CHAPTER 1

The evolution of ancient Greek musical notation

By the middle of the third century BC, from which the first preserved documents of ancient Greek written music date, musical notation was already firmly established; it had acquired much of the inner structure that emerges from the full account given in Alypius’ handbook, compiled perhaps half a millennium later. Thus we are not in a position to directly observe the evolution of this system out of more primitive precursors; its origins are the object of speculation. Although some work has been done on this subject, there are several details for which no adequate explanation has been proposed so far. In the following chapter a new theory of the original conception and early evolution of ancient notation will be derived mainly from internal structural evidence. In accord with the nature of such an approach, this initial argument evolves on rather abstract lines. This might seem hardly appropriate for a musical subject, but it allows the development of a consistent view from a very limited body of evidence. It will be left to the later chapters to embed the conclusions, sometimes with modifications, into a broader, more practical and historical, picture.

THE NOTATION

Greek notation was based on letters or letter-like signs, each one designating a certain functional position within a network of musical scales. It is one of the more complex aspects of the system that this functional position cannot be determined unambiguously from any sign in question, but must be derived by reference to the musical context, i.e. from the general tonality of a given piece. Moreover, the exact pitch of a sign depended not only on

1 For the contexts in which notation was used, cf. e.g. Pöhlmann 1976; 1986; 2005; Prauscello 2006.
The evolution of ancient Greek musical notation

positioning the scalar network within the frequency continuum (for instance by means of a ‘concert pitch’ as in the tuning of modern instruments and orchestras),² but in many cases also on the ‘genus’ and fine tuning, for which the scores provide no information.

The signs of the system come in two complete sets, which are associated with vocal and instrumental music respectively, although the extant scores do not maintain the distinction throughout.³ But the instrumental notation is obviously the older one, so it is very likely that it was originally used for vocal melodies also.⁴ Both sets employ letter-like signs. But while the vocal notation consists simply of the letters of the Ionian alphabet in their canonical order, duplicated in slightly modified form outside the central region of the system, the identification of the instrumental signs has raised difficulties. Speculations about an origin in a Semitic alphabet have been rightly rejected.⁵ Certain Greek local scripts seemed more promising; but although many identifications of signs with letters can be made plausible, others remain problematic. Above all, no meaningful series emerges. Neither are the supposed letters arranged in alphabetic order, nor does their order make any musical sense (by expressing intervallic relations, for instance), nor can they be accounted for as abbreviations for degrees of the scale,⁶ nor as meaningful numbers. So the series of instrumental note signs remains a riddle yet to be solved.

In the developed state in which we know it, the system combines two ideas: that of the regular model scale (sýstêma), and that of keys (tónoi or trópoi), which merge into a comprehensive description of the tonal space of Greek music. The model scale is an abstract set of notes defined within a

² For arguments for the system being more or less fixed in pitch see AGM: 273–6; the topic is discussed in more detail below, pp. 68 ff.
³ Cf. e.g. Aristid. Quint. 1.11, p. 23.18–22; Gaud. 21, p. 150.9–11. The instrumentalist Limenios used ‘instrumental’ notation for his paean (DAGM Nº 21). Barker (1995: 48–9) argues that one major motive for developing an alternative vocal notation was probably the need to distinguish vocal and instrumental ‘parts’ within one score.
⁴ Cf. AGM: 261. — Throughout this book I use ‘notation’ without article for the practice of notating music and its appearance in documents; where I talk about ‘the notation’, a specific system (generally the ancient one) is meant.
⁶ Identification with letters from the Argive local script was favoured by West 1992a: 38–41; AGM: 260–3. When it comes to interpreting the series, however, West admits that “likely their meanings will remain forever hidden” (1992a: 41). The Indic and Western medieval heptatonic note names cited by him can hardly serve as parallels: can we assume that in the fifth century there existed some nomenclature for no fewer than a dozen degrees of the scale (involving, by lucky coincidence, no two or three terms with the same initial letter), which left no single trace in later treatises? On top of this, the series of notational signs did not even denote a ‘scale’ in a practical sense, as will become clear below.
The notation

