Maths and Music Theory

Adam Townsend

February 2011

I gave this 45-minute talk to the University College London Undergraduate Maths Colloquium (`http://ucl.sneffel.com`) in February 2011, using slides and some practical demonstration by means of a keyboard and bad singing. These are the notes I compiled to make the slides and so go into considerably more detail than I did in the talk, particularly in the first and second halves.

0 References

I owe a lot of this talk to the work of Thomas M. Fiore at the University of Michigan. This talk then, is a distillation of the following (excellent) material:

- Fiore, *Music and Mathematics*
  `http://www-personal.umd.umich.edu/~tmfiore/1/musictotal.pdf`
- Fiore et al., *Musical Actions of Dihedral Groups*
  `http://www-personal.umd.umich.edu/~tmfiore/1/CransFioreSatyendra.pdf`
- D. Benson, *Music: a Mathematical Offering*
  `http://www.maths.abdn.ac.uk/~bensondj/html/music.pdf`

1 Introduction to music theory

Music theory is a big field within mathematics and lots of different people have taken it in different directions. This is good news for us, because it means I can show you some of these directions, hopefully at least one of which, you will find interesting.

I'll start by re-introducing sound in a very brief mathematical way that I'm sure the physicists will have seen before, but it's worth reminding ourselves of where it comes from. From there, we'll have a wave equation and we'll take a look at which properties of this equation give us different outcomes. (I didn't go very deeply into this in the talk, in order to leave room for what I think is the more interesting stuff to come.)

Then we'll take a different direction entirely, looking at how we can bring in some group theory by modelling our familiar music scale to the cyclic group of order 12. We'll have a look at some initially simple actions on groups which correspond to both classical and popular music.

I chose this topic because I knew fairly little about it and thought it might be something that we can all follow without too much difficulty. It brings together applied and pure mathematics in a fairly satisfying way.

I hasten to add that music theory won't allow you to plug in loads of *na-na-na*'s and get *Hey Jude* out of it. Music is primarily a creative pursuit but mathematicians are pattern-spotters and there are plenty of patterns for us to spot here, so let's get on with it.

2 A little physics

2.1 Introduction

Sound 'happens' when air vibrates. You're familiar with the idea that air is loads of molecules whizzing around at about 1000mph bumping into each other. This bumping happens a lot (about 10 billion times a second) and this combined effect is air pressure. When an object vibrates, then, it causes waves of increased and decreased pressure in the air. The ear picks these up through some clever biology (which I won't explain since we're interested more in the mathematics) and interprets it as sound.

Sounds waves travel as longitudinal waves, that is to say they travel in the same direction as they propagate.

We should get rid of the idea of vibrations having single frequencies: on the whole they consist of more than one frequency and so defining one is difficult. Furthermore, it's possible to interpret sound at a pitch which isn't possible in the waveform at all, instead being the combination of two or more. There's a whole branch here called *psychoacoustics* which deals with this kinda thing. But, having said that, initially we'll use the one-freqency theory.

2.2 Why sine waves?

You're probably familiar with the notion of sound being a bunch of sine waves, but why is this the case? If you'll allow me to indulge in a little A-level physics, I shall remind you:

Consider a (light) vibrating string, anchored at both ends. Placing a heavy bead of mass m in the middle, the string exerts a force F on the bead back towards the equilibrium position with a magnitude (for small displacements, anyway) proportional to the perpendicular distance away from the equilibrium position, x:

$$F = -kx$$

Since we also have Newton's second law,

$$F = ma = m\frac{\mathrm{d}^2x}{\mathrm{d}t^2}$$

We get

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \frac{k}{m}x = 0$$

If you don't mind me renaming $\frac{k}{m}$ as ω^2, then we have

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} + \omega^2 x = 0$$

which has solution

$$x = A\cos\omega t + B\sin\omega t$$

which we can also write (if you remember those trig identities from A-level maths) as

$$x = C\sin(\omega t + \phi).$$

This is the reason that sine waves, and not any other sort of periodically oscillating wave, is the basis for our analysis of harmonics. It governs the sound being produced, and the movement of any point on what's called the *basilar membrane* in the ear, the bit which governs our perception of sound.

