Music Theory

Kyle Sherbert

March 21, 2018

Phrase A

Phrase B

1 Introduction

The building block of a musical composition is normally the chord. Composers stick different chords together in interesting ways to form a musical phrase, and then move on to the next phrase. Of course, *good* composers write each phrase with the whole song in mind, letting themes meet and interact in interesting ways. That is, the building block of a *good* composition is the phrase.

I could write “good” music too, if, instead of writing a whole bunch of notes, I could merely describe how my themes interact for each phrase. For example, my first phrase could be “theme A”, and my second, “theme B”, then, um, “a twiddle of A” and finally, er, “A star-melded with B”. So I’d like to define an algebra of phrases which defines unary operators like “twiddle” ($\sim$) and binary operators like “star-meld” ($*$).

Well-established music theory already provides many sorts of operations. “Melody inversion”, for instance, refers to a reflection of each pitch around some axis pitch. That is, a “3rd up” becomes a “3rd down”, and the series “C-E” becomes “C-A” (if C is your axis). This inversion is equivalent to numbering each pitch from the axis and taking the additive inverse mod 12. This is all well and good, but *multiplicative* inverses in mod 12 are remarkably uninteresting: when they *do* exist, they equal themselves! Perhaps there are other ways of representing phrases as mathematical objects, which allow for a more complete algebra.

First let us define what we mean by a phrase. The phrase P has a certain key, a certain duration, and a bunch of notes. We would like to break P up into t equal-sized “chords”, picking t so that individual notes never start or end in the middle of a chord (even though they may be distributed over several chords). Once you have found the smallest possible value t_0, you know any number $t = n \cdot t_0$ is also valid: it just means breaking each chord up n more times.

For example, a common-time phrase with four measures and nothing but quarter notes has $t_0 = 16$. Syncopate your rhythym halfway through with an eigth rest and $t_0 = 32$. Change a couple quarter notes to a triplet and $t_0 = 96$: the eigth rest needs three chords, and each note in the triplet needs four.

Note that our definition of phrase doesn’t deal with articulation, or tempo variation, or much of anything really that makes music great. These are left entirely to the composer.

We will explore three algebras. For each one, we will define the phrase sum ($A + B$), negation ($-A$), product ($A * B$), and inverse ($\sim A$). We shall get our twiddles and star-melds after all! Each algebra is accompanied by attached .wav files performing each operation on the following two phrases A ( ) and B ( ) from the song *In Dreams*.[*]

[*] Composed by Fran Walsh and Howard Shore for *Lord of the Rings*. Piano arrangement by Dan Coates.

2 Linear Algebra

In this algebra, we represent phrases as matrices over $\mathbb{R}$. In this section, indices start from 1.

Represent the ith chord of the phrase P with a column 12-vector $p_i \in \mathbb{R}^{12}$, so that the jth number in p_i indicates the relative dynamic at which to play the jth pitch in the circle of fifths when starting from the key of P. For example, if your phrase P is in the key of C, the chord "C-E-G-C"is represented by:

$$p^\top = [\ 2\ \ 1\ \ 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ ]$$

C is "twice as loud", because it is played twice in the chord, but note that p preserves no information on each note's octave, or even *whether* a note is played in multiple octaves. The chord "C-E-G" would map to the same p as long as we play the C extra loud.

We define the mapping[†] $P \to \mathbf{M} \in \mathbb{R}^{12\times t}$, where the ith column of $\mathbf{M}$ represents the ith chord of P. We define the reverse mapping $P' \leftarrow \mathbf{M}'$ to refer to any phrase P' which could map to the matrix $\mathbf{M}'$.

We define a negative dynamic to *sound* exactly the same as its absolute value, but of course it adds and multiplies like a negative number. This means $-P$ will *look* identical to P, but produces a different result when subjected to additional operations.