skeleton of fixed intervals, against which the note material of any actual musical piece can be matched. It comes in the form of the so-called Greater Perfect System, which comprises two octaves, with the ‘middle’ note, mésé, at the centre. Keys, on the other hand, regulate the pitch distances between single instantiations of that scale. Changing from one key to another involves modulation. As soon as an entire set of keys comes to be regarded as more or less fixed within the frequency continuum, they can also be used for transposing a melody to a different pitch range. Although both ideas, that of the model scale and that of the key, seem inseparably entwined in the notational system, they had evolved quite independently of each other. The arrangement of tônoi originated in the practice of modulation, whereas the Perfect System was probably conceived somewhat later in the context of aulos making. Subsequently the regular scales could readily be imposed on an already established system of pitch relations; this was apparently not done before Aristoxenus. The work of this outstanding theoretician is, no doubt, the major landmark on the way to the fully regularised scheme, although his own ‘multi-key diagram’ (diagramma polýtropon) did not yet display the entire Perfect System for each key.

The notation in its evolved form relies on the Aristoxenian system in many respects; Aristoxenus himself, however, rather despised musical notation as contributing nothing to the understanding of the art. From his words it becomes clear that not all of his colleagues thought in similar ways; and we will see that the architects of the notation were always at the forefront of the musical science of their time. Still we must bear in mind that the whole process of finding a proper definition of the tonal material of ancient Greek music was perfectly possible without resort to notation. It was not until late antiquity that the note signs found their way into handbooks on music of Aristoxenian hue; writers with philosophical pretensions such as Ptolemy would not use note signs even then. And indeed in many cases the unequivocal note names were preferable over the signs, ambiguous as these were in respect both to pitch and to musical function.

8 Hagel 2005a. Aristoxenus could already refer to the tetrachords of the Perfect System as recognised entities (Harm. 2.40, p. 50.4–7; cf. also ps.-Plut., Mus. 1117d). For the aulos’ significant role in, and the story of its rejection from music theoretical discourse, cf. Wilson 1999 (with emphasis on the paradoxes associated with the instruments’ new negative image in mid-fifth century Athens); Wallace 2003. There were cities more fond of this instrument than Athens; but even there, according to Aristotle, “practically the majority of the free men” embraced its art in the earlier fifth century (Pol. 1341a).
9 Cf. Hagel 2000: 183–8. The diagramma polýtropon is mentioned in Adrastus ap. Theon, Util. math. 6.4.1–4; Procl., in Tim. 35b. 2.170.7–12; Vitruv. 5.5.6; cf. also Plut., De cohob. ira 455d.
10 Aristox., Harm. 2.39–41, p. 49–51.
The evolution of ancient Greek musical notation

There are twelve semitones to the octave: accordingly, twelve scales arranged at semitone steps seem sufficient to account for all possible notes and tonal relations. Aristoxenus, however, devised a system of thirteen tónoi, so that the highest replicated the lowest at the octave. Thus the completeness of the cycle became immediately visible from the diagram; at any rate the octave must be associated with thirteen notes separated by semitones in exactly the same way as it is with eight notes of a heptatonic scale (whence its English name). Moreover, Aristoxenian theory defined a kind of ‘modulation to a scale an octave apart’, which could be exemplified only if two such scales existed in the diagram. Such a modulation seems a nonsensical conception for those accustomed to treating notes an octave apart as functionally completely interchangeable; but this is not the way the Greeks felt about it.

