perception of harmony as a whole.
[2] Terhardt, 1984, p. 283.
[3] The phenomenon of virtual pitch was described in Chapter Four (Section 1).
[4] Ibid., p. 292.
[5] Ibid., p. 292.

Terhardt's account of the sensory component as reduced to (i) is given in terms of an "inverse relationship between roughness and evaluated consonance"[6] – which is effectively an iteration of the interference hypothesis. Indeed, as seen in Chapter Eight (Section 2.1), there seems to be ample psychoacoustic evidence that destructive interference is one of the causal mechanisms underlying dissonance perception. However, Terhardt goes further to draw the conclusion that the absence of sensory consonance is reducible to the presence of roughness – a conclusion which I take to be misguided for three main reasons.

First, the move from dependence to reducibility is unmotivated. Terhardt initially describes sensory consonance as "actually dependent on roughness" but subsequently argues that sensory consonance is "basically dependent on other sensory parameters as well".[7] It is therefore unclear how empirical findings that point to a range of different parameters would warrant the leap from ontological dependence to metaphysical reducibility, in any case. In order to do so, Terhardt would be expected either to identify a psychoacoustic property to which sensory consonance is reducible or explain how it may be reduced to a combination of sensory parameters – but neither explanation is given.

That said, it could be argued that Terhardt's reducibility claim may be interpreted differently. Granted, given that he is not a metaphysician, it seems appropriate to consider how his reductionist approach may be otherwise defended. Even if reduction is understood in terms of identity of properties, the objection above does not in itself undermine such an approach. Instead, the main difficulty concerns the fact that reducing consonance and dissonance to the psychoacoustic base would entail that consonance and dissonance *qua* aesthetic properties are *identical* with psychoacoustic properties – which is precisely the view I am arguing against in this section. Yet, one might suggest that Terhardt's reducibility thesis only amounts to a replaceability of terms, instead. On this view, his is a conceptual form of reducibility, rather than the metaphysical kind. But this would also be problematic, and is best explained under the second reason why I reject his view, as follows.

Although Terhardt understands sensory consonance in terms of roughness, it is unclear how the reducibility claim is entailed by that inverse relationship. On the one hand, the psychoacoustic property of consonance would not be reducible to that of roughness itself, in that it is the absence of roughness – under the interference hypothesis – that indicates the presence of consonance. On the other hand, even if one takes sensory consonance to be dependent upon the absence of roughness, to treat the latter as a property is a questionable move, in that it is the absence of something else (i.e. beats).

[6] Ibid., p. 284.
[7] Ibid., p. 285.

Hence, what Terhardt is effectively suggesting is that consonance – under that inverse relationship – is reducible to a double absence, i.e. the absence of the sensation of roughness which is due to the absence of beating between frequency components. But this is where the alternative interpretation of Terhardt's reducibility thesis as solely conceptual seems to be not only inadequate but also metaphysically obscure. It is inadequate because it does not account for the perception of consonance/dissonance *qua* aesthetic properties, in that they are taken to be reducible to the absence of psychoacoustic properties – so they are still accounted for in those terms. In addition, I take it to be metaphysically obscure for not explaining how the absence of a certain property – whether it be the absence of beating or of the auditory sensation of roughness – enables an aesthetic property to be instantiated.

However, although Terhardt writes of sensory consonance in particular, another way of considering the reducibility claim is in terms of sensory dissonance, instead. As outlined in Chapter Eight (Section 2), it is dissonance that is associated with the sensation of roughness, so it could be argued that it is dissonance that is reducible to roughness – which would equally support the view stated in (i) above. This modification would seem to be a more plausible proposal, since it also conforms to evidence for the interference hypothesis. But there is a further difficulty with this counterargument, which concerns the relationship between the acoustic base and the audible properties of chords. This constitutes the third reason why I reject the reductionist approach, and is best understood by considering (ii).

As seen in Chapter Four, in spite of the perceived correlation between the acoustic properties of frequency and amplitude/intensity with the audible properties of pitch and loudness, respectively, the latter is not normally taken to be reducible to the former. This is even more salient in the case of timbre, which is not single-handedly associated with any acoustic property in particular. Specifically, if dissonance is reduced not only to sound fluctuations but also sharpness of timbre – as per Terhardt's thesis expressed in (i) and (ii) above – then a further difficulty arises. As indicated in Chapter Two, there are several factors underlying timbre perception, including the duration, intensity and frequency of tones as well as the auditory experience of the subjects. This is what has led Stephen Handel to describe timbral experience as context-dependent and conclude that "no known acoustic invariants" can be said to underlie it.[8] On that basis, if one were to maintain that sensory dissonance is reducible to sharpness of timbre, then one would effectively be attempting to reduce it to some property to which timbre itself is *not* reducible. Hence, this would amount to yet another counterintuitive consequence of the reductionist approach.

[8] Handel, p. 433.

Hence, whether we speak of sensory consonance or dissonance, the reducibility thesis is at best unmotivated and, at worst, an inadequate attempt at giving a psychoacoustic account of dissonance perception. As noted earlier, although Helmholtz had endorsed the correlation between the perception of dissonance and the auditory sensation of roughness in virtue of the phenomenon of beating, his view is not – or at least not overtly – committed to ontological reducibility.[9] In following Helmholtz's footsteps, it could still be suggested that Terhardt is not, either; but his attempt to reduce 'harmony' to the perception of virtual pitch is, I believe, an unmistakable indicator that he is.

As explained in Chapter Four, the phenomenon of virtual pitch results from the workings of the auditory system in tracking the fundamental frequency that is the 'lowest common denominator' of a series of frequency components – i.e. the overtones of the tones that it detects. As such, it could be classed as one possible explanation under the harmonicity hypothesis, in that the virtual pitch of a chord *qua* Gestalt-based auditory object is precisely the 'missing fundamental' – i.e. that lowest common denominator. Terhardt therefore claims that virtual pitch, which was not known to Helmholtz, provides a psychoacoustic explanation for 'harmony': in his words, "the fact that a pitch can be perceived although there does not exist any spectral component of the frequency corresponding to that pitch provides the key to understanding the whole harmony phenomenon, that is, tonal affinity, compatibility, and fundamental-note relation".[10]

By indicating these three principles identified above as the basis for the whole phenomenon of harmony, Terhardt's understanding of it is considerably more restrictive than the three approaches identified by Rehding, as outlined in Chapter Eight (Section 3.1). This is in itself a weakness, in that his conception of 'harmony' dismisses significant developments in the aesthetics of consonance/dissonance perception even though it is aimed at accounting for the musical component of 'musical consonance'. Most importantly, however, his attempt to reduce those principles – which may be taken as alluding to the foundations of traditional tonal harmony – to a *single* psychoacoustic property is particularly problematic. Similarly to the case of sensory consonance, his account is somewhat vague but seems to revolve around what he refers to as the principle of 'tonal meaning'. Under this principle, in processing the pitches of the complex tones that jointly constitute a chord, the auditory system 'assigns' a tonal meaning to the virtual pitch that it identifies.

---

[9] Helmholtz's claim that the modern tonal system "was not developed from a natural necessity, but from a freely chosen principle of style", for instance, could be taken to indicate that he was not committed to the view that 'harmony' is reducible to psychoacoustic facts. Helmholtz, p. 249.

[10] Terhardt, 1984, pp. 287-288.

For instance, in the case of a major triad such as C4–E4–G4, its virtual pitch is equivalent to "two octaves below the lowest one of the primary fundamentals", namely, C2.[11] This being the case, upon being presented with a token of that chord subtype, the auditory system would attribute the tonal meaning of C4–E4–G4 to the virtual pitch 'C' – which would enable us to identify the chord as a triad in 'C' (Major).[12] In Terhardt's view, this relationship between virtual pitch and tonal meaning similarly applies to other chords and, most importantly, is taken to provide a psychoacoustic basis for 'harmony'.

However, there are two key difficulties with Terhardt's analysis. First, it is at best oversimplifying, in that reducing a given chord to the tonal meaning of its fundamental does not seem to account for differences between tokens of chord inversions of their respective subtypes (e.g. E4–G4–C4 and G4–C4–E4), and neither does it account for differences between the two modes (e.g. C Major and C minor, which would share the same virtual pitch in C2). The second difficulty concerns not so much what Terhardt says about it, but what he does *not* explicitly say when he states that consonance is reducible to virtual pitch. While he acknowledges that virtual pitch was originally developed to further the understanding of pitch phenomena in particular, he does not make the case as to how it can be said to explain why consonance *depends* on it, let alone why 'harmony' is *reducible* to it. In addition, even if his example of major triads were developed sufficiently enough to substantiate his claim to reduction, his attempt at extrapolating it based on a single psychoacoustic fact falls short of the complexity of the 'whole harmony phenomenon'.

Hence, Terhardt's attempt to reduce 'harmony' to the perception of virtual pitch is fraught with difficulties. Over and above the issues identified above, it is unclear what it would mean for consonance *qua* distinctively musical property to be reduced to virtual pitch, which is a psychoacoustic property the ontology of which is somewhat peculiar – given that, as a 'missing' fundamental, it has no actual counterpart in the acoustic base. Yet, Terhardt's account is silent on this issue. Furthermore, with his restrictive view of 'harmony' – as indicated earlier – Terhardt's approach becomes an easy target for the objections raised by those who favour a contextualist approach to consonance/dissonance. Worse still, Terhardt's two-component concept of 'musical consonance' is undermined as a result. Rather than a conciliatory approach, his reducibility thesis only feeds the conflict between music and psychoacoustics – which, perhaps ironically, is what he sought to put an end to with his account.

---

[11] Terhardt, 1974, p. 1067.
[12] Terhardt's own example was that of a G Major triad (specifically: G5–B5–D6, with G3 as its virtual pitch). Terhardt, 1984, p. 291.

In view of these difficulties, the claim that 'musical consonance' can be reduced to psychoacoustic properties does not seem to hold water. In the case of 'harmony', if one conceives of it in Terhardt's terms, one is faced with the challenge of explaining how the ontological reduction to virtual pitch is warranted. On the other hand, if one takes 'harmony' to encompass a broader understanding of it that includes the voice-leading or the contextualist approaches, then the reducibility thesis is quite simply untenable. This is because it does not account for some distinctively musical phenomena such as resolution and the emancipated dissonance, as described in Chapter Eight (Section 3.4). Conversely, in the case of the sensory component, one may object that an alternative approach that focuses on sensory dissonance, instead, may prove more fruitful. Yet, it is questionable to what extent reducing dissonance to the auditory sensation of roughness can illuminate the nature of consonance itself. Furthermore, as seen earlier, reducing dissonance to sharpness of timbre equally leads to counterintuitive results that undermine the reductionist view.

Consequently, if neither consonance nor dissonance is reducible to any psychoacoustic property, then there must be another way to account for the metaphysics of 'musical consonance'. Having rejected the reducibility thesis, in the subsequent sections I shall attempt a solution to the problem of consonance that accounts for the differences between the sensory component and 'harmony', in line with my thesis of the dual phenomenology of consonance/dissonance as outlined in Chapter Eight (Section 3.4).

2. The metaphysics of sensory consonance/dissonance

Dovetailing with the account I offered in Chapter Eight (Section 2), it is with the consonance and dissonance of dyads/chords *qua* isolated auditory objects that this section is concerned. Having previously argued that sensory consonance/dissonance must be understood *both* in terms of judgements of auditory sensations *and* responses to those phenomena, my argument in this section will incorporate these two elements into my assessment of the relationship between consonance/dissonance and the psychoacoustic base of dyads/chords. Specifically, I will argue that they are best understood via certain response-dependence biconditionals for sensory consonance and dissonance, which I will subsequently illustrate with the case of triads. This will form the basis for my analysis of the apriority of the concepts of 'concord' and 'discord' *qua* chord types, as an attempt to provide a characterisation of those concepts that is grounded in the understanding of the aesthetics of consonance and dissonance *qua* sensory properties.

Given that the sensory component is closely related to its psychoacoustic base, my account in this section is particularly concerned with the relationship between aesthetic responses to consonance and dissonance and the non-aesthetic base of dyads and chords. In Chapters Seven and Eight (Section 2.1), I indicated that two psychoacoustic mechanisms in particular – namely, harmonicity and interference – are the most likely physical candidates to be causal mechanisms in the perception of sensory consonance and dissonance. Thus, in spite of Rehding's dismissal of the harmonicity hypothesis and Scruton's resistance to the interference hypothesis that underlies Helmholtz's theory, the latest psychoacoustic evidence – as collated by Harrison and Pearce – seems to support *both* of them as causal factors underlying the aesthetics of sensory consonance and dissonance.

Beyond its empirical basis, this is a metaphysical claim that has two important consequences for my account of the metaphysics of sensory consonance/dissonance. First, given the psychoacoustic evidence for the roles of harmonicity and interference in consonance/dissonance perception, I will take these to constitute the basis for a set of non-aesthetic conditions that would enable appropriate assessments of the consonance or dissonance of specific chord tokens. Secondly, a further consequence of that claim is that verdicts[13] of consonance/dissonance may be predicted. Although this may *prima facie* seem restricted to psychoacoustics and thus removed from the aesthetics of tonal harmony as such, I will argue that this predictive element is what has enabled compositional practice to be developed in terms of the concepts of 'concord' and 'discord', in particular.

It is in light of these observations that I intend to account for the metaphysics of sensory consonance/dissonance. Unlike what Rehding professes in dismissing the harmonicity hypothesis as useful fiction, it seems that describing any relevant findings from psychoacoustic research as such is tantamount to neglecting the role of sensory consonance/dissonance in shaping the foundations of traditional tonal harmony. Those who favour a contextualist view may similarly choose to relegate it to the realm of fiction; yet, the context-sensitivity of the musical experience of harmony does not entail that consonance and dissonance must be *essentially* context-sensitive. Hence, it would seem an unjustifiable and neglectful choice to dismiss the perception of dyads/chords *qua* isolated auditory objects, since they form the basis of those foundational principles. In other words, prior to considering consonance and dissonance within a distinctive musical context, it seems not only appropriate but also fundamental that an account of the aesthetic experience of isolated chord tokens be provided in relation to their psychoacoustic base.

---

[13] As indicated in Chapter Eight (Section 2.4), the term 'verdict' – in the sensory case – encompasses *both* judgements *and* responses. Yet, I will revisit the usage of the term under the metaphysics of 'harmony'.

2.1. Response-dependence

Having rejected the possibility of reduction in Section 1, in this section I will provide an account of the metaphysics of sensory consonance/dissonance in terms of a particular kind of ontological dependence, namely, that of response-dependence. In order to substantiate my solution to the problem of consonance, the argument that I shall present here is twofold. First, I will propose an account of sensory consonance/dissonance *qua* properties, since I take the problem of consonance to be primarily of a metaphysical nature. I will subsequently point to some epistemological consequences of it by considering the nature of sensory consonance/dissonance *qua* concepts, which should thus address the conceptual dimension of the sensory component of 'musical consonance'. It is in view of this twofold proposal that I will argue that the notion of response-dependence is best equipped to account for the nature of concords and discords, which are normally taken to refer to certain consonant and dissonant dyad/triad types, respectively.