We call C the peak amplitude, $\frac{\omega}{2\pi}$ the frequency, and ϕ the phase.

2.3 Harmonics

More vibrational modes are possible than just one. In fact, the middle can stay stationary while the two halves vibrate with opposite phases. In guitars, this is what happens when we hold down half the string. The sound produced is exactly twice the frequency, or a *whole octave* higher. More on that later but what we get mathematically is

$$x = A\cos 2\omega t + B\sin 2\omega t.$$

If we hold the string a third of the way along we get

$$x = \tilde{A}\cos 3\omega t + \tilde{B}\sin 3\omega t$$

which is equivalent to a *whole octave and a perfect fifth.*

Where our university mathematics starts to come in to play is what happens in the general case. Plucked strings normally vibrate with a *mixture* of all the modes described by multiples of the natural frequency with various amplitudes. All this depends on *how* the string is made to vibrate (are we plucking a string or throwing a guitar off a cliff?).

It is not difficult to see that the general equation, then, of a typical point on the string is

$$x = \sum_{n=1}^{\infty} \left(A_n \cos(n\omega t) + B_n \sin(n\omega t) \right).$$

This should send Fourier alarm bells ringing, which is good because that is what we can use to analyse these waves. However, I didn't deal with this in my talk for lack of time (plus I might want to do more on in next year).

When a note on a stringed or wind instrument sounds at a certain pitch, say with frequency ν, sound is essentially periodic with that frequency. Fourier series then tells us that such a sound can be decomposed into the sum of sine waves with various phases, at integer multiples of the frequency ν.

Frequency ν gives us the *fundamental*, and frequencies $m\nu$ give us the mth *harmonic*. Why do notes an octave apart sound good, but notes slightly less make a horrible noise? We've seen already that an interval of one octave means doubling the frequency of vibration: the standard example is the A above middle C being 440 Hz, so the A below middle C is therefore 220 Hz.

When we play one of these notes, not only these frequencies come out but *also* the multiples of this frequency, so for the two A's we get (in Hertz):

440, 880, 1320, 1760, ...

220, 440, 660, 880, ...

But if we play notes of frequencies 220 Hz and 445 Hz, we get (in Hertz):

445, 890, 1335, 1780, ...

220, 440, 660, 880, ...

(these are called partials)

The presence of the 440/445 Hz and the 880/890 Hz etc. gives that unpleasant sensation.

On the other hand though, the niceness (music theorists call this *extreme consonance*) of two notes an octave apart means we perceive the two notes as being 'the same', but higher. Such has it been for ever, and it is engrained into our music. This will lead us shortly into looking at scales and then applying some (simple) group theory.

3 Scales

The perfect fifth (C–G) comes from the ratio $\frac{3}{2}$. The third partial of the lower note coincides with the second parial of the upper note, which gives us a pleasant sound. The Pythagoreans discovered this in the 6th Century BC, by plucking strings of small integer lengths: in fact, this was the first known example of a law of nature ruled by the arithmetic of integers, something which greatly influenced the Pythagoreans.

They concluded that a convincing set of notes could be made just by using the ratio $\frac{3}{2}$, and then by repeated use of the ratio $\frac{2}{1}$ if we need to change the octave of the note. If we take C as the ratio $\frac{1}{1}$, and continue to multiply (and divide) by $\frac{3}{2}$, we get

$\left(\frac{3}{2}\right)^{-5}$	$\left(\frac{3}{2}\right)^{-4}$	$\left(\frac{3}{2}\right)^{-3}$	$\left(\frac{3}{2}\right)^{-2}$	$\left(\frac{3}{2}\right)^{-1}$	$\left(\frac{3}{2}\right)^{0}$	$\left(\frac{3}{2}\right)^{1}$	$\left(\frac{3}{2}\right)^{2}$	$\left(\frac{3}{2}\right)^{3}$	$\left(\frac{3}{2}\right)^{4}$	$\left(\frac{3}{2}\right)^{5}$	$\left(\frac{3}{2}\right)^{6}$
D♭	A♭	E♭	B♭	F	C	G	D	A	E	B	F♯