In the final stage of the notation, another two keys were added to Aristoxenus’ thirteen, so that there were now not one but three pairs of scales that merely extended each other to a total range of three octaves. Perhaps this expansion was caused by musical needs, but more likely it was conceived out of purely aesthetic motives. The resulting fifteen keys were renamed to form five triads, each of which associated a basic scale with neighbouring ‘Hypo-’ and ‘Hyper-’ scales one fourth below and above respectively. The relation of ‘Hyper-’ keys was apparently invented in analogy to the ‘Hypo-’ scales, which looked back on a respectable history: ‘Hypophrygian’ and ‘Hypodorian’ were already parts of pre-Aristoxenian systems, while ‘Hypolydian’ is Aristoxenian at the latest. Three triads retained the old designations as ‘Dorian’, ‘Phrygian’ and ‘Lydian’; for the remaining two, which had no comparable roots in traditional musical practice, names had to be invented. To supplement the set of ethnic designations, the old names ‘Iastian/Ionian’ and ‘Aeolian’ were adopted, which had once stood for musical styles now forgotten.

11 Cleonid. 13, p. 205.10–11.
12 For note signs in practical use, but not part of the tónos system, cf. Aristid. Quint. 1.11, p. 24–7 together with DAGM № 41 (cf. below, pp. 300ff.); the notation could be expanded independently of the scale system.
13 The motivation is expressed by Aristid. Quint. 1.10, p. 21.1–4: ...δῖως γ’ ἀν ἕκαστος βορύτητα τε ἔχει καὶ μεσότητα καὶ ὀξύτητα “that each participates in low, central, and high pitch”.
14 Aristox., Harm. 2.37, p. 47.1–13.
16 Cf. AGM: 231. Here the antibarbarian construction of Heraclides Ponticus probably played a role – he had even reserved the term harmonía for the Greek modes, Dorian, Ionian, and Aeolian (Ath. 624C, reflected in Pollux 4.63; cf. below, p. 61 n. 22 and pp. 430ff.): with the fifteen-keys system, a Greek majority was restored. Throughout this book I use the form “Iastian” rather than “Ionian”; in the sources, both are used indifferently for the respective tónoi.
The notation

Theorists before and after Aristoxenus contented themselves with fewer keys. Three older systems mentioned by Aristoxenus himself consist of five or six tónoi. Other authors refer to musical styles using seven keys, or even merely the three basic ones of Dorian, Phrygian and Lydian.17 Ptolemy, constrained mainly by the limitations of his ‘Pythagorean’ viewpoint,18 reverted to seven keys, in accordance with the seven diatonic ‘tunings’ or octave species, rejecting even an eighth that other anti-Aristoxenians had admitted. But his objections against the extended system of practical music and Aristoxenian theory do no justice to its motivation and structural foundation: to account for every possible kind of modulation.

In addition to the tónoi-based account, one very common type of modulation was usually described in a different way. Many theorists perceived it not as a change of key at all, but merely as the employment of two different options within the same tónos.19 To describe this relation, the ‘Greater Perfect System’ was combined with a ‘Lesser Perfect System’ into one tonal structure, called the ‘Unmodulating System’ (sýstēma ametábolon, Diagram 1).20 Historically such a combination was purportedly favoured by the existence of two standard lyre tunings that shared their lower range from the lowest note, hypátē, up until the central mész, from which they continued upwards with a ‘disjunctive’ whole tone or a ‘conjoint’ tetrachord respectively, ending with two different nētai.21

Modulation between the two parts of the combined system was so common that it received a name of its own: ‘modulation according to scale’, as opposed to ‘modulations according to key’, i.e. modulations that could not be described without resorting to the combination of two or more such systems, with two or more distinct mészai.22 Such extensive combinations were called ‘modulating systems’, which explains the name ‘Unmodulating System’ for the simple one-mész type – a terminology which at first glance