Due to the narrow scope of this chapter, however, three preliminary remarks must be made concerning the notion of response-dependence as applied to isolated chord tokens. First, my account of sensory consonance/dissonance is neutral on the broader metaphysical question around the nature of properties as such. Whether they are taken to be dispositions or just generally conceived as 'ways that things are', I want my proposal to work whatever the best account of the ontology of properties turns out to be. Second, I shall restrict my proposal to considering the conditions under which aesthetic responses to single dyad/chord tokens and judgements thereof may be understood in terms of a given biconditional. Third, it is not my intention to enter into the broader debate around the metaphysics of response-dependence as such. Instead, my account will be restricted to its most conventional aspects, such that it may work under whichever version of response-dependence one chooses to favour.

In view of the above, I am going to argue for a response-dependent view of sensory consonance and dissonance *vis-à-vis* the psychoacoustic base of dyads and chords. To that end, the specific conditions under which their basic formulations will be given are fundamentally non-aesthetic, in that they concern the harmonicity and interference hypotheses as outlined in Chapter Seven. Although the distinction between aesthetic and non-aesthetic properties will be considered at greater length in the context of the metaphysics of 'harmony' (Section 3), for the present purposes I take 'non-aesthetic' to refer primarily to the psychoacoustic properties underlying chord perception.

My account of the sensory component of 'musical consonance' is thus circumscribed to the psychoacoustic domain, thus conforming to the distinction I adopted in Chapter Six. As seen in Chapter Eight (Section 2.1), psychoacoustic experiments are taken to provide empirical evidence for consonance/dissonance perception. Upon being presented with the sound of a dyad or a chord, participants are asked to make a judgement. In spite of certain nuances of specific studies, their ratings are normally computed in terms of consonance or dissonance judgements which can often be predicted by means of computational models. Beyond the divergence in some of the studies alluded to in that chapter, what their findings indicate is that it is possible to express those predictions in terms of certain biconditionals that relate the consonance or dissonance of a given dyad/chord *qua* auditory object to the nature of the participants' judgements. Yet, as I have argued in Chapter Eight (Section 2.4), they must also include reference to their *responses* in terms of pleasantness or unpleasantness as part of their predictions. It is under this view that the biconditionals will be formulated here. It is worth noting that, although they may also apply to the aesthetic responses that listeners in general may have to tokens of isolated dyads/chords, the focus of this main section is on the parameters of psychoacoustic experiments, in particular.

2.1.1. Terms of the biconditionals

Biconditionals are at the centre of any response-dependence account, in that they set both the individuation conditions for a given property and the truth conditions of claims around its ascription to a certain object. The first part of my argument in this section will thus be concerned with consonance/dissonance *qua* properties, so that their conceptual dimension may be subsequently considered in light of it. In order to keep to my earlier commitment to a common-ground understanding of response-dependence, the version of the biconditional I will adopt here is that suggested by Jussi Haukioja in his overview of different notions of response-dependence, which is stated thus: (1) "*x* is *F* iff *x* would elicit response *R* from subjects *S* in conditions *C*".[14] Accordingly, I shall adapt this formulation of the response-dependence biconditional for both parts of my main argument, such that *F* may stand both for consonance/dissonance *qua* property and *qua* concept.

[14] Haukioja, J., 'Notions of Response-dependence', In Hoeltje, M., Schnieder, B., Steinberg, A. (eds.), *Varieties of Dependence*, Munich: Philosophia, 2013, p. 168. Different versions of response-dependence offer different ways of stating the basic formula, which may refer either to properties or concepts.

In the case of consonance/dissonance *qua* property, *x* refers to any given chord token, but it may also refer to that of a dyad (i.e. a harmonic interval). For instance, *x* may refer to the interval of a third or a fifth, or a major or a minor triad. In turn, *R* refers to the aesthetic response that is evoked once the dyad or chord is sounded. In tandem with my account in Chapter Eight (Section 2.4), I will take *R* to be characterised in terms of pleasantness or unpleasantness. Next, the subject *S* refers not only to a given participant in a psychoacoustic experiment but also, in principle, any given listener. However, given that my account is restricted to the case of tonal harmony in particular, I will take *S* to refer to Western listeners, as described in the previous chapter under the terms of Harrison and Pearce's study. In addition, given the historical variances in tuning and temperament – which will be considered in Chapter Ten – I will take *S* to refer to subjects that have only ever experienced harmony under the system of twelve-tone equal temperament.

Arguably, the most complex term in a response-dependence biconditional is the set of conditions *C*; and given what has been outlined in previous chapters, this is most certainly the case with the sensory component of 'musical consonance'. As stated earlier, I take the conditions under which an isolated dyad/chord may be said to be consonant or dissonant to be determined by their psychoacoustic base. In particular, the growing consensus in psychoacoustic circles – as seen in Chapter Eight (Section 2.1) – is that both interference and harmonicity are causal factors in consonance/dissonance perception. This, however, would seem to create a problem, given that these acoustic properties may be said to be intrinsic to the nature of dyads and chords. As such, one might argue that they are not the kind of conditions normally dealt with in the metaphysics of response-dependence.

In order to explain why I take interference and harmonicity to be the basis for a set of conditions for the sensory-component biconditionals, it is essential to recapitulate some aspects of my account of the ontology of chords from Part I. Specifically, I then argued that chords are best understood both as events and as auditory objects – the analyses of which were undertaken in Chapters Three and Four, respectively. This distinction is not only relevant but essential for the response-dependence biconditional to work. This is because both harmonicity and interference are acoustic properties which pertain to chords *qua* sound events, whereas the isolated dyads and chords that are being considered under the biconditional must be understood as Gestalt-based auditory objects, instead. Under my earlier characterisation, chords *qua* auditory objects have distinctive properties which are not shared by chords *qua* sound events, such as Gestalt-based qualities and psychoacoustic properties associated with both linear and non-linear mechanisms of auditory perception.

Hence, I will consider the aesthetic experience of dyads and chords *qua* isolated auditory objects under those non-aesthetic conditions. Specifically, interference conditions will be identified with reference to the phenomenon of beating and the notion of critical band, while harmonicity conditions will essentially concern the presence of whole-number relations between frequency components of a given overtone series and their fundamental.

Prior to formulating the response-dependence biconditionals, two important observations on the nature of those non-aesthetic conditions seem pertinent. First, although both harmonicity and interference may be said to provide causal explanations for the perception of sensory consonance/dissonance, I do not take them to be the sole basis for accounting for the aesthetics of the sensory component as such. To do so would mean that they are not merely conditions but, essentially, that they may be taken to account not only for judgements of dyads/chords but also aesthetic responses to them in terms of pleasantness and unpleasantness, respectively. This, however, is *not* a view that I share. As I have previously argued, consonance and dissonance are not reducible to any acoustic or psychoacoustic properties. Furthermore, I have also distinguished between the content of aesthetic judgements and the nature of aesthetic responses to consonance/dissonance. With this distinction, I aimed to differentiate between aesthetic judgements of the sonorous quality of dyads/chords and aesthetic responses to them whilst also holding psychoacoustic theories to account, in view of my argument against the reductionist approach.

Secondly, given that the non-aesthetic conditions include not only acoustic properties of dyad/chord events but also their psychoacoustic properties *qua* auditory objects, two further questions seem to arise. These concern their relevance for the phenomenology of consonance/dissonance perception and the epistemic access that listeners may have to such conditions. On the one hand, one may object that those conditions, whether acoustic or psychoacoustic in nature, may not be relevant to the way listeners *experience* the sounds they hear – seeing as no amount of psychoacoustic information is sufficient to characterise the phenomenal qualities that dyads and chords may be perceived as having. On the other hand, one may equally question the relevance of such non-aesthetic conditions for consonance/dissonance perception given the limited, if not negligible, epistemic access that listeners may have to the specificities of those conditions under the relevant parameters of psychoacoustic experiments. This objection may even be made more poignant when one considers the kind of conditions that response-dependence accounts may usually include – such as 'under normal daylight' in the case of colour perception, for instance – which are both accessible and phenomenologically relevant.

Yet, the first objection would only apply if one were attempting to provide a *psychoacoustic* analysis of the experience of harmony – which is not the case at present. Rather, in pursuing an account of the metaphysics of its sensory component, it is the nature of the auditory sensations that listeners have that is of central importance, since what is being attended to in experiments is the sonorous quality of a given dyad/chord considered as an isolated auditory object. But since this results from certain non-aesthetic conditions, I take these to be not only relevant but fundamental to assessing whether those sensations are experienced as smooth or rough, for example. Similarly, the responses of pleasantness/unpleasantness that are evoked are also due to the nature of the experience of those auditory sensations – but it is these that are determined by the psychoacoustic base.

Although the issue of epistemic access, in turn, would seem to pose a more serious difficulty, I want to argue that such is not the case. This is because it is those who undertake experiments that are required to have epistemic access to the psychoacoustic details of the conditions, rather than the participants themselves. Indeed, the access that participants have to those conditions is via their perceptual awareness of them, which is no different from the case of normal daylight conditions in the example of colour perception. Over and above that, whether or not listeners have epistemic access to the specificities of the non-aesthetic base does not detract from the experience that they have, and neither does it raise any issues for the researchers' ability to assess the relationship between the participants' auditory sensations and the psychoacoustic base underlying them.

Having considered these preliminary observations, in what follows I intend to provide an account of sensory consonance/dissonance in terms of certain response-dependence biconditionals. As indicated earlier, I take verdicts of consonance or dissonance to encompass both aesthetic judgements of auditory sensations and the corresponding aesthetic responses. Hence, in order to incorporate this distinction between aesthetic judgements of the smoothness or roughness of tokens of dyads/chords and aesthetic responses to them in terms of pleasantness and unpleasantness, I propose a new iteration of the earlier formulation of the biconditional as follows: (2) x is F iff x would elicit an auditory sensation A and evoke an aesthetic response R from subjects S in conditions C. It is under this refined proposal that I shall consider some specific cases of sensory consonance/dissonance perception in the next subsections. Since my account is applicable both to dyads and chords, for ease of exposition I shall use examples of dyads first, which should convey a clearer analysis of cases of chords. For the sake of consistency, I shall favour examples that have been previously used in Chapters Seven and Eight.

2.1.2. *The sensory-consonance biconditional*

In this section, I aim to provide an account of sensory consonance *qua* response-dependent property. If we apply (2) above to the perception of isolated dyads/chords, the formulation of the biconditional for sensory consonance could be stated as follows: (3) a given dyad/chord is consonant if, and only if, it elicits an auditory sensation of smoothness (or equivalent) and evokes a favourable aesthetic response of pleasantness (or equivalent) from Western-type listeners under the following conditions: (i) the frequency components of each constituent complex tone are integer multiples of their fundamental (or at least perceived as such);[15] (ii) the fundamental frequencies of each constituent complex tone fall outside the critical band region;[16] (iii) it is *either* the case that the frequency components of each constituent complex tone do not beat against one another's *or*, in case they do, they are not perceived as such.[17] While I take these conditions to be necessary for an assessment of sensory consonance, they are insufficient as they must be considered in tandem with the aesthetic judgement of *A* and the aesthetic response *R* – which must be satisfied as such.

To illustrate it with tokens of dyad types, Helmholtz's perfect consonances such as the fifth and the fourth[18] would be accounted for by (3), in that both *A* and *R* are normally favourable and all three conditions are, in principle, met. The same, however, cannot be said of thirds and sixths – which Helmholtz deemed less perfect ones. In the case of thirds, as detailed in Chapter Seven, some variations in the width of the critical band mean that (ii) would not be met by third intervals on the lower end of the register, such that the interval between C3 and E3, for example, would not be deemed consonant on this account.[19] Similarly, low-register sixths (such as C3–A3) would not be perceived as consonant, in that they do not meet (iii), but they otherwise would when sounded on the upper register, seeing as the second component of the disjunction tends to be the case.

---

15 As seen in Chapter Seven (Section 3), although the auditory system is able to distinguish between resolved and unresolved harmonics, there are occasions in which a lack of harmonicity may not be perceived as such.

16 Under the critical band theory developed by Plomp and Levelt, harmonic intervals comprising frequencies that fall *outside* the bandwidth are deemed consonant. Conversely, those that fall within would be classed as dissonant; in particular, if the frequency difference between them is less than 25 percent of the critical band, the interval is typically judged as dissonant. Plomp, Levelt, p. 560. Although there is no consensus around the threshold figure, for the purposes of this section I shall adopt Plomp and Levelt's as the standard view.

17 This could be due to the fact that lower frequencies have masked the beating partials, or the acoustic power of the latter is insufficient for them to be perceived as beating – or even to be perceived at all.

18 The harmonic status of the fourth is historically controversial. In general, the fourth has been treated as a consonance, but it has also been classed alongside dissonant intervals. Although these historical divergences lie outside the scope of the present section, some aspects of this controversy will be discussed in the next chapter instead, as part of my assessment of the impact of certain differences in tuning and temperament.

19 Under Plomp and Levelt's theory, the major third between C3 and E3 (just as C1–E1 and C2–E2) has its component tones falling within the critical band, so strictly speaking it *cannot* be classed as a consonance. By contrast, C4–E4 may be deemed consonant, in that both components fall outside the critical band.

Counterintuitive though it may appear to some to consider thirds as 'not consonant', this consequence of subjecting it to (3) indicates two advantages of the biconditional. First, it conforms to the psychoacoustic evidence that low-register thirds are not perceived as consonant, in contrast to their counterparts on upper sections of the register. Second, and most importantly, although the biconditional may be applied to any given intervallic relationship, it is the consonance of tokens of specific *subtypes* of intervals that is at stake when assessing cases of dyads. Hence, two additional observations must be considered at this stage, which concern the type/token distinction I adopted in Chapter Five.

First, I have previously argued that the type/token distinction is not only suited to account for chords *qua* musical individuals but is equally applicable to the harmonic intervals which they comprise. Yet, further to the first-order distinction between dyad/chord types – i.e. the patterns – and their respective tokens – i.e. their audible instances – I have also identified another level of specificity for those types, in that they depend on the octave within which (or across which) they may be found. For instance, the interval of a major third (e.g. C–E) is a type of interval; but, for the response-dependence biconditional to do its work, it is necessary that its subtype be specified (e.g. C3–E3, and *not* C4–E4). Hence, in view of (3), any apparently inconsistent judgement concerning a given interval type is explained away once one specifies that *x* is not just the interval type found between C and E but precisely the subtype C3–E3, as per the example above.