Sum. *Given phrases* P_1 *and* P_2*, take* $t = \mathrm{lcm}(t_{0_1}, t_{0_2})$ *and perform the mappings* $P_1 \to \mathbf{M}_1$*,* $P_2 \to \mathbf{M}_2$*. The sum* $P' = P_1 + P_2$ *is given by* $P' \leftarrow \mathbf{M}_1 + \mathbf{M}_2$*.*

Negation. *Given the phrase* $P \to \mathbf{M}$*, the negation* $P' = -P$ *is given by* $P' \leftarrow -\mathbf{M}$*.*

Product. *Given phrases* P_1 *and* P_2*, take* $t = \mathrm{lcm}(t_{0_1}, t_{0_2})$ *and perform the mappings* $P_1 \to \mathbf{M}_1$*,* $P_2 \to \mathbf{M}_2$*. The product* $P' = P_1 * P_2$ *is given by* $P' \leftarrow \mathbf{M}_1\mathbf{M}_2^\top$*.*

Inverse. *Given the phrase* $P \to \mathbf{M}$*, the inverse* $P' = \sim P$ *is given by* $P' \leftarrow (\mathbf{M}^+)^\top$*, where* $\mathbf{M}^+$ *is the pseudoinverse of* $\mathbf{M}$*.*[‡]

Artistic Freedom The composer is free to select any valid t for hir phrase: operations will generally produce different results for matrices formed from different t. The composer is also free to interpret the octave(s) in which each note should be played, given the reverse mapping $P' \leftarrow \mathbf{M}'$.

This algebra seems fairly intuitive, especially when P_1 and P_2 are the same duration. In this case, taking the least common multiple t sets each chord to the same duration, and addition is equivalent to playing the two phrases simultaneously.

Our definition of multiplication requires transposing $\mathbf{M}_2$ because, first, you can only multiply matrices if their inner dimensions agree, and second, because matrix multiplication mixes rows of the first factor with *columns* of the second factor. Note that $\mathbf{M}_1\mathbf{M}_2^\top$ must have 12 rows and 12 columns, so P' is guaranteed to have $t = 12$. Also note that the commutation $P_2 * P_1$ produces the transpose matrix of $P_1 * P_2$. You're just swapping your pitch and time axes!

[†] The notation $\mathbf{M} \in \mathbb{R}^{12\times t}$ means $\mathbf{M}$ is a matrix of real numbers with 12 rows and t columns.

[‡] As opposed to the true inverse $\mathbf{M}^{-1}$, which exists only if $\mathbf{M}$ is square. You probably know that the inverse $\mathbf{M}^{-1}$ is that matrix which satisfies $\mathbf{M}\mathbf{M}^{-1} = \mathbf{1}$. The pseudoinverse has several similar but weaker criteria, such as $\mathbf{M}\mathbf{M}^+\mathbf{M} = \mathbf{M}$. If $\mathbf{M}^{-1}$ exists, $\mathbf{M}^+ = \mathbf{M}^{-1}$.

$\mathbf{M}^+$ is easily calculated through singular value decomposition. If $\mathbf{M} = \mathbf{L}\boldsymbol{\Delta}\mathbf{R}^\top$ where $\boldsymbol{\Delta}$ is a diagonal square matrix, then $(\mathbf{M}^+)^\top = \mathbf{L}\boldsymbol{\Delta}^+\mathbf{R}^\top$, where $\boldsymbol{\Delta}^+$ is $\boldsymbol{\Delta}$ with each non-zero value replaced by its reciprocal.

3 Galois Algebra

In this algebra, we represent phrases as polynomials over $\mathbb{F}_2^{12}$.[§] In this section, indices start from 0.

Represent the ith chord of the phrase P with the binary 12-vector $p_i \in \mathbb{F}_2^{12}$, so that the jth binary value in p_i indicates whether or not to play the jth pitch in the circle of fifths when starting from the key of P. For example, in the key of C, the chords "C-E-G-C" and "C-E-G" are both represented by:

$$p = [\ 1\ \ 1\ \ 0\ \ 0\ \ 1\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ ]$$

The binary 12-vectors preserve no information on each note's dynamic, octave, or even *whether* a note is played in multiple octaves.