17 Ps.-Plut., Mus. 1134.ab; Ptol., Harm. 2.6, p. 56.4–6; Bacchius 46, p. 303.3–6; cf. also Aristid. Quint. 1.11, p. 23.1; Ath. 655c; Frag. Cens. 12, p. 74.11–12; Schol. Dion. Thrax, Gramm. Gr. 1.3, p. 476.33.
18 Cf. below, pp. 56f.
19 So Ptol., Harm. 2.6, p. 54.7–11; p. 56.3–17.
20 For the accentuation of the feminine genitives plural (which are often found printed differently), cf. Hdn., Pros. cath. 3.1, p. 426.
21 Cf. Diagram 25 on p. 104 below. The terminology, which assigns the notion of ‘low’ to high pitches and ‘high’ to low pitches, is based on the physical position of the strings on the lyre in tilted playing position (cf. e.g. Baud-Bovy 1978: 164; AGM: 64). This is best illustrated by Plut., Plat. quaest. 1008c, where the analogy to the aulos makes it clear that the notion of ‘topmost and first’ applies to the entity next to the player: τὴν ὑπάτην ὑφή, aîn mèn lúra τὸν ἀνωτάτον καὶ πρῶτον [tō-pov], ἐν δ’ αὐλαι τὸν κάτω καὶ τὸν τελευταίον ἐπέχουσαν “seeing that the hypátē holds the topmost and first position on the lyre, but on the aulos the bottommost and final” (the hole for the lowest note is situated at the remote end of the wind instrument).
The evolution of ancient Greek musical notation

must appear peculiar for a structure that, in our understanding, already contains a modulation.\(^{23}\)

Once the number of keys had been extended to the full circle of fifths, all the relations were describable by modulations between different tónoi. Even so, the traditional ‘conjunct’ scales, now entirely redundant, remained in

\(^{23}\) Ptolemy, acknowledging the synemménón tetrachord as a modulating element, restricts the term systéma amétabolon to the Greater Perfect System. The Division of the Canon seems to preserve a pre-Aristoxenian usage where it designated merely the ‘fixed’ notes of the double octave (Sect. can. 19, p. 163.11–165.2; cf. Barker 2007: 400, and the arguments in Hagel 2005a for the importance of this scale skeleton in fourth-century bc music theory).
The notation

use. In the system of fifteen tônoi, the crucial ‘conjunct’ tetrachords of the five keys with plain names are part of their respective ‘Hyper-’ keys: Hyperlydian contains the ‘conjunct’ scale of Lydian, Hyperphrygian that of Phrygian, etc.

The existence of the rudimentary ‘conjunct’ approach to modulation side by side with an extended system of keys that would supersede it shows that the latter is younger. So it is not unlikely that the ancient notion of music ‘in three tônoi’ already refers to tonal structures with two branches such as the Unmodulating System. If so, the three keys in question would already have incorporated the most important notes of their later ‘Hyper-’ scales within their respective conjunct branches. The same holds true for seven-tônoi music. But here the old Dorian, Phrygian and Lydian were expressly provided with their ‘Hypo-’ counterparts also, so that all the relations of the later triads were already present. This explains the later appearance of ‘Hyper-’ keys as separate entities: thanks to the old conception of a ‘conjunct’ alternative, they had been included implicitly. Only once the complete set of modulating scales had been laid down by Aristoxenus, would it become obvious that part of them was structurally related to the conjunct branch of the old keys. Consequently, the last revision of the nomenclature represented this relation by the invention of names with the prefix ‘Hyper-’.

As a result, the set of tônoi consists of several layers. Some still bear their pre-Aristoxenian names; others seem to have been implicitly present before Aristoxenus, but explicitly added as keys either by him or his successors; some were conceived and baptised by Aristoxenus, and renamed afterwards. Table 1 provides an overview of this evolution. It starts with the traditional three-tônoi music, for which we can already compare the famous nómos trimelês, attributed to the early sixth century, consisting of a Dorian, a Phrygian and a Lydian part. The two systems mentioned by Aristoxenus are distinguished mainly by the harmonic relationship they attribute to the Mixolydian.