Secondly, an important element of the type/token distinction for my account of the response-dependence of sensory consonance *qua* property concerns what should feature on the left-hand side of the biconditional (i.e. the terms prior to *iff*). Specifically, given that responses and judgements involve the occurrence of actual sound events, I want to argue that *x* must refer to *tokens* of dyads/chords, rather than types. While I shall attempt to address it more fully later in this section, this question has implications that are relevant for an assessment of the suitability of (3). This is because what listeners are presented with are tokens of certain dyad or chord subtypes (e.g. the audible instances of C4–E4 or C4–E4–G4, respectively), but since subtypes are abstract entities, it is to tokens thereof that the left-hand side of the biconditional must refer. That said, this may seem counterintuitive to some in that it could be argued that taking *x* to refer to tokens, in particular, is tantamount to dismissing an important aspect of sensory consonance *qua* concept that the response-dependent biconditional should also capture, namely, its apriority. Yet, rather than dismissing it, I will indeed be addressing the conceptual dimension of sensory consonance/dissonance in Section 2.3, under my account of concords and discords.

Given the need to distinguish between the perception of thirds on the lower and upper register, I will incorporate the type/token distinction into the sensory-consonance biconditional – where *F* concerns consonance *qua* property, in particular. Hence, I am proposing another iteration of the response-dependence biconditional, as follows: (4) a given token of a dyad/chord subtype possesses consonance if, and only if, it elicits an auditory sensation of smoothness (or equivalent) and evokes a favourable aesthetic response of pleasantness (or equivalent) from Western-type listeners under the conditions specified in (3). Thus, on this new iteration, the consonance of the third interval C–E may be assessed in any given octave, and any perceptual differences may be accounted for by reference to its relevant token. Consequently, under (4) a token of the C4–E4 interval subtype may be deemed consonant, but the same cannot be said of that between C3–E3.

Beyond the import of the type/token distinction, the outcome of the application of the response-dependence biconditional to cases of low-register thirds points to another important consequence for the account I am pursuing here. Specifically, it concerns the nature of that outcome, and how it may be expressed. In saying that a given dyad – such as C3–E3 – is 'not consonant', two potential interpretations of this predicate may be considered. First, by 'not consonant' it may be meant 'not consonant at all', but it may also mean 'not *as* consonant' – which is what Helmholtz seems to have intended by calling major and minor thirds medial and imperfect consonances, respectively.[20] Yet, 'not consonant' may be construed neither as the lack of a certain quality nor as degrees thereof but, instead, as the presence of another property – namely, that of dissonance.

2.1.3. The sensory-dissonance biconditional

Having set out the sensory-consonance biconditional above, in this section I aim to account for the case of sensory dissonance. As I have indicated in earlier chapters in relation to the dual phenomenology thesis, the distinction between consonance and dissonance is not solely a conceptual one – under which 'dissonance' is taken to be the antonym of 'consonance'[21] – but is fundamentally a metaphysical distinction concerning their nature *qua* properties, and that is for two main reasons. First, I do not take sensory dissonance to be the lack of consonance, just as I do not take sensory consonance to be the absence of dissonance; rather, both of them are sensory properties in their own right.

---

20 This possibility of there being degrees of consonance will be considered in the next subsection, instead.
21 This view was adopted by Harrison and Pearce as part of their account. Harrison, Pearce, p. 217.

In support of that view, the second reason concerns precisely the nature of the psychoacoustic evidence towards each of those phenomena. While the presence of interference is taken to be the main causal factor for dissonance perception, it is to the presence of harmonicity that the growing consensus in psychoacoustic circles points in the case of consonance. In other words, once the criterion for property individuation *vis-à-vis* the psychoacoustic base is established in terms of the presence of something, rather than the absence of something else, it becomes clear that consonance and dissonance cannot be characterised in terms of a reciprocal negative function, i.e. by treating one as the absence of the other. This, however, does not entail that they are reducible to the presence of those psychoacoustic properties, for the same reasons as outlined in Section 1.

Hence, in order to account for that distinction between the two sensory properties, there must also be a version of (3) for sensory dissonance, which I am stating as follows: (5) a given token of a dyad/chord subtype possesses dissonance if, and only if, it elicits an auditory sensation of roughness (or equivalent) and evokes an unfavourable aesthetic response of unpleasantness (or equivalent) from Western-type listeners under the following conditions: *either* (i) the frequency components of its constituent tones beat amongst themselves, and they *are* perceived as an auditory sensation of roughness (or equivalent);[22] *or* (ii) the fundamental frequencies and/or particularly strong partials fall within 25% of the critical band, in particular. Unlike the case of sensory consonance, these conditions for a given dyad/chord to be deemed dissonant need not be jointly met – a contrast which I take to be a further indicator that they should be treated as properties in their own right.

In view of (4) and (5), my response-dependent account of tokens of dyad types is able to provide a metaphysical underpinning for the 'engineering view' of harmony described in Chapter Eight (Section 2). For instance, under (5), dyad types that are traditionally considered as dissonances – such as seconds, sevenths and the tritone – all meet at least one of the two conditions above; and assuming both *A* and *R* are satisfied, they would be deemed dissonant. Yet, the case of tokens of chord subtypes, in particular, is significantly more complex. Given my earlier account of chords as Gestalt-based auditory objects consisting of *three or more* simultaneously occurring complex tones, it is beyond the scope of this section to offer a sufficiently large sample that could illustrate the application of the biconditionals to a variety of possible chord tokens. Instead, for the present purposes, my account will be focused on those that have traditionally been seen as foundational elements of tonal harmony – namely, triads and their respective inversions.

[22] This may be because they hold sufficient acoustic power and/or they are not masked by other frequencies.

2.2. *The case of triads*

In this section, I will consider the consonance or dissonance of triads as a case study. Triads are tertian chords, which means that they are formed by the superposition of two thirds: in the case of a major triad, a major third from the bass note (i.e. the root) is superimposed by a minor third, such that the interval between the bass and the uppermost note is that of a perfect fifth; conversely, by reversing the order of the thirds a minor triad is formed (though the interval of the fifth stays the same)[23]. Since both major and minor triads are tertian chords, the same issue that arises for the perception of thirds on the lower register also affects tokens of triad subtypes. For example, while a token of a C Major triad given in C4–E4–G4 may be deemed consonant, others found on lower parts of the register – i.e. C1–E1–G1, C2–E2–G2 and even C3–E3–G3, as implied from the earlier example – would not come out as consonant under (4), in that they would not meet (ii).

Yet, triads can only be said to be tertian chords when they are in root position, as shown in the examples above. Alternatively, they may be encountered under two variants, as their first and second inversions. Despite being designated under the same triad name (e.g. the first or second inversion of the C Major triad), they have different chord patterns, given that they comprise different intervals.[24] Hence, tokens of first and second inversions of major/minor triads must be subjected to (4) separately; and despite sharing the same chord name as, say, a C Major triad, the qualifiers 'first inversion' or 'second inversion' must be used to designate their individual chord structures. As such, they both constitute chord types in themselves with their own subtypes across the different octaves. This being the case, given that both types of inversion only include harmonic intervals (i.e. dyads) that were, in principle, deemed consonant in Section 2.1.2, their tokens would also be expected to come out as consonant under (4) – provided both *A* and *R* are also satisfied. But the by-now familiar exception occurs when thirds are found on the lower register, in which case (ii) is once again not met. One such example is particularly relevant to illustrate the difference in consonance perception that may arise between the two, as follows.

---

[23] This is the generic chord pattern for any given major or minor triad, respectively, regardless of which pitch class is taken as its root (e.g. whether a C Major triad or a G Major triad).

[24] The system of chord notation developed during the contrapuntal/figured-bass period is particularly helpful for understanding these chord structures. In the case of the first inversion, the intervals that form the chord pattern are a third from the root, superimposed by a fourth, which gives a sixth interval above the bass. In figured-bass, this is represented as '6 over 3', or simply '6'. The second inversion of any given triad is, in turn, notated as '6 over 4', since it comprises a fourth and a sixth interval in relation to the new root (as well as a third 'sandwiched' in between the two). Kostka et al., p. 44.

In comparing the sound of a token of E3–G3–C4 (i.e. a subtype of the first inversion of the C Major triad) and one of the neighbouring G3–C4–E4 (i.e. a subtype of its second inversion), one may observe an audible difference in terms of consonance perception, despite their relative proximity as inversions of the same chord type. This can be explained by the fact that the interval of the third – which is what creates the difficulty on the lower register – is found between E3 and G3 in the case of the first inversion, and thus falls *within* the critical band. However, the same cannot be said of the interval between C4 and E4. This, however, does not mean that 'E3–G3–C4' is a dissonant chord, in that under (5) an aesthetic response of unpleasantness (or equivalent) to it must be evoked, which may not be the case – even if *A* is taken to be satisfied as such and the conditions are equally met.

The example above raises an important question regarding aesthetic verdicts of sensory consonance. As indicated earlier, one possible interpretation of an outcome of the sensory-consonance biconditional is that a dyad/chord token may be said to be 'not as consonant' as another. Hence, one may question whether (4) enables us to discern between different 'degrees of harmoniousness of consonance' – to borrow Helmholtz's term, as described in Chapter Eight (Section 2). For instance, Rigden takes C4–E4 to be "definitely a consonance", whereas C3–E3 is treated as "marginally a consonance or perhaps even a dissonance".[25] Yet, although it may sound counterintuitive not to class E3–G3–C4 as possessing consonance – seeing as it is a token of the first inversion of the C Major triad – the purpose of the biconditional is precisely that of ascertaining whether tokens of dyad/chord subtypes can be said to possess consonance *qua* sensory property – which requires that both *A* and *R* are satisfied but also that the relevant conditions are met.

However, it seems that this kind of ambiguity in some cases of aesthetic verdicts of dyads/chords cannot be so easily dismissed, in that it points to the very nature of the perception of consonance/dissonance *qua* sensory properties. On the one hand, I have argued that sensory consonance and sensory dissonance must be treated as distinct properties when understood in terms of the presence of certain psychoacoustic properties. But, on the other hand, the perceptual assessment of those sensory properties may not always be 'clear-cut'. This being the case, to characterise them in comparative terms (e.g. 'not as consonant/dissonant as') is tantamount to accepting different gradations under which the properties of sensory consonance and dissonance may be perceived. Hence, on that basis, one may be inclined to object to my attempt at accounting for the sensory component of 'musical consonance' by means of the metaphysics of response-dependence.

[25] Rigden, p. 200.

Yet, such an objection may be countered when we consider the nature of psychoacoustic experiments as well as the workings of response-dependence. First, although psychoacoustic findings on consonance/dissonance perception may be given in comparative terms, the biconditionals are meant to be applicable to dyads/chords *qua* isolated auditory objects in their own right. As indicated earlier, I take verdicts on their consonance or dissonance to be given on the basis of *individual* judgements of auditory sensations and aesthetic responses in terms of pleasantness and unpleasantness. Thus, to include verdicts of 'not as consonant/dissonant as' is to assume that they are not being individually assessed, which defeats the very purpose of isolating them in the first place.

Furthermore, in view of the semantic difficulties involving a range of single-word definitions of 'consonance' and 'dissonance' in psychoacoustic studies which equally do not discriminate between judgements and responses – as detailed in Chapter Eight (Section 2.2) – I take the purpose of the biconditionals to be precisely that of providing a standard for consonance and dissonance ascription under the sensory component. Specifically, this would be aimed at preventing cases of ambiguity from being construed as inconclusiveness of findings, instead. This is particularly relevant when one considers the ways in which the psychoacoustics of consonance/dissonance perception has been dismissed by those who hold a contextualist view – such as Scruton – and those who take a subjectivist approach – such as Rehding. Hence, in resisting the possibility of a gradation of consonance or dissonance in response-dependence biconditionals, my account is also aimed at deflecting the challenge posed by contextualist and subjectivist views on the experience of harmony.

Beyond these counterarguments, there is another important observation concerning how consonance/dissonance verdicts on tokens of triads are given. As I argued in Chapter Eight (Section 2.4), in the context of psychoacoustic experiments it is essential that participants are prompted by the distinction between judgements of auditory sensations and their aesthetic responses in terms of pleasantness and unpleasantness. This is because it is effectively both *A* and *R* that are primarily the basis for property ascription, given the participants' limited epistemic access to *C*. In addition, as stipulated in (4) and (5), both *A* and *R* must be satisfied in relation to *tokens* of dyad/chord subtypes for the response dependence to function accordingly. This, however, would seem to pose a difficulty, in that a given triad *type* could not, in principle, be said to be consonant or dissonant as such. Yet, I want to argue that this difficulty may be addressed by considering the nature of sensory consonance/dissonance *qua* concepts. Specifically, I take the concepts of 'concord' and 'discord' – as they have come to be known – to be particularly illuminating here.

2.3. *Concords and discords*

Having set out the biconditionals for sensory consonance and dissonance above, in this section I aim to account for their conceptual dimension. Specifically, I want to argue that the biconditionals in (4) and (5) provide the basis for the concepts of concord and discord. As indicated in Chapter Seven (Section 2), Helmholtz sought to account for the conditions under which chords are consonant, such that they may be termed 'concords'. In his view, the key condition for a chord to be a concord is that each constituent complex tone should "form a consonance" with one another. Conversely, if two constituent tones are "consonant to the root, but dissonant to each other", what results is a discord.[26]

Helmholtz's definitions are not only minimal but also vague. Far from an isolated case, however, a similar degree of imprecision may be observed in more recent definitions of 'concord' and 'discord'. For instance, the *Concise Oxford Dictionary of Music* has it that 'concord' refers to a chord which "seems satisfactory in itself", adding that "what constitutes a concord is not strictly laid down and must often depend on individual assessment".[27] By contrast, 'discord' is taken to refer to "a chord which is restless, jarring to the ear, requiring to be resolved in a particular way if its presence is to be justified by the ear".[28] In view of these definitions, it is perhaps unsurprising that the aesthetics of sensory consonance/dissonance is beset with issues of equivocation in assigning them to chords.

In order to address those issues – as identified in Chapter Eight (Section 2.2) – the account I want to pursue in this section should be able to set out the terms under which a given dyad/chord may be said to be a concord or a discord. In light of the analysis I have offered above, it seems that one of the advantages of the response-dependence biconditionals is that they are able to provide the logical formulations under which the concepts of 'concord' and 'discord' may be more adequately described. Along with the terms of (4) and (5), my proposal is equally based on two fundamental claims. First, contrary to the generality of the definitions cited above, I take 'concord' and 'discord' to refer to dyad/chord *subtypes*, rather than unspecified types. Second, and as a consequence of it, this means that x should refer not to tokens but the subtypes themselves.