Does this go round in a circle? Unfortunately not for the Pythagoreans, although it is very close.

$$1.404\ldots = \left(\frac{3}{2}\right)^{-6} \times \left(\frac{2}{1}\right)^{4} \approx \left(\frac{3}{2}\right)^{6} \times \left(\frac{2}{1}\right)^{-3} = 1.423\ldots$$

where we're multiplying by powers of $\frac{2}{1}$ (which remember doesn't change the note, only the octave) to bring it down so that all our notes have frequency between 1 and 2.

So the Pythagoreans would have seen musical intervals as continued subtraction, later forming the basis of Euclid's algorithm for finding the highest common factor of two integers.

4 Pitch as a group

We now look at our current model of pitch, which we can translate into the integers modulo 12 (we'll use the notation $\mathbb{Z}_12$), in order to get some cool properties out of it. We'll use this conversion:

C	C♯	D	E♭	E	F	F♯	G	A♭	A	B♭	B
0	1	2	3	4	5	6	7	8	9	10	11

This translation also has the convenient property of being understood worldwide—different countries have different naming conventions for the notes. As an example, let's see what that makes music look like. 'Twinkle, twinkle, little star' goes from

$$(\mathrm{C, C, G, G, A, A, G; F, F, E, E, D, D, C})$$

to

$$(0, 0, 7, 7, 9, 9, 7; 5, 5, 4, 4, 2, 2, 0)$$

So let us define a few simple but useful functions which we can perform:

Transposition: Let $T_n : \mathbb{Z}_{12} \to \mathbb{Z}_{12}$ be such that $T_n(x) = x + n \mod 12$.

Inversion: Let $I_n : \mathbb{Z}_{12} \to \mathbb{Z}_{12}$ be such that $I_n(x) = -x + n \mod 12$.

Music theorists like to transpose and invert segments of music by applying these functions to each element in the set. Sometimes it works, sometimes it doesn't.

So transposing 'Twinkle, Twinkle' by $n = 7$ (the operation T_7) gives us

$$(7, 7, 2, 2, 4, 4, 2; 0, 0, 11, 11, 9, 9, 7)$$

or

$$(\mathrm{G}, \mathrm{G}, \mathrm{D}, \mathrm{D}, \mathrm{E}, \mathrm{E}, \mathrm{D}; \mathrm{C}, \mathrm{C}, \mathrm{B}, \mathrm{B}, \mathrm{A}, \mathrm{A}, \mathrm{G})$$

which sounds like 'Twinkle, Twinkle', but just higher. This is a *key change*, in fact T_7 is a key change of a *perfect fifth*.

But inverting it about $n = 0$ (the operation I_0) gives us

$$(0, 0, 5, 5, 3, 3, 5; 7, 7, 8, 8, 10, 10, 0)$$

or

$$(\mathrm{C}, \mathrm{C}, \mathrm{F}, \mathrm{F}, \mathrm{E}\flat, \mathrm{E}\flat\mathrm{F}; \mathrm{G}, \mathrm{G}, \mathrm{G}\sharp, \mathrm{G}\sharp, \mathrm{B}\flat, \mathrm{B}\flat, \mathrm{C})$$

which sounds awfully strange.

5 The T/I group

Having looked at some examples of transpositions and inversion, we take a look at how they interact with each other. The collection of transpositions and inversions end up forming a mathematical group, and we call it the T/I group.

Quickly recall that a group G is a set G and a function $* : G \times G \to G$ s.t.

1. $\forall a, b, c \in G, (a * b) * c = a * (b * c)$
2. $\exists e \in G$ s.t. $a * e = a = e * a \forall a \in G$
3. $\forall a \in G \ \exists a^{-1}$ s.t. $a * a^{-1} = e = a^{-1} * a$

And let me (somewhat cavalierly) define a **chord** as a set of three notes (x, y, z). As a friend of mine pointed out, there are many forms of chord: this is specifically a *triad*.