---

24 They are not entirely identical, because the Lesser Perfect System, with its succession of three conjunct tetrachords without any intervening disjunctive tone, and therefore without repetition of the notes at the octave, is not compatible with the regular Greater Perfect System. To establish identity, the lowest tetrachord has to be ignored.

25 Aristoxenus had used the same prefix in his ‘Hypermixolydian’, where it indicated a simple pitch relationship: the scale ‘even beyond’ the Low and the High Mixolydian. The resemblance to the later triads is only semantic (cf. the discussion below, pp. 429 ff.).

26 Ps.-Plut., Mus. 1132d; 1134ab.

27 Cf. n. 90 on p. 32 below.
The evolution of ancient Greek musical notation

<table>
<thead>
<tr>
<th>three tónoi</th>
<th>pre-Aristox.</th>
<th>seven tónoi</th>
<th>Aristroenus</th>
<th>post-Aristoxenian</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Low Mixolydian</td>
</tr>
<tr>
<td>5</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Dorian</td>
</tr>
<tr>
<td>4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Hypodorian</td>
</tr>
<tr>
<td>3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Hypermixolydian</td>
</tr>
<tr>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Hypophrygian</td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Lydian</td>
</tr>
<tr>
<td>−</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Hypolydian</td>
</tr>
<tr>
<td>1 #</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>High Mixolydian</td>
</tr>
<tr>
<td>2 #</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Low Phrygian</td>
</tr>
<tr>
<td>3 #</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Low Hypophrygian</td>
</tr>
<tr>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>4 #</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Low Lydian</td>
</tr>
<tr>
<td>5 #</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>Low Hypolydian</td>
</tr>
</tbody>
</table>

Table 1  Concordance of tónoi systems

The canonical seven tónoi seem to have been widely acknowledged by the time of, or not long after, Aristroenus,\(^{18}\) who takes it over and adds six new keys to fill in the extant semitone gaps. Four of them he labelled simply after their higher neighbours; similarly, there were two variants of Mixolydian, each reflecting one of the two older systems. There remained the highest scale, which he called ‘Hypermixolydian’: that ‘exceeding the Mixolydian’. The ultimate revision of the notation brought about two new doublet scales and the triadic terminology.

The table is arranged not according to pitch (as it commonly is in the ancient lists), but according to scalar relations: notes of similar designation in neighbouring scales are always a fourth or a fifth apart.\(^{19}\) For each tónos

---

\(^{18}\) Aristroenus describes the earlier systems not to give a historical overview, but to exemplify the former disagreement about tónoi relations. It is therefore possible that he deliberately omitted the most widely acknowledged account(s).

\(^{19}\) Extant treatises generally give lists and/or diagrams of tónoi ordered by pitch. More in-depth works must have included others that displayed the inherent harmonic relations and consequently fol-
The problem: Dorian marginalised

its conventional modern equivalent key is indicated by the corresponding number of flats and sharps. These correspondences between tónoi and modern keys have nothing to do with pitch or modality, nor are they anyhow inherent in the abstract theoretical scheme of tónoi. They result from structural features of the ancient notation – which are ultimately equivalent to our system of accidentals: there is a natural key, namely Hypolydian, which corresponds to the signs of the notation in much the same way as our natural key corresponds to the letters we use to designate notes, or to the white keys on the piano. Similarly, our accidentals have their counterparts in certain complications in the usage of the ancient note signs.30

THE PROBLEM: DORIAN MARGINALISED

These facts are unequivocal, and they have troubled scholars quite a lot.31 ‘Dorian’ is in many respects central to ancient Greek music: as a lyre tuning, it was probably the first to be learnt by the novice;32 as a mode, it was most highly esteemed by both Plato and Aristotle;33 as an octave species, it