---

26 Helmholtz, p. 338. Helmholtz also refers to discords as 'condissonant triads'.

27 Kennedy, M., Kennedy, J. B., 'Concord (consonance)', *The Concise Oxford Dictionary of Music*, 2013, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199203833.001.0001/acref-9780199203833-e-2087.

28 Kennedy, M., Kennedy, J. B., 'Discord', *The Concise Oxford Dictionary of Music*, 2013, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199203833.001.0001/acref-9780199203833-e-2661#.

In view of those claims, I propose that the left-hand side of the biconditional – in the case of sensory consonance/dissonance *qua* concepts – should be adapted from (4) to "*x* is a concord *iff*...", and as "*x* is a discord *iff*..." from its original expression in (5). Yet, this also means that the right-hand side of those biconditionals must equally be revisited. To that end, I propose that the response-dependence biconditionals for sensory consonance *qua* concept be stated as follows: (6) *x* is a concord if, and only if, its designated or salient token elicits an auditory sensation of smoothness (or equivalent) and evokes a favourable aesthetic response of pleasantness (or equivalent) from Western-type listeners under the conditions specified in (4). For instance, C4–E4–G4 is a concord if, and only if, an instantiated token of it is deemed smooth and it evokes a pleasant response from Western listeners under the conditions specified earlier. Similarly, the biconditional for sensory dissonance *qua* concept could be stated thus: (7) *x* is a discord if, and only if, its designated or salient token elicits an auditory sensation of roughness (or equivalent) and evokes an unfavourable aesthetic response of unpleasantness (or equivalent) from Western-type listeners under the conditions specified in (5).

The formulations in (6) and (7) are central to the second part of my argument, in that they point to the conceptual nature of sensory consonance and dissonance as pertaining to dyad/chord subtypes, in particular. Specifically, one of the reasons why I have sought to account for the metaphysics of the sensory component of 'musical consonance' via the notion of response-dependence concerns the *a priori* nature of the relationship between verdicts of the consonance and dissonance of dyads/chords and the concepts of concord and discord, respectively. As Haukioja points out, "to say that a concept or a property *F* is response-dependent is to claim that there is an *a priori* connection between *F* and our responses (with respect to F) in normal, favourable or ideal conditions".[29]

On that basis, if (6) and (7) are accepted as workable response-dependence biconditionals, I take them to be indicative of the apriority of the concepts of concord and discord for two main reasons. First, in view of their individual conditions arising from the non-aesthetic base of dyads/chords, the relationship between aesthetic verdicts and the parameters of the psychoacoustics of consonance/dissonance perception is warranted by the correspondence between the auditory sensations that dyads/chords elicit and respective judgements of their smoothness or roughness. Second, and most importantly, the *a priori* connexion between the non-aesthetic base and aesthetic verdicts may be observed in the fact that the consonance or dissonance of specific dyad/chord subtypes may be predicted.

[29] Haukioja, p. 168.

This claim to apriority is fundamental to understanding the conceptual nature of sensory consonance and dissonance. As seen in Chapter Eight (Section 2.1), computational models deployed in several psychoacoustic experiments are used to predict consonance or dissonance verdicts with reference to chord types, not tokens. This is because, in the particular case of subtypes, it is the relationship between the specific types of intervals that they comprise and the conditions specified under the relevant psychoacoustic properties that seems to provide the criteria for consonance and dissonance predictions. Similarly, this may also be said of traditional methods of compositional practice, given that composers utilise certain chord subtypes – whether in root position or in their inversions – in the firm expectation that their tokens would be instantiating consonances or dissonances when executed. It is, therefore, on that basis that I take (6) and (7) not only to address the earlier concern around the inability of (4) and (5) to assign consonance or dissonance to chord types but also provide workable definitions of 'concord' and 'discord', respectively.

Overall, there are at least three key advantages to understanding the sensory component of 'musical consonance' in terms of response-dependence. First, given that the biconditionals incorporate not only aesthetic responses to chords *qua* auditory objects but also aesthetic judgements of their auditory sensations, my account shows that the aesthetics of sensory consonance/dissonance goes beyond the psychoacoustic base, and is therefore irreducible to it. Second, the biconditionals address the question of the relationship between consonance/dissonance *qua* property and the psychoacoustic base, as well as our ability to predict aesthetic verdicts as such. Third, they also enable an understanding of sensory consonance/dissonance *qua* concepts, thus providing appropriate definitions of 'concord' and 'discord' – which I take to be another virtue of using response-dependence.

3. The metaphysics of 'harmony'

Having characterised the metaphysics of sensory consonance/dissonance above, in this section I aim to account for consonance and dissonance *qua* musical properties, instead. First, I shall consider the nature of aesthetic verdicts of consonance and dissonance in the specific case of 'harmony' in view of the relationship between aesthetic and non-aesthetic properties. Subsequently, whilst recognising that the metaphysics of the sensory component bears implications for that of 'harmony', I will argue that the notion of response-dependence cannot account for the latter. Instead, I will suggest that it is via the notion of aesthetic supervenience that the metaphysics of 'harmony' is best understood.

3.1. Aesthetic verdicts

In order to account for the metaphysics of 'harmony', it seems important to consider the nature of aesthetic verdicts on the consonance or dissonance of dyads/chords in a distinctively musical context *vis-à-vis* their non-aesthetic base. Although the relationship between aesthetic and non-aesthetic properties is well-established in the literature on philosophical aesthetics, the specific cases of consonance and dissonance point to an important aspect of that relationship which must be addressed first for my proposal to work satisfactorily. Specifically, it concerns the consequences of the contrast between the phenomenology of the sensory component and that of 'harmony' for the nature of aesthetic verdicts of consonance and dissonance in those respective domains.

First, since the phenomenology of the sensory component is distinct from that of 'harmony' – as I have characterised them in Chapter Eight (Section 3.4) – the relationship between aesthetic and non-aesthetic properties in those two domains must also be of a different kind. As seen in that chapter, this may be observed in terms of the nature of responses to the consonance or dissonance of dyads/chords *qua* isolated auditory objects, on the one hand, and those resulting from experiencing them within a musical context, on the other. Yet, it seems important to consider whether those differences may also be observed in the aesthetic judgements that are elicited by certain auditory sensations, given that I have taken both judgements and responses to form the basis for an aesthetic verdict. To that extent, I want to argue that there is an important distinction to be made with regard to the nature of verdicts of consonance or dissonance in the two domains.

In the sensory case, I characterised the nature of verdicts in terms of consonance and dissonance assessments in the specific context of psychoacoustic experiments. As I argued in Chapter Eight (Section 2.4), participants should be able to distinguish between their rating of the sonorous quality of the sounds of dyads/chords, on the one hand, and their rating of the aesthetic experience of those auditory sensations, on the other. To that effect, the response-dependence biconditionals I proposed in Section 2 include references to both, under A and R, respectively. Both of these, I have maintained, must be satisfied alongside the conditions that stem from the nature of the non-aesthetic base, whether or not participants have epistemic access to the specificities of it. It was under this framework that I have characterised aesthetic verdicts of consonance/dissonance as pertaining to the sensory component. Conversely, I take the view that aesthetic verdicts in the case of 'harmony' should not be accounted for under those terms.

As indicated in Chapter Eight (Section 3.4), the auditory experience of tokens of isolated dyad/chord subtypes is not replicated as such when they are perceived within a musical context. This can be seen in the example of Schoenberg's emancipated dissonance, as described in that chapter. Specifically, the emancipation of the dissonance challenges the need for resolution both in a narrow sense – such that they need not be resolved by concords, in particular – and in a broad sense – under which the resolution constraint is altogether eliminated. This elimination of the resolution constraint marks a significant contrast between the sensory and the musical experience of dissonance. If dissonances are accepted and enjoyed *as such* in the musical experience of dyads/chords, the correlation between roughness and unpleasantness seen in assessments of sensory dissonance is absent. In other words, the emancipated dissonance represents the dissolution of the correlation between aesthetic judgements and responses in the context of 'harmony'.

Given that *A* and *R* are dissociated in a distinctively musical experience of harmony, I take the view that aesthetic verdicts on the consonance or dissonance of dyads/chords perceived within a musical context must be given in a different way. Specifically, I am adopting Frank Sibley's definition of verdicts as purely evaluative aesthetic judgements.[30] Under this proposal, the descriptive component of *A* in the sensory case is either irrelevant or of secondary importance in the case of 'harmony', such that it is the evaluative component of *R* that becomes central to aesthetic verdicts of the musical experience of consonance/dissonance. This, however, indicates that the response-dependence biconditionals are inadequate to account for the experience of 'harmony'.

3.2. The inadequacy of response-dependence

The contrast between the phenomenology of sensory consonance/dissonance and that of 'harmony' also has implications for the conceptual dimension of the relationship between aesthetic and non-aesthetic properties in those two domains. In this section, I will therefore revisit the concepts of concord and discord, which will be considered in light of Sibley's characterisation of the relationship between aesthetic and non-aesthetic qualities. In view of the conditions under which that relationship may be said to obtain for each component of 'musical consonance', I will subsequently argue that the metaphysics of response-dependence is not suitable for accounting for the experience of 'harmony'.

[30] Sibley, F., 'Aesthetic Concepts', In Benson, J., Redfern, H. B., Roxbee Cox, J. (eds.), *Approach to Aesthetics*, Oxford: OUP, 2001, p. 33.

Earlier in Section 2.3, I offered an analysis of the conceptual dimension of sensory consonance and dissonance. Under my proposal, the concepts of concord and discord are best described under (6) and (7), as a consequence of the terms of the sensory-consonance and sensory-dissonance biconditionals in (4) and (5), respectively. My account of sensory consonance/dissonance *qua* concepts is thus derived from their nature *qua* properties, thus entailing a correspondence between (4) and (6), on the one hand, and (5) and (7), on the other. In that section, I also argued that (6) and (7) are able to provide clear definitions of 'concord' and 'discord' by focusing on dyad/chord subtypes, whilst still maintaining the response-dependence of consonance/dissonance as expressed under *A* and *R*.

However, such correspondence is not observed in the case of 'harmony'. This disparity may be illustrated by considering, once again, the musical experience of resolution. If, on the narrow sense identified above, a discord need not be resolved by a concord but may instead be resolved by a less dissonant dyad/chord, then the correspondence between (5) and (7) is violated. This is because the latter would not be experienced with an unfavourable aesthetic response of unpleasantness, since the dissonance is in effect performing the role of repose traditionally assigned to concords, instead. This example thus indicates how dissonances acquire distinctive properties in the musical experience of dyads/chords, in spite of the roughness of their auditory sensations.

Hence, it seems that the relationship between aesthetic and non-aesthetic properties in the context of 'harmony' must be accounted for differently. In order to do so, my analysis will draw from Sibley's understanding of how that relationship may obtain on a conceptual level. Although Sibley rejects the possibility that non-aesthetic properties may in any way provide sufficient conditions for the ascription of aesthetic ones, he acknowledges that aesthetic concepts seem to be governed by some logical conditions. Specifically, Sibley identifies three ways in which the relationship between aesthetic and non-aesthetic qualities may be conceived, namely in terms of: (i) logical necessity; (ii) logical presupposition; and (iii) in terms of a less stringent relationship between aesthetic and non-aesthetic qualities.[31] These conceptual distinctions are particularly illuminating for considering the contrast between the two components of 'musical consonance'. In view of the account I have pursued in this chapter, I take sensory consonance/dissonance *qua* concepts to be best understood under (ii). This is because the psychoacoustic properties of concords and discords may be appropriately spoken of as logically presupposed by the aesthetic properties manifested in *A* and *R*. Yet, the same cannot be said of the case of 'harmony'.

[31] Sibley, F., 'Aesthetic and Non-aesthetic', In Benson, J., Redfern, H. B., Roxbee Cox, J. (eds.), *Approach to Aesthetics*, Oxford: OUP, 2001, p. 47.

Dovetailing with my account of the dual phenomenology of consonance/dissonance in Chapter Eight (Section 3.4), I want to suggest that this conceptual contrast is equally indicative that the two components must be accounted for differently. Whether or not the concepts of concord and discord are familiar to listeners, it is not necessary for them to *understand* the concepts of harmonicity and interference, in particular, in order to submit aesthetic verdicts that would conform to the characterisation of those concepts under biconditionals (6) and (7) above. Hence, it does not seem to be the case that the sensory component is to be understood under (i). Instead, the assumption held by psychoacoustic researchers is best encapsulated by (ii), in that the presence of harmonicity, for example, can be said to be logically presupposed from the auditory sensation of smoothness and corresponding aesthetic response of pleasantness to concords.

Yet, neither (i) nor (ii) would serve as a characterisation of the relationship between aesthetic and non-aesthetic properties in the case of 'harmony'. Unlike the case of concords and discords under the sensory component, the psychoacoustic properties of dyads/chords as experienced within a distinctively musical context are not logically presupposed by their consonance or dissonance. Rather, there can only be said to be, in Sibley's own terms, "a characteristic association or relationship which, though still not merely contingent, is much less stringent than logical necessity or presupposition" – as indicated in (iii) above.[32] Hence, given that consonance and dissonance *qua* musical concepts cannot be conceived under (ii) as per the sensory component, they cannot be adequately described under the metaphysics of response-dependence – so they must be accounted for in a different way.

One of the implications of this contrast between the two components of 'musical consonance' is that, under my proposal, the concepts of concord and discord cannot be used indiscriminately across the sensory and musical domains. But the sensory component and 'harmony' do share the same psychoacoustic base. As seen in my account of sensory consonance/dissonance as response-dependent properties, I took the non-aesthetic base to consist of certain acoustic and psychoacoustic properties which determine the quality of certain auditory sensations – and these are also part of the non-aesthetic base in the case of 'harmony'. Hence, whether one takes the musical experience to be independent from the circumstances of sound production or not, the psychoacoustic basis of it remains as such.[33] Yet, there is more that can be said of the non-aesthetic base in the case of 'harmony'.

---

[32] Ibid., p. 48.

[33] Although there is an important contrast between the technical nature of psychoacoustic experiments and the contexts in which music is experienced, it is to the psychoacoustic properties at work in sound perception – both linear and non-linear ones, as detailed in Chapter Four (Section 1) – that I am referring here.

3.3. Special aesthetic supervenience

While the sensory component is best accounted for via the workings of response-dependence, the same cannot be said of the case of 'harmony'. Hence, in this final section I want to suggest that the consonance or dissonance of tokens of dyad/chord subtypes when perceived within a musical context can only be said to supervene on the subvenient base.[34] Specifically, I will argue that the supervenience relation between consonance/dissonance *qua* musical properties and the subvenient base is best described under an interpretation of Jerrold Levinson's thesis of special aesthetic supervenience for the aesthetics of 'harmony'.