Now select a generator α of $\mathbb{F}_2^{12}$. We define the mapping $P \to F(x)$, where $F(x)$ is a polynomial over $\mathbb{F}_2^{12}$ of degree $t-1$, and $F(\alpha^i) = p_i$. If we define the coefficients c_k so that $F(x) = \sum c_k x^i$, then they can be calculated from the system of equations $p_i = \sum \alpha^{ik} c_k$. The properties of finite fields guarantee that this system is solvable as long as $t < 2^{12}$.

Finally, we must select a "fundamental" polynomial $M(x)$; this polynomial has no direct dependence on any musical phrase and is left to the discretion of the composer. If $M(x)$ is irreducible, then every phrase has an inverse (but that doesn't mean they'll sound good).

Sum. *Given phrases $P_1 \to F_1(x)$, $P_2 \to F_2(x)$, the sum $P' = P_1 + P_2$ is given by $P' \leftarrow F_1(x) + F_2(x)$.*

Negation. *The negation of any phrase P is equal to itself. That is, $P = -P$.*

Product. *Given phrases $P_1 \to F_1(x)$, $P_2 \to F_2(x)$, the product $P' = P_1 * P_2$ is given by $P' \leftarrow F_1(x) \cdot F_2(x)$.*

Inverse. *Given the phrase $P \to F(x)$, the inverse $P' = {\sim}P$ exists iff $\gcd(F, M) = 1$. Let the polynomials $q(x)$ and $r(x)$ be defined by $F = q \cdot M + r$ and $\deg(r) < \deg(M)$. Now let r^{-1} be defined by $r \cdot r^{-1} \equiv 1 \pmod{M}$ and $\deg(r^{-1}) < \deg(M)$. Then P' is given by $P' \leftarrow q \cdot M + r^{-1}$.*

Artistic Freedom All operations depend on the selection of α, and inverses depend on the selection of M. The composer is free to select any valid t for hir phrase: operations will generally produce different results for polynomials formed from different t. The composer is also free to interpret the dynamics and octave(s) in which each note should be played, given the reverse mapping $P' \leftarrow F'(x)$.

The mathematical abstraction in this algebra has very little to do with traditional music theory. Even if P_A and P_B are only slightly different musically, they will behave vastly differently in this algebra. The inherent chaos means most operations produce rather ugly sounds. However, the freedom to select α and M means opportunities are endless: an inverse ${\sim}P$ may sound completely awful in one choice of parameters, but there *exists* a different choice in which ${\sim}P$ is absolutely beautiful. It's up to the composer to find it.

Of course, if you know ahead of time what you want ${\sim}P$ to be, you can select your fundamental to make it happen: if you want $P_1 = {\sim}P_2$, then take $M = F_1 \cdot F_2 + 1$.

Our definition of inverse makes use of the quotient q, rather than just r, simply to avoid the length of ${\sim}P$ being bounded by the selection of M. If this bothers you, you can select an M whose degree is greater than any t you plan to work with, so that $q = 0$.

[§] The notation $\mathbb{F}_2^{12}$ refers to the finite field with 2^{12} elements, each of which is a vector with 12 binary values.

4 Harmonic Algebra

In this section, we treat phrases like they are, well, sound waves. In this section, indices start from 0.

Unlike the other algebras, we will define operations on the chord level. Since our results depend on the actual frequency of each pitch, chords cannot be "stretched": $C_1 + C_2$ and $C_1 * C_2$ exist if and only if C_1 and C_2 have the same duration.

The phrase operations simply apply each chord operation sequentially. This means two phrases P_1 and P_2 must have the same number of chords $t = \mathrm{lcm}(t_{0_1}, t_{0_2})$ to be combined. Taken with the chord duration constraint above, we see $P_1 + P_2$ and $P_1 * P_2$ exist if and only if P_1 and P_2 are the same duration.