30 These relations were determined independently by Bellermann 1847 and Fortlage 1847. Bellermann (if I understand the principle followed by him at all) mistakenly notates Dorian with seven sharps (43) because it contains pitches only available as reverted forms (see below), which he associates with sharps. But Phrygian and Hypodorian also include reverted forms; the mere fact that the same pitches would have been available as different notes cannot justify a transcription as if these other forms had been used (cf. e.g. Π Ὀ transcribed as a on p. 39 and Beilagen Blatt 1f in “Dorisch” and “Hyperdorisch”, but as b in “Phrygisch” etc.); on the contrary, it makes them stand out even sharper. In any case, the association of ancient keys with modern sharps and flats reflects only one aspect of the former, whereas it obscures the internal relations between the single note signs (cf. Fortlage 1847: 136 n. 1). Moreover, a one-to-one match between ancient and modern notation cannot be achieved anyway: Π Ὀ should be transcribed by b on systematic grounds in the mentioned keys, while one could argue for writing a in chromatic Lydian; as Hypoaeolian paramêsis the latter rendition is scarcely avoidable.

31 Cf. especially Riemann 1902 (followed most prominently by Düring 1934; criticised by Sachs 1925: 192.4: 289 n. 1). The solution Riemann proposes is logically sound as regards a synchronous description of pitches and notational signs, but inconsistent as regards the relation of the bounding notes of the tetrachordal framework to the positions of the respective notes within the triplets of the notation; on top of this, Riemann cannot explain the evolution of the notation of his presumed original Dorian octave (according to his hypothesis, the triplets ΔΕΖ and ΝΔΟ would have been reserved, from the very start, for an expansion to take place only later; the Dorian synêmménon tetrachord, on the other hand, would not have been provided for at all, contrary to what one would expect from a Dorian-centred scheme).


33 Plato, Lach. 188d; 193d; Rep. 399a–c; Aristot., Pol. 1340b; 1342ab.
The evolution of ancient Greek musical notation gave the model for the central octave of the Greater Perfect System, between hypâtê and nêtê, the old limits of the octave harmonia. But in the notation, it is by no means the natural scale, as one should expect, but lies at the outskirts of the diagram, to be transcribed with five flats. Consequently it was suspected that there is something wrong with the notation as we have it; that it underwent a profound change after the classical period.\footnote{Most prominently Henderson 1957: 359–67 (cf. Winnington-Ingram 1958: 244–7).} It is one of the major purposes of this chapter to show that nothing of that kind was the case, but that we can understand the marginalised position of the Dorian tônos without resorting to unfounded speculation.

THE EVIDENCE OF THE DEVELOPED NOTATION

Before we can proceed to investigate the evolution of the notation, we must first inspect the organisation it displays in its evolved state. Its structure is determined by the ancient practice of analysing the tonal material in terms of tetrachords: four-note units spanning a fourth, which could be concatenated either immediately or by means of a ‘disjunctive’ whole tone. The Unmodulating System of Diagram 1, for instance (above, p.6), consists of five tetrachords and two disjunctive tones. Modulations are produced by adding a conjunct tetrachord where otherwise a disjunctive tone would have been, and vice versa. This overall structure defines the cardinal points of every musical system, the so-called ‘fixed’ notes. The relative position of the remaining inner two notes of each tetrachord determines the ‘genus’ (génos) of the scale. An (ascending) sequence of a semitone and two whole tones gives the diatonic genus, which predates Hellenic culture considerably\footnote{Diatonic music is attested in Old Babylonian cuneiform tablets, but probably goes back at least to Sumerian music; cf. e.g. Kilmer 1997; Kilmer 2001. The diatonic is acknowledged as older than the other Greek genera in Aristox., Harm. 1.19, p.24.20–25.4.} and was the only one to survive into Western middle ages (e.g. $e–f–g–a$). Possibly Greek innovations were the other two genera, which are characterised by pykná, ‘crowdings’ of the notes at the lower end of the tetrachord. In the enharmonic, which flourished in the fifth century, we are told that the three notes are separated merely by quartertones (e.g. $e–e^{b}–f–a$). The chromatic, which gained prominence among the composers of the late fifth century together with extensive modulation, used semitones