As indicated in the previous section, some important aspects of the non-aesthetic base of dyads/chords are determined by certain acoustic and psychoacoustic properties. But, in the case of 'harmony', the understanding of the non-aesthetic base must be expanded – which I will seek to undertake with reference to Levinson's taxonomy. Specifically, he identifies three different kinds of non-aesthetic properties: (a) structural ones, which are perceptible attributes regarded as fundamentally intrinsic to the object (e.g. the presence of harmonicity, in the case of sensory consonance); (b) substructural ones, which consist in any physical attribute that is not perceived as such under normal or appropriate conditions (e.g. beating partials that may be completely masked, in the case of sensory dissonance); and (c) contextual ones, which Levinson describes in terms of an "important relation of the object to the artistic context in which it occurs".[35]

Yet, if aesthetic supervenience is to work for my account of the metaphysics of 'harmony', then Levinson's taxonomy must be refined further. Beyond the fact that his characterisation of (c) is somewhat vague (which may be a deliberate attempt to render it applicable to any given aesthetic context), the main question seems to be whether or not (c) should be treated as a non-aesthetic property. Although it is not my intention to challenge its suitability to account for other aspects of the musical experience, in the specific case of 'harmony' it seems that the context should not be taken as such. Indeed, *context-sensitivity* is precisely what distinguishes it from the sensory component, thus making it distinctive of musical aesthetics. For that reason, I want to argue that (c) should not be grouped with structural and substructural attributes – which are, instead, typically non-aesthetic.

---

[34] I am adopting O. R. Jones's definition of 'supervenience' in the *Oxford Companion to Philosophy* as a relation of ontological dependence whereby one set of properties is said to supervene (i.e. depend) on another set if there could not be a difference in the first set (in the present case, the aesthetic properties of consonance/dissonance) without there being a difference in the second (i.e. the properties in the subvenient base), although there could be a difference in the second without a difference in the first. Jones, O. R., 'Supervenience', In Honderich, T. (ed.), *The Oxford Companion to Philosophy*, Oxford: OUP, 2005, p. 903.

[35] Levinson, J., *Music, Art, and Metaphysics*, Ithaca: Cornell University Press, 1990, p. 135.

This modification, however, raises an important question of its own. If contextual attributes are *not* non-aesthetic, then it is either the case that they are themselves aesthetic attributes or they must be understood in another way. In response, although it may seem counterintuitive not to conceive of them as aesthetic attributes, I take it to mean that they are not intrinsically aesthetic. Still, whether they are taken as aesthetic attributes or not, they will be *phenomenologically attributed* to the perception of dyads/chords in a given musical context. For instance, this may be observed in operas in which certain visual elements may contribute to the perceived tension of particular chords – an example of which is one of the long harmonic progressions of Wagner's *Tristan und Isolde*, as per the earlier example.

Having refined the understanding of (c), I want to suggest that Levinson's thesis of 'special aesthetic supervenience' may be taken as a baseline for assessments of the consonance or dissonance of dyads/chords in a distinctively musical context. Levinson gives his iteration of a special kind of aesthetic supervenience as follows: "two objects ... that differ *aesthetically*, but neither contextually nor (purely) substructurally, necessarily differ *structurally* (i.e., in some perceivable but non-aesthetic feature); that is, there could not be two contextually and substructurally identical objects that were aesthetically *different*, and yet structurally *identical*".[36] Specifically, I take this iteration of aesthetic supervenience to be particularly illuminating for the metaphysics of 'harmony' for three main reasons.

First, it enables us to incorporate acoustic properties under structural attributes, in that they are the physical properties underlying sound production without which no chord type can be instantiated and no consonance or dissonance can be perceived. Second, by including reference to substructural attributes, Levinson's thesis allows for the presence of psychoacoustic properties that may or may not be perceived as such, which may or may not contribute decisively to consonance and dissonance perception. Most importantly, it emphasises the asymmetry between structural attributes and substructural ones in the subvenient base. Third, the special thesis allows for contextual attributes to perform a key role in differentiating between auditory experiences of chords *qua* musical individuals. Specifically, contextual attributes may be designated with reference to the properties that tokens of dyads/chords acquire in musical contexts, such as harmonic function and that of progression. Overall, under the special thesis, the interdependence between music and psychoacoustics is still present – in terms of structural and substructural attributes – but consonance and dissonance *qua* musical properties are also accounted for in terms of contextual attributes, without being in any way determined by psychoacoustic properties.

[36] Ibid., p. 136.

Hence, unlike the response-dependent aesthetics of the sensory component of 'musical consonance', there cannot be any biconditional that can be used to predict consonance and dissonance verdicts in the case of 'harmony', given the specificities of the context in which they may be given. Instead, I have suggested that the consonance or dissonance of dyads/chords in a musical context can be said to supervene upon a combination of structural, substructural *as well as* contextual attributes, in that the context-sensitivity of 'harmony' affects the perception of the properties of the subvenient base. Therefore, the presence of structural attributes is a necessary but insufficient condition for the attribution of consonance or dissonance to chords *qua* musical individuals. Indeed, it is solely in terms of contextual attributes that their musical properties may be best discerned.

It is under the approach I have sought to describe in this chapter that, I believe, the metaphysics of 'harmony' is best distinguished from that of the sensory component of 'musical consonance'. In spite of this contrast, another consequence of the thesis of aesthetic supervenience is that the two components – i.e. the psychoacoustic and the musical – can be seen to coexist under the same subvenient base, thus conforming to my thesis of the reconciliation between music and psychoacoustics whilst also respecting the dual phenomenology of consonance/dissonance *qua* sensory and musical properties. Yet, there is more that can be said of the consonance or dissonance of tokens of dyads/chords – and this concerns both the psychoacoustic and the musical domains. Specifically, this arises from differences in tuning and temperament methods, as will be considered next.

Chapter Ten: Tuning and Temperament

There is more to the metaphysics of tonal harmony than meets the 'ear'. Much of what has been said about consonance and dissonance in previous chapters revolves around a tacit assumption about the nature of chords: namely, that they result from fixed pitches. Granted, as outlined in Chapter Five, chord types may be described as atemporal and modally fixed structures. Yet, their instances are ontologically dependent upon specific tuning conditions. Indeed, beyond the confines of the system of twelve-tone equal temperament is a universe of complex tuning possibilities. Hence, underlying the metaphysics of tonal harmony is a crucial question concerning the ontological instability of chord tokens, which stems from the lack of a single definitive tuning system. The purpose of this closing chapter is thus to consider some of the issues posed by the problem of tuning for consonance/dissonance assessments. To that end, my main proposal constitutes a refinement of my earlier thesis of the dual phenomenology of consonance/dissonance. Specifically, I will propose that a further distinction ought to be made in order to address some divergences presented by alternative tunings and temperament systems.

1. The problem of tuning

Historically, the aim of most tuning systems has been that of generating the most agreeable – or, perhaps, the most versatile – musical sounds as physically possible, depending on which methods are used for arriving at the desired outcome. Yet, underlying the plethora of tuning systems that have been devised is a problem that has not been settled by the adoption of twelve-tone equal temperament as the standard in Western music. Hence, in this section I will provide an outline of that problem in view of the issues posed by variances in tuning and temperament. While I do not intend to account for the minutiae of the history of the problem itself, I aim to describe it in terms of what, I believe, still constitutes the crux of it to this day. Specifically, I take the problem to comprise two horns. On the one hand, it stems from a theoretical emphasis on specific numerical relationships for certain musical intervals. On the other hand, it results from the practical demands of musicianship regarding the appropriate tuning of instruments and, most importantly, what kinds of harmonies they should be expected to produce. It is, therefore, under this twofold approach that I shall characterise the problem of tuning in what follows.

1.1. Intervals and ratios

As I understand it, there are two horns to the historical problem of tuning. First, it results from an impossibility of reconciling different intervallic relationships in terms of numerical ratios. Yet, the problem becomes even more complex when those intervals are 'tempered' (i.e. they are either narrowed or widened) in an attempt at reaching a compromise which is not afforded by a strict observance of ratios – which is the second horn of the problem. This section is concerned with the first horn, which is historically at the root of the problem itself and long predates the discussion around temperament.

The first horn of the problem of tuning concerns the importance of numerical ratios for the understanding of music and, consequently, that of harmony – both of which were considered by the Ancient Greeks, although 'harmony' then had a different connotation.[1] In particular, certain intervallic relationships were exalted as specimens of *symphonon* (i.e. they were deemed agreeable). The most well-known accounts that have reached us point to the principles of harmonics offered by those generally referred to as 'the Pythagoreans'. They are said to have been concerned with the correlation between certain intervals and whole-number ratios. Specifically, it was the ratios between numbers in the *tetrachtys* – namely, the first four integers – that were taken to be the key parameters for tuning.

As described by later theorists such as Ptolemy and Boethius, two types of ratios may be formed by those numbers: they may be multiple – namely, 2:1, 3:1 and 4:1 – or superparticular – such as 2:1, 3:2 and 4:3. On the Pythagorean view, these ratios were taken to correspond to certain intervals: the first one results from dividing a string into two by creating a node in the middle, which corresponds to the octave at 2:1, followed by further divisions of the string that established the fifth at 3:2, the fourth at 4:3, the twelfth at 3:1 and the double octave at 4:1.[2] Given that the last two may be taken as instances of the first two intervals but an octave above, the Pythagorean parameters of tuning came to be defined in terms of the *diapason* (i.e. the octave), the *diapente* (i.e. the fifth) and the *diatessaron* (i.e. the fourth). From these, all other intervallic relationships were derived, including the tone and the semitone.

[1] Gerald Abraham notes that the pluralised term '*harmoniai*' generally referred to tonal systems which resulted from combinations of tetrachords. The latter consisted of four notes spanning the interval of a fourth, and were seen as the foundation of Greek music. That said, as Tenney points out, the term '*harmonia*' was primarily associated with the notions of balance and order, which were originally extraneous to music. The term eventually came to be understood under a "purely melodic" connotation, rather than the vertical aspect of music with which it later became a synonym. Abraham, G., *The Concise Oxford History of Music*, Oxford: OUP, 1979, pp. 28-30; Tenney, pp. 9-10.

[2] Boethius, A. M. S., *Fundamentals of Music*, New Haven/London: Yale University Press, 1989, pp. 77-81; Solomon, J., *Ptolemy Harmonics: Translation and Commentary*, Leiden: Brill, 1999, p. 17.

Yet, upon considering the sequencing of these intervals along a succession of octaves – as found on a standard keyboard instrument – an irreconcilable difficulty arises. Although this may not have created an immediate problem for Greek musicians at the time, a mathematical inconsistency lurked within that system. If we take the note C1 on a keyboard, which is on the low-end of the pitch register, and subsequently progress from it for seven octaves, we arrive at the same pitch class on the high-end of the register, at C8. If we were then to start again at C1 and move up twelve fifths instead, (i.e. following what has come to be known as the 'circle of fifths', namely: C–G–D–A–E–B–F♯–C♯–G♯/A♭–D♯/E♭–A♯/B♭–F–C), we should arrive at the same note as in the first case. However, although it would appear that we do – since on the keyboard the end destination is the same note (i.e. C8) – the problem is that by following the Pythagorean tuning principles we do *not* – and this may be understood by doing some mathematical calculations.

Under the Pythagorean system, the ratio of the interval between C1 and C8 in the first instance is $(2:1)^7$, which equals 128:1 (where the exponent represents the number of octaves). In the case of the circle of fifths, the same interval is numerically expressed as $(3:2)^{12}$, the simplification of which approximately equals 129.75:1. As seems evident, the ratios do not coincide, which means that twelve Pythagorean fifths do not mathematically equate seven octaves – but, both theoretically and musically, they would be expected to. The last fifth of the series exceeds the last octave by a small margin, preventing the circle of fifths from being completed at its original starting point. This discrepancy, which came to be known as the 'Pythagorean comma', is one of the sources of the problem of tuning, seeing as the Pythagorean fifth that closes the circle is somewhere between C and C♯.

In addition to the Pythagorean comma, there was yet another source of difficulty for a tuning system based on whole-number ratios: namely, the 'syntonic' comma, also known as the Ptolemaic or Didymic comma. This time, the irreconcilable difficulty involves major thirds, in particular. Although the Pythagoreans did not class is as a concord, subsequent historical developments in the understanding of harmony – as indicated in Chapter Six (Section 2) – led to the belief that thirds were also consonances, and that they too were correlated with superparticular ratios. The interval of a major third was then said to correspond to the ratio of 5:4 – whereas the minor third was to be found at 6:5. The possibility of adding these intervals, especially the major third, to the existing group of concords lies at the heart of another tuning system, namely, that of just intonation. In particular, one of the key principles underscoring just intonation is that the *purity* of those intervals expressed in low-number integer ratios must be preserved.

The defining feature of just intonation is the notion of purity. Specifically, pure concords are those which correspond to the ratios of small integers, namely, the octave (2:1), the perfect fifth (3:2), the perfect fourth (4:3), the major third (5:4) and the minor third (6:5). The possibility of achieving these intervals on a monochord[3] turned just intonation into the ideal system against which instruments were frequently measured from the fifteenth century onwards. In particular, the discovery of the harmonic series came to reinforce this ideal of sonorous purity. But it probably did not take long for early theorists to realise that introducing a different ratio for the interval of a third would result in further mathematical difficulties. Instead of 5:4, the Pythagorean major third was at a significantly larger ratio of 81:64 – which was evidently far removed from their notion of a concord.

Hence, to insist on the purity of that interval, alongside that of fifths and fourths (at the ratios of 3:2 and 4:3, respectively), meant that the difference between the Pythagorean major third and its 'just' counterpart would constitute yet another unwanted remainder in the system: that is, the syntonic comma, calculated at the ratio of 81:80. In view of the agreeable sonority of the just major third, this comma became another mathematical constant with which music theorists had to grapple, and thus came to compound what I am characterising as the problem of tuning.[4] Most importantly, the possibility of preserving the purity of those intervals can be said to be at the root all historical forms of tuning and temperament that ever came to be conceived.

To be clear, any tuning system that is aimed at simultaneously preserving the ratios of the octave and the fifth (or the octave and the fourth, for that matter) cannot be a perfectly closed system, just as one designed to maintain the purity of major thirds in addition to those concords is equally mathematically unworkable. As Kyle Gann has described it, "those ratios are incommensurate because no power of 5 [...] will ever be divisible by 3 or 2, and no power of 3 will ever be divisible by 5 or 2, and so on".[5] Hence, in view of the implications of these inconsistencies for music making, it became clear to many that the only way to make tuning viable was by tampering with those ratios, and thus tempering their corresponding intervals – which is not a solution, but a practical compromise.