T and I act component-wise, that is to say

$$T(x, y, z) = (T(x), T(y), T(z)) \qquad I(x, y, z) = (I(x), I(y), I(z))$$

We say **major** chords are of the form

$$(x, x + 4, x + 7) = (x, x - 8, x - 5)$$

and **minor** chords are of the form

$$(x, x + 8, x + 5) = (x, x - 4, x - 7)$$

where x is called the *root note*. So if you invert a major chord

$$I_0(x, x + 4, x + 7) = (-x, -x - 4, -x - 7)$$

we get a minor chord with root note $-x$, and vice versa.

The C-major chord is (C,E,G) or $(0, 4, 7)$ using our notation. Let S be the set of transposed and inverted forms of the C-major chord $(0, 4, 7)$:

Prime forms:

C (0,4,7)	D♭ (1,5,8)	D (2,6,9)	E♭ (3,7,10)	E (4,8,11)	F (5,9,0)
F♯ (6,10,1)	G (7,11,2)	A♭ (8,0,3)	A (9,1,4)	B♭ (10,2,5)	B (11,3,6)

Inverted forms:

Fm (0,8,5)	F♯m (1,9,6)	Gm (2,10,7)	G♯m (3,11,8)	Am (4,0,9)	B♭m (5,1,10)
Bm (6,2,11)	Cm (7,3,0)	C♯m (8,4,1)	Dm (9,5,2)	D♯m (10,6,3)	Em (11,7,4)

where we've used the convention that 'm' signifies 'minor' (not a convention undertaken by music theorists, who use little letters instead, but one that matches what you'll find online when looking up guitar tabs). This is the set of the 24 minor and major chords.

Now, let G be the group of the 24 functions $T_n : S \to S$ and $I_n : S \to S$ where $n = 0, 1, 2, \ldots, 11$. Let the operation be function composition, $\circ$.

Note that

$$T_m \circ T_n = T_{m+n};$$

$$T_m \circ I_n = I_{m+n};$$

$$I_m \circ T_n = I_{m-n};$$

$$I_m \circ I_n = T_{m-n}.$$

where n, m are mod 12.

Note that the result of composing transpositions and inversions is itself a transposition or inversion, so we can be convinced that $\circ$ really is an operation on G. We call this group the T/I group.

This group is really important because it allows us to see things in music that we otherwise wouldn't be able to see!

6 The PLR group

The set of functions who input and output chords are the PLR group, dating back to a chap called Riemann (but not *that* Riemann) in the late 19th Century. So this type of theory is called *Neo-Riemannian theory.*

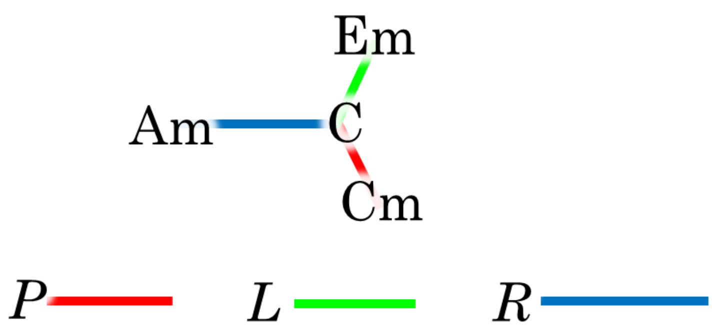
6.1 Some more functions

Let me introduce three functions:

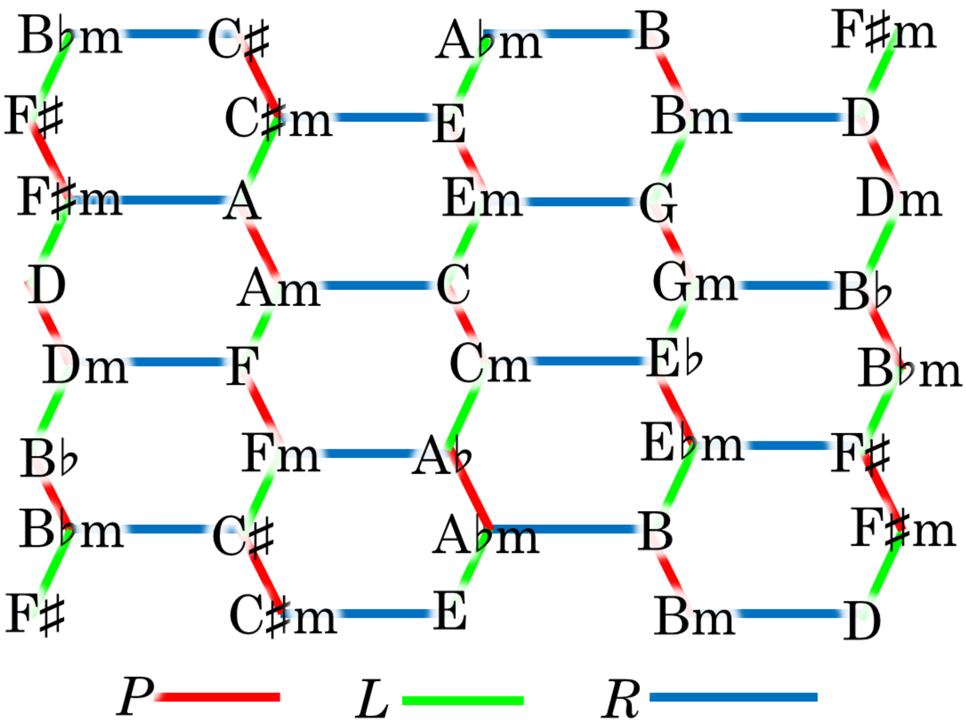
- Let
$$P(x, y, z) = I_{x+z}(x, y, z)$$
$$\text{e.g. } P(0, 4, 7) = (7, 3, 0)$$
which takes C to Cm.
- Let
$$L(x, y, z) = I_{y+z}(x, y, z)$$
$$\text{e.g. } L(0, 4, 7) = (11, 7, 4)$$
which takes C to Em.
- Let
$$R(x, y, z) = I_{x+y}(x, y, z)$$
$$\text{e.g. } R(0, 4, 7) = (4, 0, 9)$$
which takes C to Am.

6.2 Plotting the chords

If we write down C-major in the centre of the page, and draw these three functions and their results as a graph, we get the following:



And if we keep doing this with each chord in turn, we get this big network:



but notice that the top and bottom, and left and right, meet up. This means that we can join these bits of the network up, and what 3D shape is this equivalent to? Mathematicians may call it $\mathbb{S}^1 \times \mathbb{S}^1$ (pronounced *torus*), which looks like a doughnut with a hole in it. Cool, huh?

6.3 Musical interpretation

Note musically this means:

- P takes a chord and maps it to its parallel major or minor
- L is a leading tone exchange, for more theoretical reasons
- R takes a chord to its relative major or minor, e.g. it takes C to Am and Am to C.

6.4 Definition of the group

We shall now define the PLR group as follows:

- The set of all possible compositions of P,L and R.
- With operation function composition $\circ$.

What is $|G|$? Actually only 24, since we get things like $L \circ L(x) = x = R \circ R(x)$.