Consider a note as a pitch and an octave (ex. F♯5). To every note assign the integers p, which counts chromatically the number of pitches between C and the note's pitch, and o, which indicates the octave. Now define the integer $n = 12o + p$. For example, F♯5 has:

$$n = 12(5) + 5 = 65$$

The pure-tone sound wave of a single note n is given by:

$$\phi_n = \cos[2^{(n-n_A)/12} \cdot \omega_A t]$$

where $n_A = 57$ corresponds to A4, and $\omega_A = 2\pi(440\text{Hz})$ is its angular frequency.

4.1 Pure Harmonic Algebra

We define the mapping $C \to f(t)$ so that $f(t) = \sum A_i \phi_i$, where the number A_i represents the relative dynamic for each note $n = i$. If the note does not appear in the chord, then $A_i = 0$. Negative dynamics are perfectly permissible: $A_i < 0$ simply produces the note π radians out of phase with its absolute value.

If we are capable of playing chords directly from their sound wave functions, defining the operations is simple. We will christen this the "pure harmonic" algebra:

Sum. *Given chords $C_1 \to f_1(t)$, $C_2 \to f_2(t)$, the sum $C' = C_1 + C_2$ is given by $C' \leftarrow f_1(t) + f_2(t)$.*

Negation. *Given the chord $C \to f(t)$, the negation $C' = -C$ is given by $C' \leftarrow -f(t)$.*

Product. *Given chords $C_1 \to f_1(t)$, $C_2 \to f_2(t)$, the product $C' = C_1 * C_2$ is given by $C' \leftarrow f_1(t) \cdot f_2(t)$.*

The obvious definition of inverse would use the reciprocal of $f(t)$, but as it turns out this (usually) produces *silence*. Every node of $f(t)$, (and there are many, for all but the simplest chords) turns into an infinity in $\frac{1}{f(t)}$. We can approximate them with "very big numbers", fine, but these high points are so close together that we are effectively describing a pure tone with *very* high frequency - too high to be heard by the human ear!

To make the result of our inverse operation interesting, we will just reciprocate each note individually, approximating each secant function ϕ_n^{-1} by expanding it into a series of cosines, each a harmonic ϕ_{n+h_i} of the original function:¶

Inverse. *Define the inverse of a note ϕ_n^{-1} to be:*

$$\phi_n^{-1} \equiv 2 \sum (-1)^i \phi_{n+h_i}, \quad \text{where} \quad h_i = 12 \log_2(2i+1)$$

Given the chord $C \to \sum A_i \phi_i$, the inverse $C' = \sim C$ is given by $C' \leftarrow \sum A_i \phi_i^{-1}$

¶Secant functions contain infinities and discontinuities, so it isn't *rigorously* correct to take the Fourier decomposition. But we'll overlook that. Even so, the series never converges: every term is just as "important" as the next, so there's never a good point to stop taking harmonics. Still, after just a few, the decomposition is recognizable as an *attempt* at a secant function, so we'll overlook that too. Take as many terms as you feel compelled to.

Interestingly, our definition of inversion essentially means we play the same chord but with a different *timbre*, like we're playing an instrument! Personally I think it sounds like a bagpipe...

4.2 Chromatic Harmonic Algebra

If we *must* write our phrases out as sheet music, the reverse mapping $C' \leftarrow f'(t)$ is not self-evident. In this case, we would do better to define the mapping $C \rightarrow A_n$, where A_n refers to the whole sequence of relative dynamics A_i described above. We can define our operations in terms of A_n by applying a bit of algebra and trigonometry to the operations defined above.

We must on occasion "round" to the nearest chromatic pitch. This introduces some error as compared with the pure harmonic algegra. Since, for whatever mysterious reason, our sense of beauty is more dictated by factors of 2 than factors of 12, products and inverses in this "chromatic" harmonic algebra are liable to sound horribly out of tune.