[3] As J. M. Barbour notes, "the seeds of just intonation had been sown early in the Christian era, when Didymus and Ptolemy presented monochords that contained pure fifths and major thirds" – long before the system eventually came to be regarded as the ideal tuning method for creating the most perfect harmonies. Barbour, J. M., *Tuning and Temperament: a Historical Survey*, Mineola, New York: Dover, 2004, p. 89.
[4] Furthermore, another disruptive constant was generated by the admission of the syntonic comma into the problem of tuning. The difference between the two commas gave rise to another remainder which came to be known as the 'schisma', calculated to be at the ratio of 32805:32768. But the presence of the two commas – in any tuning system that is fundamentally based on superparticular ratios – was sufficient to create a mathematical conundrum in itself.
[5] Gann, K., *The Arithmetic of Listening*, Urbana: University of Illinois Press, 2019, p. 16.

1.2. The art of temperament

The second horn of the problem of tuning, which I will aim to characterise in this section, concerns the attempt to forge a practical compromise so as to render the long-favoured intervals as close to pure as possible – whilst acknowledging that the tempering of any given concord will inexorably bear consequences for other intervals. Briefly, to engage in the art of temperament is to accept that a strict observance of whole-number ratios is a theoretical ideal which, albeit feasible on a monochord, is musically impracticable.

The problem of tuning gained a new layer of complexity with the introduction of tempered intervals. According to Barbour, the defining characteristic of a *tuning* system, understood in historical terms, is that those relationships are or can be expressed in rational numbers, whereas *temperament* refers to a modification of a tuning where that is not the case – or at least one in which only some consonant intervals may be expressed with integers.[6] In particular, one of the main concerns consisted in calculating how the Pythagorean comma could be divided and distributed across the register in order to accommodate those intervals that were not deemed consonant under the Pythagorean system. Indeed, the possibility of having a tuning system which could incorporate major thirds that were as close to pure as possible became highly desirable. This was one of the principal aims of theorists who advocated what came to be known as 'mean-tone temperament'.

Given the prevailing influence of Pythagorean principles during the mediaeval period, the interval of a third was generally taken to be dissonant. As seen earlier, the ratio of a major third, in particular, was far removed from the Pythagorean notion of a concord, thus departing from its just counterpart by a significant degree. Yet, changing attitudes to the classical ideal of harmony increasingly led to a new understanding of consonance in some quarters. As Scruton points out, as early as the twelfth century both the major and the minor thirds were already taken to be 'imperfect consonances'.[7] Accommodating thirds as concords, in particular, was precisely what some theorists set out to achieve in devising a range of mean-tone temperament methods. Invariably, these entailed widening the size of a just major third, but only to a tolerable level, in order to preserve the sonority of the pure interval as much as possible. Although most thirds were still tempered by a certain degree across the varieties of the mean-tone system, one in particular – namely, that known as 'quarter-comma' – managed to include pure thirds, whilst others comprised a mixture of both tempered and pure forms of that interval. This, however, came at a certain cost.

[6] Barbour, p. xii.
[7] Scruton, 1997, p. 241.

In order to incorporate the agreeable harmony of purer thirds, mean-tone theorists conceded that perfect fifths should be narrowed by a small margin. Although this would have been seen as anathema to those of a classical inclination, the practice became the hallmark of mean-tone temperament, such that its methods came to be referred to by the numerical fraction of a Pythagorean comma by which the fifths should be tempered.[8] In addition to flat-sounding fifths, the main difficulty arising from this new temperament method was a consequence of compressing those intervals. Specifically, one of those fifths – frequently placed between G♯/A♭ and D♯/E♭ but not exclusively – turned out to be considerably larger than the others, so that the circle of fifths could be completed. This, however, effectively meant that a 'problem interval' was introduced into the system.

In essence, the issue was that this new interval was less of a sharp fifth but, instead, closer to a diminished sixth. Reportedly, this distinctive interval is said to have sounded so unpleasantly that one sought to avoid it at all times. Hence, even though mean-tone systems had the advantage of admitting both major and minor thirds into the group of concords whilst accommodating the Pythagorean comma, they brought with them a highly undesirable consequence – namely, an interval beset by the acoustic phenomenon of beating to such a degree that its sound became associated with the howling of a wolf. This interval of 'the wolf' – as it came to be known – would then lurk around any mean-tone system, thus adding another layer to the problem of tuning.

Given that temperament is an art, some sought to find ways of addressing that difficulty. Several proposals were made to handle the wolf, whether by means of narrowing fifths by different fractions of a comma or by placing the wolf under intervals less frequently used. A common way of minimising its impact was by avoiding that interval altogether, which meant that composers and musicians refrained from composing and playing in certain tonalities. Yet, this was perceived by some as an increasing limitation on the possibilities of musical expression, which led some theorists to seek ways to remove the wolf from the system once and for all. In order to do so, they experimented with more irregular forms of temperament, which would include both pure and tempered fifths so that all keys could be playable on a standard keyboard instrument. These new attempts at accommodating one of the most undesirable consequences of mean-tone temperaments came to be referred to as 'well temperaments', of which there is also a large variety. But the most distinctive feature of this incongruous family of temperaments is fundamentally the irregularity of interval sizes, which rendered each method a universe of its own.

[8] Mark Lindley identifies one-quarter, one-fifth, one-sixth, two-sevenths and two-ninths of the comma as the most common fractions. Lindley, M., *Lutes, Viols and Temperaments*, Cambridge: CUP, 1984, p. 44.

Granted, the very nature of the art of tempering intervals entails that making some of them sound better most likely means that others will not. Yet, amongst other difficulties stemming from the lack of uniformity of interval sizes, the main problem with well temperaments was that some tonalities would sound more agreeable than others, and others still were said to sound most unpleasantly. Underlying these differences was a more pressing concern affecting all forms of irregular temperament. Given that the purity of the octave was held to be fundamental, different sizes of fifths and thirds implied different sizes of fourths and sixths – seeing as they are their respective octave complements. Worse still, this also meant that different sizes of tones and semitones were distributed across the register – which, to some, was too high a price for eliminating the wolf.

In view of these difficulties, music theorists and mathematicians alike started to contemplate the possibility of a system that would achieve what, in their view, would be a better form of compromise. Given the limitations of the classical, ratio-based tunings and the shortcomings of different varieties of temperament, a tuning method that prioritised uniformity and utility whilst preserving some of the most desirable consonances as much as possible must have seemed like the most pragmatic solution, which they found in what came to be known as (twelve-tone) 'equal temperament'. This current system is fundamentally based on the uniformity of tone and semitone sizes across the register, such that the octave could comprise twelve equal semitones – whence it received its name.

This, however, encountered a considerable degree of resistance in many circles due to other undesirable outcomes that equal sizes of semitones created. For instance, mathematicians such as René Descartes and Isaac Newton both decried the system as an unsatisfactory compromise, to say the least. As Stuart Isacoff notes, while Descartes took it to constitute a violation of the correct proportions of concords, Newton similarly claimed that it was "unworthy of philosophers to contrive the corrupting of the true proportions".[9] Long after the system became firmly established as the standard method in Western music, this indictment of equal temperament for violating the pure ratios of intervals has also echoed through to the twenty-first century. In a recent work, Ross W. Duffin has condemned equal temperament for incurring a "terrible musical cost" and, quite simply, ruining harmony.[10] Furthermore, James Young has suggested that it deprives Western music of those microtonal differences that give different keys a distinctive character.[11]

---

[9] Isacoff, S., *Temperament*, London: Faber and Faber, 2002, p. 196.
[10] Duffin, R. W., *How Equal Temperament Ruined Harmony*, New York: W. W. Norton and Co., 2007, p. 27.
[11] Young, J. O., 'Key, Temperament and Musical Expression', *The Journal of Aesthetics and Art Criticism*, 49, 3, 1991, p. 238.

2. The consequences of uniformity

Having characterised the problem of tuning above, in this section I aim to consider some of its implications for the metaphysics of tonal harmony. As seen earlier, the problem of tuning stems from two different priorities. Specifically, they concern the observance of certain numerical ratios in intervallic relationships, on the one hand, and the practical need to temper interval sizes in order to accommodate the commas and the wolf, on the other. In view of the clash between these priorities, I will identify some consequences of the uniformity of equal temperament which have a bearing on assessments of consonance and dissonance. In particular, I will point out why the adoption of equal temperament has not settled the question of the observance of specific ratios and proportions.

Granted, by enabling the uniformity of interval sizes across all octaves, equal temperament eliminated not only the commas and the wolf but also did away with the need for any split keys, which became an occasional alternative as the art of temperament developed.[12] Under equal temperament, certain notes became equivalents with one another (e.g. F♯ and G♭ became equal in pitch, which was not always the case under other systems). These advantages conferred a degree of versatility to equal temperament that is probably unparalleled by any other tuning. Yet, although its emergence was heralded by many as a culmination of the art of tempering, equal temperament did not solve the problem of tuning as such but only shifted it into other kinds of undesirable consequences.

Specifically, one of its several consequences for tonal harmony stems precisely from the fact that it renders all semitones – and hence all tones – equal in size. Even though the unequal tempering of intervals virtually meant that some of the most cherished numerical ratios had to be abandoned, in preserving the inequality of semitone sizes the other forms of tuning and temperament were able to retain the ratios of major and minor semitones in terms of rational numbers, which were typically expressed in superparticular ratios. But, by establishing the equal division of the whole tone, equal temperament would seem to have eliminated the relevance of observing numerical ratios altogether[13] – the only exception in an equal-tempered system being the octave, which was still kept at the ratio of 2:1.

[12] As Isacoff indicates, one solution to the problem of tuning was that of constructing split-key keyboards, which differentiated between accidentals – such as G♯/A♭ – so that purer intervals could be sounded. Yet, although some specimens were built, they were taken to be cumbersome and unsatisfactory. Isacoff, p. 104.

[13] A new way of measuring intervals emerged in the late nineteenth century when Alexander Ellis divided the octave into 1200 units which he called 'cents'. Barbour, p. vi. On that basis, a semitone is equivalent to 100 cents, regardless of which pitch classes are separated by it. This system also meant that interval sizes could be comparable across different systems. For instance, while an equal-tempered fifth equals 700 cents, the perfect fifth is measured at 701.955 cents and the mean-tone fifth of the quarter-comma system contains 697 cents.

Hence, the octave can be said to be the only true concord within an equal-tempered system, in that all other intervals are tempered by a certain amount when compared against the benchmark of just intonation. Most importantly, the widespread acceptance that the tone could be split into equal semitones carries significant importance, which is nonetheless understated in the literature. This is because of its consequences for the foundational principles of traditional tonal harmony that were proposed by Rameau, as outlined in Chapter Six (Section 2). Although Rameau's views on tuning itself are uncertain, some of his observations in his treatises have been taken as evidence that he favoured equal temperament.[14] But, if that is the case, two important questions arise which have a bearing on the perception of the consonance/dissonance of dyads and chords.

First, not only did Rameau acknowledge the different sizes of tones and semitones but he also stated that the "perfect diatonic system" is one which follows the pattern 'mT–MT–S–MT–mT–MT–S' (where 'MT' stands for 'major tone', 'mT' for 'minor tone' and 'S' for 'major semitone').[15] Even more controversial is his description of a chromatic scale, which contains not two but three different sizes of semitones, corresponding to the ratios of 25:24 (minor semitone), 16:15 (major semitone) and 27:25 (the maximum semitone, which he indicates as the distance between D–E♭, F♯–G and A–B♭).[16] As a consequence of these irregular scales, both major and minor harmonies would be affected as and when tokens of triad types are sounded across their respective diatonic scales. It is, therefore, puzzling to imagine how Rameau must have been able to reconcile these observations with the system of equal temperament. Perhaps unsurprisingly, these remarks do not feature in a later publication of his principles of composition in *A Treatise of Music* (1779).

Secondly, Rameau's treatises still echoed the classical practice of establishing a correspondence between certain intervals and specific numerical ratios. In particular, he was concerned with the fact that there can be two different ratios for all intervals but the octave, which result from "the difference between the major tone and the minor, both of which are present in the diatonic system; [namely] the difference is a *comma* whose ratio is 80:81". To illustrate it, Rameau explains that the interval of a fourth between C and F and that between D and G correspond to two different ratios.[17] Similarly, he also expressed a concern with the observance of certain proportions within perfect intervals, such as that of

---

[14] Barbour notes that "he vacillated ... in his adherence to it", having also suggested that "'the most perfect of all' temperaments [is] that in which 'the fifth is diminished by the ¼ part of a comma'" – even though he knew of the shortcomings of mean-tone systems. Barbour, p. 135.
[15] Rameau, 1971, p. 28.
[16] Ibid., p. 33.
[17] Ibid., p. 32.

major chords at 20:25:30, and that of perfect minor chords, at 20:24:30. These observations are particularly problematic when we consider the impossibility of realising these proportions under equal temperament given the uniformity of semitone sizes.

Whether Rameau eventually subscribed to an equal-tempered system or not, these considerations are particularly relevant for the metaphysics of tonal harmony. His observations concerning different sizes of tones and semitones, as well as his preoccupation with ratios and proportions, have consequences for our understanding of harmony, especially considering that equal temperament does not eliminate the difficulties arising from the problem of tuning. In particular, the instability of tokens of triad types comprising tempered intervals of thirds and fifths has an important bearing on the experience of sensory consonance/dissonance. It is in light of these considerations that I will revisit the dual phenomenology of consonance/dissonance in the final section below.

3. Consonance and *armonia concinnentia*

Having discussed the consequences of equal temperament in parallel with Rameau's principles, my argument in this final section comes in three stages. First, I will consider the relevance of different interval sizes for assessments of sensory consonance, with particular focus on the concords of traditional tonal harmony. Specifically, I will maintain that the sensory-consonance biconditional formulated in Chapter Nine is able to accommodate those differences; yet, in order to do so, I will propose a more refined reading of the biconditional in terms of a distinction between consonance and 'concinnousness', as part of the second stage of my argument. Correspondingly, I will subsequently propose a distinction between concords and 'concinnous' concords. It is in light of these distinctions that I will revisit the question of the importance of whole-number ratios in intervallic relationships *vis-à-vis* the current Western standard of equal temperament.

The first stage concerns the ways in which intervallic relationships came to be measured – as a result of being either narrowed or widened, and made either regular or irregular. Perhaps unsurprisingly, this has some important consequences for assessments of the sensory component of 'musical consonance', as described in Chapter 9 (Section 2). Specifically, I shall focus on those intervals that are taken to be the most basic consonances in traditional tonal harmony. Given that Rameau conceived of triads and their inversions as containing the most fundamental elements of harmony, the first stage of my argument concerns the intervals they comprise – namely, thirds, fifths, fourths and sixths.