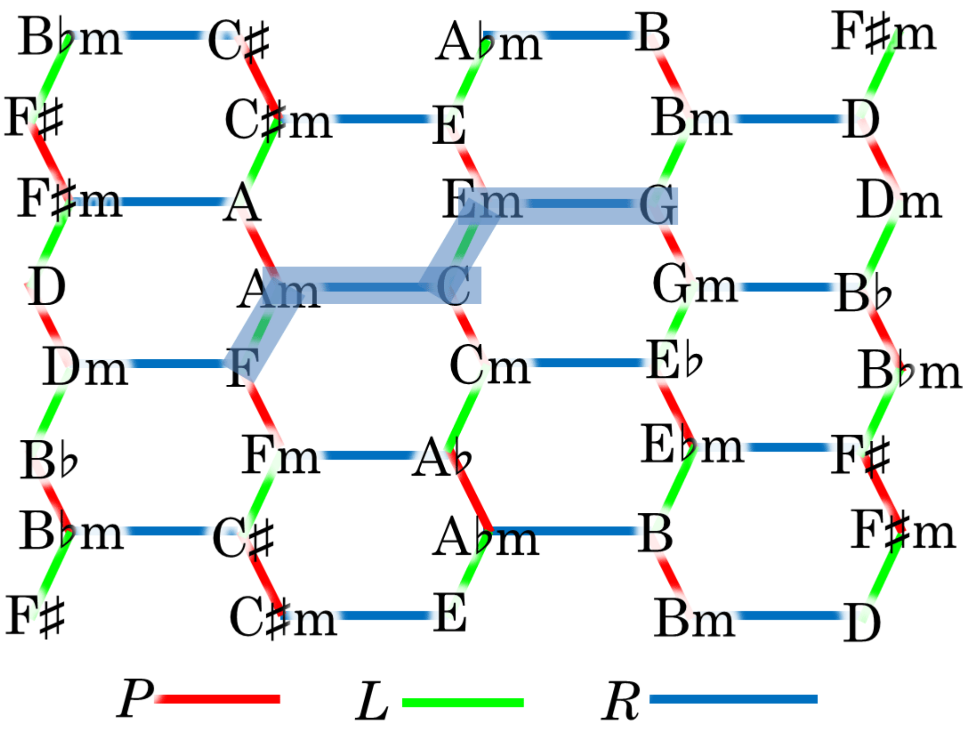
7 Examples

7.1 50's progression (I-vi-IV-V)

In C major: C, Am, F, G (sometimes G7)

$$\bigcirc \xrightarrow{R} \bigcirc \xrightarrow{L} \bigcirc \xrightarrow{R\circ L\circ R\circ L} \bigcirc \xrightarrow{L\circ R} \bigcirc$$

This is a classic chord sequence, whicih you can hear in many old songs, including *Earth Angel, Stand By Me, Wonderful World (Sam Cooke), Grease, Nothing's Gonna Stop Us Now*. What does this look like on the graph?



It's very stable—it oscillates along this small section of the network.

7.2 Pachelbel

In D major: D, A, Bm, F♯m

$$\bigcirc \xrightarrow{R\circ L} \bigcirc \xrightarrow{R\circ L\circ R} \bigcirc \xrightarrow{L\circ R} \bigcirc \xrightarrow{L} \bigcirc$$

Pachelbel's famous *Canon in D* (I'm sure you've heard it, it's all over YouTube) uses this chord sequence over and over again in the bass line. There's a brilliant comedy routine about it by Rob Paravonian which you can see at `http://www.youtube.com/watch?v=JdxkVQy7QLM`. This tracks the same stable section of the network as the 50's progression if we plot it.

7.3 Beethoven's Ninth Symphony

This is a classic example and we'll single out bars 143–176. Check out the sequence of 19 chords in the bass line.

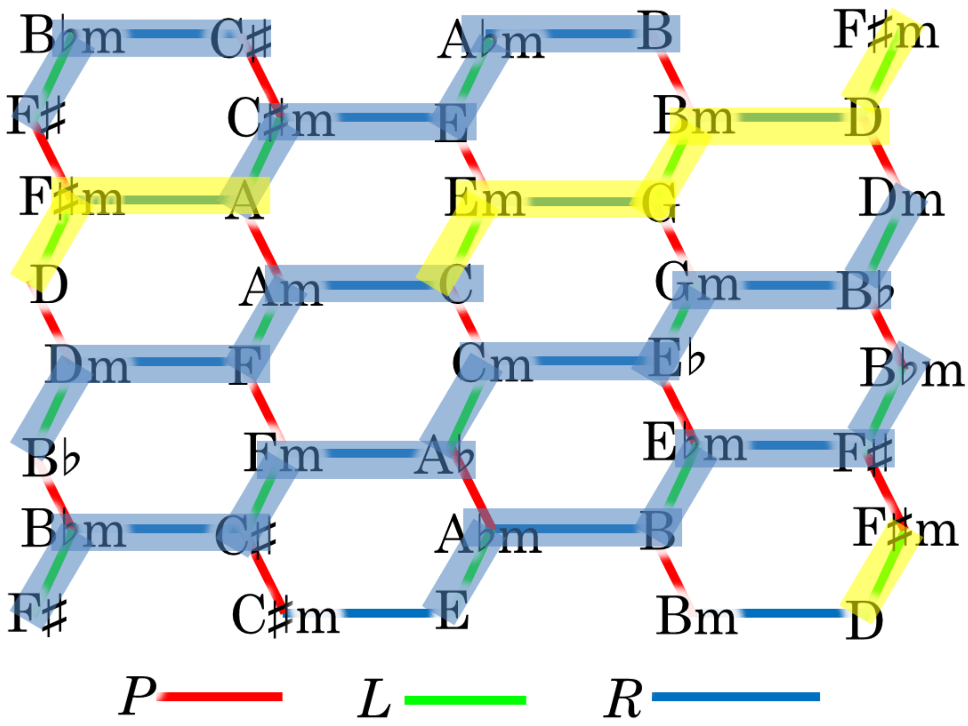
They are

$$\text{C}, \text{Am}, \text{F}, \text{Dm}, \text{B}\flat, \text{Gm}, \text{E}\flat, \text{Cm}, \text{A}\flat, \text{Fm}, \text{D}\flat, \text{B}\flat\text{m}, \text{G}\flat, \text{E}\flat\text{m}, \text{B}, \text{G}\sharp\text{m}, \text{E}, \text{C}\sharp\text{m}, \text{A}$$

Notice that the *whole* sequence can be obtained by applying to C the functions R and L in turn!

$$\text{C} \xrightarrow{\text{R}} \text{Am} \xrightarrow{\text{L}} \text{F} \ldots$$

If we keep going, we hit every single chord. Here's what it looks like if we plot it, where the yellow sections are the bits we've added in by continuing the sequence.



So the path that this traces out on our graph hits every single chord along the torus. Movement in music can be likened very much to movement along the surface of the torus. In fact this graph's automorphism group is D_{24}, something we'll look at now.

8 D_{24}

Now, we're going to show that there's an isomorphism between the PLR group and *the dihedral group of order 24,* D_{24}.

Recall from the Groups and Rings course that D_{24} is the group generated by two elements, s and t such that

$$s^{12} = 1 \qquad t^2 = 1 \qquad tst = s^{-1}$$

So we can think of it as the group of symmetries of a 2-sided dodecagon. Time for some traditional-style maths.

Theorem: The PLR group is generated by L and R and has at least 24 elements.

Proof: First notice that by the definitions of P, L and R,

$$PT_1 = T_1P \qquad LT_1 = T_1L \qquad RT_1 = T_1R$$

Now, if we begin with the C major triad and alternately apply R and L, we get the following sequence, which is the complete list from the Beethoven example:

$$\mathrm{C, Am, F, Dm, B\flat, Gm, E\flat, Cm, A\flat, Fm, D\flat, B\flat m,}$$

$$\mathrm{G\flat, E\flat m, Bm, G\sharp m, E, C\sharp m, A, F\sharp, D, Bm, G, Em, C}$$

which tells us the 24 bijections

$$R, LR, RLR, \ldots, R(LR)^{11}, (LR)^{12} = 1$$

are distinct and that the PLR group has at least 24 elements, and that LR has order 12.

Furthermore, note that

$$R(LR)^3(\mathrm{C}) = \mathrm{Cm}$$

and since $R(LR)^3$ has order 2 and commutes with T_1, $R(LR)^3 = P$ and so the PLR group is generated by L and R alone.