As seen in Chapter Six (Section 2), one of the key intervals in tonal harmony is that of the third; and yet, this is probably the interval that effectively gave rise to the perceived need of temperament and led to the popularity of mean-tone systems. As seen earlier, these were aimed at preserving the purity of thirds, given that just intonation was the key parameter for evaluating consonance. But most forms of mean-tone temperament were not able to do so across all octaves. In the specific case of major thirds, although they were usually much closer to purity than their Pythagorean equivalents, they can nonetheless be said to have been tuned to sound as pleasant – or as tolerable – as possible, given the impossibility of realising the just intonation ideal throughout the pitch register. In turn, equal-tempered major thirds are considerably wider, albeit less so than Pythagorean ones.

Since the perfect fifth comprises a major and a minor third, tempering a major third inexorably meant tempering the minor third. Thus, the fate of minor thirds was usually determined by whatever was the case with their major counterparts. This is because of one important difference between them, namely, that sharpened minor thirds were reportedly more tolerable than sharpened major thirds. Most importantly, tempering these intervals also had consequences for their octave complements – i.e. the major and minor sixths. Similarly to thirds, these were not originally classed as consonances but later came to be conceived as such – in that sixths are found in both inversions of perfect triads.

Unlike thirds and sixths, fourths and fifths have a longer history *qua* concords, from its respective iterations as the *diatessaron* and the *diapente* of the Pythagoreans through to mediaeval times and beyond. But with the emergence of the art of temperament, the purity of those intervals came to be sacrificed for the sake of more consonant thirds, as seen in Section 1.2. In the case of fifths, when compared against their just/Pythagorean equivalent, mean-tone fifths were usually flatter – with the exception of the wolf – and well-tempered fifths were irregularly found across the register. In equal temperament, albeit uniformly distributed, fifths are narrowly flattened. However, in spite of these fluctuations, the fifth has generally been regarded as a consonance even if it was made imperfect. The same, however, cannot be said of its octave complement – i.e. the fourth. As indicated in the previous chapter, fourths have had a more controversial history, possibly because the widened mean-tone fourth may not have been as well-received as the slightly flattened fifth. One of the likely reasons for the perception of those fourths as dissonant is their closer proximity to the infamous interval of the tritone, i.e. the augmented fourth. Yet, as seen earlier, Rameau's treatises firmly included fourths in the class of consonances, in that they are found in both inversions of the perfect chord.

In view of the summary offered above, if we turn to assessments of sensory consonance in terms of the respective response-dependence biconditional as set out in Chapter Nine, several difficulties seem to emerge when the most typical concords in traditional tonal harmony are evaluated. In that chapter, I proposed under (4) that a given token of a dyad/chord subtype possesses consonance if, and only if, it elicits an auditory sensation of smoothness (or equivalent) and evokes a favourable aesthetic response of pleasantness (or equivalent) from Western-type listeners under certain non-aesthetic conditions – conditions which, as it happens, conform to the parameters of just intonation. Hence, tempered tokens of dyads/chords depart not only from the just intonation ideal but, under my account, from the conditions upon which sensory consonance is based.

As a consequence, there may also be a disparity between aesthetic judgements of auditory sensations of tempered tokens of concords, on the one hand, and the aesthetic response that may ensue, on the other. Specifically, although the aesthetic response evoked by tempered tokens of concords may still be favourable, their auditory sensations may not be described as smooth, given that the harmonicity of tempered intervals is to a greater or lesser degree compromised. This is particularly important when we consider that it is normally under equal temperament that perception of consonance is assessed in psychoacoustic experiments, since it is currently the standard tuning method. Since all intervals in equal temperament are indeed tempered – at the exception of the octave, which is preserved at the 2:1 ratio[18] – even small differences in interval sizes may prevent tokens of dyad/chord subtypes from meeting the non-aesthetic conditions as indicated in (4).

Furthermore, another important consequence of divergences in tuning and temperament concerns the ways in which participants may come to perceive certain dyads/chords that are tuned under alternative methods, when they are most likely to be familiar with equal temperament. This is because familiarity with the latter may lead some to perceive them as deviations from equal temperament and therefore 'out-of-tune'. This would be an undesirable consequence for the response-dependence biconditional, since its purpose is to establish a criterion of objectivity to assessments of the perceived consonance of chords *qua* isolated auditory objects. In order to address this potential issue of subjectivity in assessments of the consonance of dyads/chords, it seems that a further distinction must be made for the biconditional set out above to do its work properly.

---

18 That said, soundboard instruments such as the piano are subject to what Terhardt has described as the 'octave enlargement' phenomenon. As he explains it, the frequency ratio between two successive tones of equal pitch class (e.g. C3 and C4) as sounded on a piano is slightly greater than 2:1, even though the interval between them is "subjectively evaluated as a correct octave". Hence, under this characterisation, equal-tempered octaves do not strictly conform to the 2:1 ratio. Terhardt, 1974, p. 1066.

While it may be said that participants experience the consonance of dyads/chords in ways that may be classed as subjective, the acoustic properties which form the basis for the conditions specified in the biconditionals point to an objective component of that experience. As seen earlier, harmonicity is determined by whole-number ratios between the fundamental and the overtone series, such that the frequencies of the relevant complex tones will not beat. In addition, as indicated in Chapter Four (Section 1), even if they beat, the beats may not be audibly perceived as such. As a result, participants should be able to hear a perfect 'blending' of tones, which the sensory-consonance biconditional is designed to describe. As seen earlier, this was enshrined in the 'just' standard of intonation that had long served as the yardstick for alternative tuning and temperament systems. However, if assessments of the consonance of tokens of dyad/chord subtypes that have been tuned under a different method may be subject to comparison under the familiarity with equal temperament, the sensory component becomes once again vulnerable to the challenge posed by the subjectivist approach defended by Rehding, as indicated in earlier chapters.

For that reason, as the second stage of my argument, I want to propose a way in which these divergences between the harmonicity of just intervals and the potential or likely inharmonicity of tempered ones may be addressed so that the integrity of the sensory-consonance biconditional is preserved. To that end, I will adopt a specific terminology to mark the contrast between the consonance of just-tuned and tempered dyads/chords. Specifically, I will use the term 'concinnous' to designate the consonance of just-tuned tokens of dyads/chords. My motivation in introducing this terminology is to illuminate further the metaphysics of the sensory component of 'musical consonance'.

The term 'concinnous' has its roots in classical writings; for instance, Boethius referred to the *armonia concinnentia* of certain superparticular ratios (e.g. 3:2). This expression has been rendered in one of the English translations of his work as "the consonance of harmony".[19] A more recent usage of the term is found in a treatise of the late-eighteenth century by Thomas Salmon, where he differentiates between concinnous and 'inconcinnous' intervals – although he does not explicitly offer definitions for those terms.[20] Similarly, John Maxwell distinguished between concords and 'concinnous discords' nearly a century later, but he does not substantiate that distinction further.[21] Hence, despite the lack of a clear definition, the term 'concinnous' was still present in the intellectual debate around the time of the emergence of equal temperament.

[19] Boethius, p. 14.
[20] Salmon, T., *A Proposal to Perform Musick in Perfect and Mathematical Proportions*, London: Printed for John Lawrence, 1688, p. 24.
[21] Maxwell, J., *An Essay upon Tone*, Edinburgh: Macfarquar and Elliot, 1781, p. 27.

In view of the vagueness around the term 'concinnous', I will propose here a specific characterisation of it in terms of a distinction between 'concinnousness' and 'consonance'. Specifically, I want this distinction to mirror the contrast in the perception of dyads/chords whose frequency components are found to be in whole-number relationships and those whose components are *not*. Given that, of all tuning and temperament methods considered in Section 1, only just intonation consistently conforms to what we understand today as the presence of harmonicity, I am going to use the term 'concinnous' to refer strictly to just-tuned dyads/chords, and I will describe 'possessing concinnousness' as strictly conforming to the acoustic principles that underpin the notion of harmonicity. On that basis, although tempered tokens of dyads/chords may be deemed consonant under (4), I want to argue that they can be said to possess concinnousness if, and only if, they meet the non-aesthetic conditions *strictly in acoustic terms*. This means that, although tempered tokens may be given verdicts of consonance – e.g. when their lack of harmonicity is not perceived as such by the auditory system – under my proposal they cannot be said to possess concinnousness, in that they do not strictly conform to whole-number ratios.

In turn, as a third stage of my argument, I want to argue that the same can be said of sensory consonance *qua* concept. Under the iteration of (6) in Chapter Nine (Section 2.3), *x* is a concord if, and only if, its designated or salient token elicits an auditory sensation of smoothness (or equivalent) and evokes a favourable aesthetic response of pleasantness (or equivalent) from Western-type listeners under the same non-aesthetic conditions that operate in (4). But, in view of the distinction between consonance and concinnousness, only just-tuned tokens of dyad/chord subtypes can be said to be concinnous concords. Conversely, although tempered tokens may satisfy the conditions to be deemed concords, they would instead be 'inconcinnous'. By 'inconcinnous' I mean that the frequency components within a given dyad/chord are *not* in whole-number relationships – i.e. they are inharmonic – such that 'inconcinnousness' refers to the presence of inharmonicity.

In view of the above, I want to revisit the question of the importance of whole-number relationships *vis-à-vis* the standard of equal temperament by way of conclusion. As seen earlier, the history of tuning and temperament is beset by the question of whether numerical ratios are relevant for understanding the nature of harmony. Yet, the rise of equal temperament has been taken by some to have rendered this question obsolete. Contrary to this view, I take the considerations in this chapter to indicate that overcoming certain practical difficulties of unequal temperaments does *not* eliminate the correspondence between certain intervallic relationships and specific whole-number ratios.

My main proposal in this final chapter has been that a distinction must be drawn between tokens of dyads/chords that respect whole-number relationships and those that do not. As seen earlier, one of the consequences of the uniformity of equal temperament is precisely that, excepting the octave, none of its intervals observe the ratios that reflect the presence of harmonicity. Hence, although one may be able to speak of the sensory consonance of equal-tempered dyads and chords, under my proposal one will never be able to speak of them as concinnous concords. Granted, this may not seem to be of any interest to composers and musicians alike. Yet, after attempting an account of the metaphysics of tonal harmony, my final conclusion is that it should be. This view echoes that of Descartes, who – as Isacoff notes – stressed the importance of distinguishing between an interval ratio that may sound pleasant and one that is a true concord: while it may not be possible to calculate which chords are the most beautiful, we are nonetheless able to determine those that are the most perfect.[22]

[22] Isacoff, pp. 174-175.

Coda

As can be seen, an enquiry into the metaphysics of tonal harmony poses some interesting philosophical questions not only for the aesthetics of music but also for the ontology of sound. In this thesis, I have sought to answer some of these questions by providing an analysis of chords *qua* sound events, auditory objects and musical individuals in Part I. Given that chords are sounds, I first considered an overview of the philosophical debate on the nature of sound in Chapter One, which I construed in terms of a problem that needs to be addressed in order to illuminate our understanding of the nature of chords. Subsequently, I proposed a solution to that problem in Chapter Two, which I characterised under a thesis of reconciliation between the Wave and the Event View, on the one hand, and the Event and the Object View, on the other. The first two chapters were thus pivotal to my account of the ontology of chords in the remainder of Part I, where I described how chords may be understood *qua* sound events (Chapter Three), auditory objects (Chapter Four) and musical individuals (Chapter Five). Taken together, these three chapters constitute the core of my proposal in the first part of this work.

Subsequently, I aimed to provide an account of the aesthetics of chords with specific focus on consonance and dissonance *qua* sensory and musical properties. The second part of this study was introduced with the problem of consonance, which I sought to address by giving an account of the aesthetics of chords *qua* auditory objects and musical individuals. My main proposal in Part II is thus a response to that problem, which was outlined in the last section of Chapter Six and substantiated over the ensuing chapters. In Chapter Seven, I offered an overview of the psychoacoustics of tonal harmony. This subsequently informed my account of sensory consonance/dissonance in Chapter Eight, which also includes an account of 'harmony' as the distinctively musical component of 'musical consonance'. In addition, I argued for a conciliatory approach between music and psychoacoustics whilst acknowledging the dual phenomenology of consonance and dissonance. My analysis of the two components was substantiated further in Chapter Nine, where I accounted for the metaphysics of the sensory component and that of 'harmony' in terms of response-dependence and aesthetic supervenience, respectively. Subsequently, in the closing chapter I revisited the metaphysics of consonance, in view of the consequences arising from differences in tuning and temperament for the experience of tonal harmony.

From the discussion presented in the preceding chapters, two brief observations seem appropriate by way of conclusion. First, my overall proposal has been structured in response to three main problems that I have identified as particularly important for an account of the metaphysics of tonal harmony. These concern the nature of sound, the perception of consonance/dissonance as well as issues of tuning and temperament. Second, in order to address these problems, my overall approach has been one of reconciliation, wherever this has been possible and desirable. Contrary to some entrenched views on the ontology of sound and on the nature of consonance and dissonance, this work is fundamentally an attempt at integration which, one hopes, will prove to be a useful contribution to the philosophy of music.

The aim of this PhD thesis has been to provide a substantial treatment of the metaphysics of tonal harmony – which, as far as my survey of the literature goes, has not been attempted as such. Yet, given the broad enquiry that an account of the ontology and aesthetics of chords inevitably involves, some areas have remained unexplored as a result. Two such areas, in particular, would benefit from further research, which the limited scope of this doctoral thesis could not comprise. They concern the implications of my proposal for other traditions within Western music (e.g. jazz harmony), as well as the consequences of the divergences in tuning and temperament for assessments of tokens of triads – across the two modes – and tetrads. These may well prove to be fertile ground for illuminating the metaphysics of tonal harmony beyond what has been considered here.

References

Abraham, G., *The Concise Oxford History of Music*, Oxford: OUP, 1979.

Audi, P., 'Grounding: Toward a Theory of the "In-virtue-of" Relation', *The Journal of Philosophy*, 109, 12, 2012, pp. 685-711.

Barbancho, A. M., Barbancho, I., Tardón, L. J., Molina, E., *Database of Piano Chords: an Engineering View of Harmony*, Springer, 2013, pp. 1-50.

Barbour, J. M., *Tuning and Temperament: a Historical Survey*, Mineola, NY: Dover, 2004.

Barnett, S. C., *What Sort of Entities Does Grounding Relate?* (PhD thesis), Manchester: University of Manchester, 2018.

Beament, J., *How We Hear Music*, Rochester, NY: Boydell Press, 2001.

Benade, A. H., *Fundamentals of Musical Acoustics*, New York: Dover, 1976.

Berkeley, G., *Three Dialogues between Hylas and Philonous*, New York: Bobbs-Merrill, 1954.

Bidelman, G. M., Krishnan, A., 'Neural Correlates of Consonance, Dissonance, and the Hierarchy of Musical Pitch in the Human Brainstem', *Journal of Neuroscience*, 29, 42, 2009, pp. 13165-13171.