So now, let's call $s = LR$, $t = L$, then $s^{12} = 1, t^2 = 1$ and

$$tst = L(LR)L = RL = s^{-1}$$

as required.

All which remains is to show that the PLR group has order 24. But first, let me give you an interesting result:

9 T/I and PLR are dual

Let me remind you of the definition of a symmetric group:

Definition: The symmetric group $\mathrm{Sym}(S)$ of a set S is the group consisting of all bijections of the set from the set to itself, under function composition.

So we can consider the T/I and PLR groups as subgroups of $\mathrm{Sym}(S)$

$$T/I, PLR \subset \mathrm{Sym}(S)$$

Then something quite cool happens. The centre of the T/I group is the PLR group, and vice-versa!

Recall the definition of the centre of a group from Groups and Rings Exercises 7A:

Definition: The *centre* $Z(G)$ of a group G is the fixed point set under the action of G on itself by conjugation

$$Z(G) = x \in G : \forall g \in G, gxg^{-1} = x$$

So in this sense, we can call the T/I and PLR groups *dual*. In fact, we can illustrate this property. If we apply T_1 to the C major triad, and then L, this is the same as applying L and then T_1:

$$S \xrightarrow{T_1} S$$

$$L \downarrow \qquad \downarrow L$$

$$S \overrightarrow{T_1} S$$

i.e. the diagram *commutes*. In fact we can see the P, L, and R all commute with T_1 and I_0. And since these are the generators of the groups, any diagram with vertical arrows in the PLR group and horizontal arrows in the T/I group will commute.

Bringing back a previous example to show you what this looks like:

- Pachelbel's *Canon in D* (chord sequence D–A–Bm–F♯m)

$$\text{D} \xrightarrow{T_7} \text{A}$$

$$R \downarrow \qquad \downarrow R$$

$$\text{Bm} \overrightarrow{T_7} \text{F}\sharp\text{m}$$

Now to prove it properly!:

Theorem: The PLR and T/I groups are the centres of each other in Sym(S). Furthermore, they are isomorphic to D_{24}.

Proof: We have determined that any element of the PLR group commutes with an element of the T/I group, i.e. $PLR \subset Z(T/I)$.

$\forall y \in S$, claim the fixed point set of y under the action of $Z(T/I)$ contains only the identity element. Suppose $h \in Z(T/I)$ and fixes y, and that g is in the T/I group, then we have:

$$hy = y$$

$$ghy = gy$$

$$hgy = gy$$

So since the T/I group acts transitively, every $y' \in S$ is of the form gy for some g in the T/I group, therefore h is the identity function on S. Hence the fixed point set of y in $Z(T/I)$ is the trivial group.

Now recall the theorem from Groups and Rings to do with the Class Equation when we said

$$|\langle x\rangle| = \frac{|G|}{|G_x|}$$

i.e. the size of the orbit of x is the size of the group G divided by the size of the fixed point set G_x.

Apply this to $G = Z(T/I)$ to give us

$$\frac{|Z(T/I)|}{|Z(T/I)_y|} = |\langle y\rangle| \leq |S| = 24$$

As $PLR \subset Z(T/I)$ and $|Z(T/I)| = 1$, we conclude

$$|PLR\text{-group}| \leq |Z(T/I)| \leq 24$$

From the bit earlier when we listened to Beethoven's *Ninth Symphony*, we know the PLR group has at least 24 elements. Hence (in the same method as Archimedes), we know the PLR group has exactly 24 elements and is equal to $Z(T/I)$.

To show the other way round, just switch the roles of the T/I and PLR groups.

Hence $PLR \cong T/I \cong D_{24}$. $\square$

This last part was mentioned as a cool result in the talk, although time (and interest, really) meant I couldn't go through this as rigorously as I have here.

If you're looking for more on this subject, then I heartily recommending reading the references at the top of this document. If you notice any mistakes (thanks, Reddit!) or wish to contact me, you can at `adam@adamtownsend.com`.