Boethius, A. M. S., *Fundamentals of Music*, New Haven/London: Yale University Press, 1989.

Bowling, D., Purves, D., 'A Biological Rationale for Musical Consonance', *Proceedings of the National Academy of Sciences*, 112, 2015, pp. 11155-11160.

Bowling, D., Purves, D., Gill, K., 'Vocal Similarity Predicts the Relative Attraction of Musical Chords', *Proceedings of the National Academy of Sciences of the United States of America*, 115, 2018, pp. 216-221.

Bregman, A. S., 'Progress in Understanding Auditory Scene Analysis', *Music Perception: an Interdisciplinary Journal*, 33, 1, 2015, pp. 12-19.

Casati, R., Di Bona, E., Dokic, J., 'The Ockhamisation of the Event Sources of Sound', *Analysis*, 73, 3, 2013, pp. 462-466.

Casati, R., Dokic, J., *La Philosophie du Son*, Nîmes: Chambon, 1994.

_________., *La Philosophie du Son*, Nîmes: Chambon, 1994, English translation, 2009, URL = https://jeannicod.ccsd.cnrs.fr/ijn_00420039.

Clark, M., Liggins, D., 'Recent Work on Grounding', *Analysis*, 72, 4, 2012, pp. 812-823.

Cohen, J., 'Sounds and Temporality', In Zimmerman, D. W. (ed.), *Oxford Studies in Metaphysics*, Vol. 5, Oxford: OUP, 2010, pp. 303-319.

Colman, A. M. 'Ohm's acoustic law', *The Oxford Dictionary of Psychology*, 2015, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199657681.001.0001/acref-9780199657681-e-5725.

Cousineau, M., McDermott, J., Peretz, I., 'The Basis of Musical Consonance as Revealed by Congenital Amusia', *Proceedings of the National Academy of Sciences of the United States of America*, 109, 2012, pp. 19858-19863.

De Clercq, R., 'Aesthetic Pleasure Explained', *The Journal of Aesthetics and Art Criticism*, 77, 2, 2019, pp. 121-132.

Dodd, J., *Works of Music: an Essay in Ontology*, Oxford: OUP, 2007.

Dokic, J., 'Two Ontologies of Sound', *The Monist*, 90, 3, 2007, pp. 391-402.

Duffin, R. W., *How Equal Temperament Ruined Harmony*, New York: W. W. Norton, 2007.

Eagle, A., 'Location and Perdurance', In Zimmerman, D. W. (ed.), *Oxford Studies in Metaphysics*, Vol. 5, Oxford: OUP, 2010, pp. 53-94.

Fallows, P., 'Chord', In Latham, A. (ed.), *The Oxford Companion to Music*, OUP, 2011, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199579037.001.0001/acref-9780199579037-e-1376.

Gann, K., *The Arithmetic of Listening*, Urbana: University of Illinois Press, 2019.

Gelfand, S. A., *Hearing: An Introduction to Psychological and Physiological Acoustics*, New York: Marcel Dekker, 1981.

Giordano, N. J., *Physics of the Piano*, Oxford: OUP, 2010.

Hamilton, A., *Aesthetics & Music*, London: Continuum, 2007.

Hamlyn, D. W., 'Gestalt theory', In Honderich, T. (ed.), *The Oxford Companion to Philosophy*, Oxford: OUP, 2005, p. 338.

Handel, S., 'Timbre Perception and Auditory Object Identification', In Moore, B. C. (ed.), *Hearing*, San Diego, CA: Academic Press, 1995, pp. 425-461.

Harrison, D., *Harmonic Function in Chromatic Music*, Chicago: University of Chicago Press, 1994.

Harrison, P. M. C., Pearce, M. T., 'Simultaneous Consonance in Music Perception and Composition', *Psychological Review*, 127, 2, 2020, pp. 216-244.

Haukioja, J., 'Notions of Response-dependence', In Hoeltje, M., Schnieder, B., Steinberg, A. (eds.), *Varieties of Dependence*, Munich: Philosophia, 2013, pp. 167-190.

Hawley, K., *How Things Persist*, Oxford: Clarendon Press, 2001.

Helmholtz, H., *On the Sensations of Tone*, New York: Dover, 1954.

Hutchinson, W., Knopoff, L., 'The Acoustic Component of Western Consonance', *Journal of New Music Research*, 7, 1978, pp. 1-29.

Isacoff, S., *Temperament*, London: Faber and Faber, 2002.

Johnson-Laird, P., Kang, O., Leong, Y., 'On Musical Dissonance', *Music Perception*, 30, 2012, pp. 19-35.

Jones, O. R., 'Supervenience', In Honderich, T. (ed.), *The Oxford Companion to Philosophy*, Oxford: OUP, 2005, p. 903.

Kameoka, A., Kuriyagawa, M., 'Consonance Theory Part I: Consonance of Dyads', *The Journal of the Acoustical Society of America*, 45, 1969, pp. 1451-1459.

Kane, B., 'The Elusive "Elementary Atom of Music"', *Qui Parle*, 14, 2, 2004, pp. 117-143.

Karp, T., *Dictionary of Music*, New York: Dell Publishing, 1973.

Kennedy, M., Kennedy, J. B., 'Concord (consonance)', *The Concise Oxford Dictionary of Music*, 2013, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199203833.001.0001/acref-9780199203833-e-2087.

__________., 'Discord', *The Concise Oxford Dictionary of Music*, 2013, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199203833.001.0001/acref-9780199203833-e-2661#.

Kennedy, J., Kennedy, M., Rutherford-Johnson, T. (eds.), 'Chord', *The Oxford Dictionary of Music*, 2013, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199578108.001.0001/acref-9780199578108-e-1857.

__________., 'Counterpoint', *The Oxford Dictionary of Music*, 2013, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199578108.001.0001/acref-9780199578108-e-2218.

Kostka, S., Payne, D., Almén, B., *Tonal Harmony*, New York: McGraw-Hill, 2018.

Kulvicki, J., 'The Nature of Noise', *Philosophers' Imprint*, 8, 11, 2008, pp. 1-16.

__________., 'Sound stimulants: Defending the Stable Disposition View', In Stokes, D., Biggs, S. (eds.), *Perception and Its Modalities*, Oxford: OUP, 2014, pp. 205-221.

Lahdelma, I., Eerola, T., 'Mild Dissonance Preferred Over Consonance in Single Chord Perception', *I-Perception*, 2016, pp. 1-21.

Latham, A., 'Arpeggio', *The Oxford Companion to Music*, 2011, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199579037.001.0001/acref-9780199579037-e-407.

Leddington, J., 'What We Hear', In Brown, R. (ed.), *Consciousness Inside and Out*, Dordrecht: Springer, 2013, pp. 321-334.

__________., 'Sounds fully simplified', *Analysis*, 79, 4, 2019, pp. 621-629.

Levinson, J., *Music, Art, and Metaphysics*, Ithaca: Cornell University Press, 1990.

Liggins, D., 'Grounding and the Indispensability Argument', *Synthese*, 193, 2016, pp. 531-548.

Lindley, M., *Lutes, Viols and Temperaments*, Cambridge: CUP, 1984.

Locke, J., *An Essay Concerning Humane Understanding*, London: Printed for Thomas Dring and Samuel Manship, 1694.

Lowe, E. J., 'Mereology', In Honderich, T. (ed.), *The Oxford Companion to Philosophy*, Oxford: OUP, 2005, pp. 587-588.

Marin, M., Thompson, W. F., Gingras, B., Stewart, L., 'Affective Evaluation of Simultaneous Tone Combinations in Congenital Amusia', *Neuropsychologia*, 78, 2015, pp. 207-220.

Max, I., 'A Molecular Logic of Chords and Their Internal Harmony', *Logica Universalis*, 12, 2018, pp. 239-269.

Maxwell, J., *An Essay upon Tone*, Edinburgh: Macfarquar and Elliot, 1781.

McDermott, J., Lehr, A., Oxenham, A., 'Individual Differences Reveal the Basis of Consonance', *Current Biology*, 20, 2010, pp. 1035-1041.

Nudds, M., 'Experiencing the Production of Sounds', *European Journal of Philosophy*, 9, 2, 2001, pp. 210-229.

_________., 'Sounds and Space', In Nudds, M., O'Callaghan, C. (eds.), *Sounds and Perception*, Oxford: OUP, 2009, pp. 69-96.

_________., 'What are Auditory Objects?' *Review of Philosophy and Psychology*, 1, 1, 2010, pp. 105-122.

O'Callaghan, C., *Sounds: a Philosophical Theory*, Oxford: OUP, 2007.

_________., 'Sounds and Events', In Nudds, M., O'Callaghan, C. (eds.), *Sounds and Perception*, Oxford: OUP, 2009, pp. 26-49.

_________., 'Hearing Properties, Effects or Parts?', *Proceedings of the Aristotelian Society*, New Series, 111, 2011, pp. 375-405.

_________., *Beyond Vision*, Oxford: OUP, 2017.

Oxenham, A. J., 'The Perception of Musical Tones', in Deutsch, D. (ed.), *The Psychology of Music*, San Diego, London: Academic Press, 2012, pp. 1-33.

Olson, H. F., *Music, Physics and Engineering*, New York: Dover, 1967.

Parker, B., *Good Vibrations: The Physics of Music*, Baltimore: John Hopkins University Press, 2009.

Parncutt, R., *Harmony: a Psychoacoustical Approach*, Berlin: Springer-Verlag, 1989.

Pasnau, R., 'What is Sound?', *The Philosophical Quarterly*, 49, 196, 1999, pp. 309-324.

__________., 'Sensible Qualities: The Case of Sound', *Journal of the History of Philosophy*, Vol. 38, 1, 2000, pp. 27-40.

__________., 'A Theory of Secondary Qualities', *Philosophy and Phenomenological Research*, 73, 3, 2006, pp. 568-591.

Plomp, R., Levelt, W. J. M, 'Tonal Consonance and Critical Bandwidth', *The Journal of the Acoustical Society of America*, 38, 1965, pp. 548-560.

Rameau, J-P., *A Treatise of Music, containing the Principles of Composition*, Dublin: Printed for Luke White, 1779.

__________., *Treatise on Harmony*, New York: Dover, 1971.

Rehding, A., 'Acoustics', In Kelly, M. (ed.), *Encyclopedia of Aesthetics*, OUP, 2014, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199747108.001.0001/acref-9780199747108-e-6.

__________., 'Consonance and Dissonance', In Rehding, A., Rings, S. (eds.), *The Oxford Handbook of Critical Concepts in Music Theory*, New York: OUP, 2018, pp. 437-466, URL = https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780190454746.001.0001/oxfordhb-9780190454746-e-14.

Ridley, A., *The Philosophy of Music: Theme and Variations*, Edinburgh: Edinburgh University Press, 2004.

Rigden, J., *Physics and the Sound of Music*, New York: John Wiley & Sons, 1977.

Roberts, P., 'Turning up the Volume on the Property View of Sound', *Inquiry*, 60, 4, 2017, pp. 337-357.

Salmon, T., *A Proposal to Perform Musick in Perfect and Mathematical Proportions*, London: Printed for John Lawrence, 1688.

Scruton, R., *The Aesthetics of Music*, Oxford: OUP, 1997.

__________., *Understanding Music*, London: Continuum/Bloomsbury, 2009.

Sheppard, A., *Aesthetics*, Oxford: OUP, 1987.

Sibley, F., 'Aesthetic and Non-aesthetic', In Benson, J., Redfern, H. B., Roxbee Cox, J. (eds.), *Approach to Aesthetics*, Oxford: OUP, 2001, pp. 33-51.

__________., 'Aesthetic Concepts', In Benson, J., Redfern, H. B., Roxbee Cox, J. (eds.), *Approach to Aesthetics*, Oxford: OUP, 2001, pp. 1-23.

Sider, T., *Four-Dimensionalism: An Ontology of Persistence and Time*, Oxford: Clarendon Press, 2001.

Simons, P., 'Real Wholes, Real Parts: Mereology without Algebra', *The Journal of Philosophy*, 103, 12, 2006, pp. 597-613.

Solomon, J., *Ptolemy Harmonics: Translation and Commentary*, Leiden: Brill, 1999.

Sorensen, R., 'Hearing Silence: The Perception and Introspection of Absences', In Nudds, M., O'Callaghan, C. (eds.), *Sounds and Perception*, Oxford: OUP, 2009, pp. 126-145.

Speaks, C. E., *Introduction to Sound: Acoustics for the Hearing and Speech Sciences*, San Diego, CA: Singular Publishing, 1999.

Stolzenburg, F., 'Harmony Perception by Periodicity Detection', *Journal of Mathematics and Music*, 9, 2015, pp. 215-238.

Tan, S-L., Pfordresher, P., Harre, R., *Psychology of Music: from Sound to Significance*, New York: Psychology Press, 2010.

Tenney, J., *A History of 'Consonance' and 'Dissonance'*, New York: Excelsior Music Publishing, 1988.

Terhardt, E., 'Pitch, Consonance and Harmony', *The Journal of the Acoustical Society of America*, 55, 1974, pp. 1061-1069.

__________., 'The Concept of Musical Consonance: a Link between Music and Psychoacoustics, *Music Perception*, 1, 3, 1984, pp. 276-295.

The Oxford World Encyclopedia, Philip's, 2014, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199546091.001.0001/acref-9780199546091-e-2402.

Thompson, W. F., 'Intervals and Scales', In Deutsch, D. (ed.), *The Psychology of Music*, San Diego, London: Academic Press, 2012, pp. 107-140.

Trogdon, K., 'An Introduction to Grounding', In Hoeltje, M., Schnieder, B., Steinberg, A. (eds.), *Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, Response-dependence*, Munich: Philosophia, 2013, pp. 97-122.

Tymoczko, D., *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice*, Oxford: OUP, 2011.

Vassilakis, P., *Perceptual and Physical Properties of Amplitude Fluctuation and Their Musical Significance* (PhD Thesis), Los Angeles: UCLA, 2001.

White, G. D., Louie, G. J., *The Audio Dictionary*, Seattle: University of Washington Press, 2005, pp. 284-285.

White, H. E., *Physics and Music: the Science of Musical Sound*, Philadelphia: Saunders College, 1980.

Whittall, A., 'Consonance and Dissonance', In Latham, A. (ed.), *The Oxford Companion to Music*, OUP, 2011, URL = https://www.oxfordreference.com/view/10.1093/acref/9780199579037.001.0001/acref-9780199579037-e-1581.

Young, N., 'Hearing Objects and Events', *Philos*, 175, 2018, pp. 2931-2950.

Young, J. O., 'Key, Temperament and Musical Expression', *The Journal of Aesthetics and Art Criticism*, 49, 3, 1991, pp. 235-242.