Sum. *Given chords $C_1 \rightarrow A_n$, $C_2 \rightarrow B_n$, the sum $C' = C_1 + C_2$ is given by $C' \leftarrow C_n$, where $C_i = A_i + B_i$.*

Negation. *Given the chord $C \rightarrow A_n$, the negation $C' = -C$ is given by $C' \leftarrow C_n$, where $C_i = -A_i$.*

Product. *Define integers $+_{ij}$ and $-_{ij}$ to be*

$$\pm_{ij} \equiv 12\lfloor \log_2 |2^{i/12} \pm 2^{j/12}| \rceil$$

Given chords $C_1 \rightarrow A_n$, $C_2 \rightarrow B_n$, the sum $C' = C_1 + C_2$ is given by $C' \leftarrow C_n$, so that

$$\sum_k C_k \phi_k = \sum_i \sum_j A_i B_j \cdot \frac{1}{2}(\phi_{+_{ij}} + \phi_{-_{ij}})$$

Inverse. *Define the inverse*[‖] *of a note ϕ_n^{-1} to be:*

$$\phi_n^{-1} \equiv 2 \sum (-1)^i \phi_{n+h_i}, \quad \text{where} \quad h_i = 12\lfloor \log_2(2i+1) \rceil$$

Given the chord $C \rightarrow A_n$, the inverse $C' = \sim C$ is given by $C' \leftarrow C_n$, so that $\sum C_k \phi_k = \sum A_i \phi_i^{-1}$.

Artistic Freedom Unlike the other algebras, there are not in fact any parameters to vary. All operations are performed element-wise, so the selection of t is irrelevant. Our mathematical representation *does* preserve information on dynamics and octaves, so the sequence C_n tells you *exactly* what C' should be. You may consider this a good thing or a bad thing.

[‖] I do not recommend taking more than 12 harmonics (ie. 13 terms) in the sum for ϕ_n^{-1}, because after that, $h_i = h_{i+1}$ on occasion, and because of the alternating factor $(-1)^m$, they cancel each other out. They only cancel out entirely if an even number of h_m land on that value, but that is simply an artifact of our rounding.

5 Results

This page contains attachments to .wav files. You *should* be able to click on the icons to open the corresponding sound. If not, you might try a different PDF viewer (perhaps a browser?). Please pardon the poor quality - they were thrown together by concatenating sin waves corresponding to each note in the phrase. The (Java) source code of my implementation of each algebra can be found at github.com/kmsherbert/composition.

The results of each operation are kept as boring as possible. If the octave is not determined by the operation, it was taken to be 4, and if the dynamic is not determined by the operation, it was taken to be uniform. In particular, note that negations in the Linear Algebra sound novel here, but they are actually just the original notes compressed to one octave.

Originals	A	B

Linear Algebra		
Sum:	$A + B$	
Negation:	$-A$	$-B$
Product:	$A * B$	$B * A$
Inverse:	A	B

Galois Algebra		
Sum:	$A + B$	
Negation:	$-A$	$-B$
Product:	$A * B$	
Inverse (Random M):	A	B
Inverse (M so $A =$ B):	A	B

Harmonic Algebra		
Sum:	$A + B$	
Negation:	$-A$	$-B$
Product (Pure):	$A * B$	
Inverse (Pure):	A	B
Product (Chromatic):	$A * B$	
Inverse (Chromatic):	A	B

6 Conclusion

We have successfully formulated three complete systems to "algebraically" manipulate musical phrases. Which one is the best? That's obviously a subjective question, but for the rest of this report, let's compare their advantages and disadvantages.

First, let's ask if each operation makes sense. In the Linear Algebra, we treat phrases as matrices, and they behave like matrices in every operation. We perform some special transpositions when taking products and inverses, but this is merely to keep the matrices in a consistent form. I feel these matrices accurately reflect the underlying geometry of music theory, in a transcendental high-dimensional sort of way. In the Galois Algebra, we treat phrases as polynomials over a finite field, and they behave like polynomials in every operation, provided you accept the use of an arbitrary polynomial M for taking inverses. Unlike the Linear Algebra, however, the polynomials do not appear to reflect the structure inherent in music theory. In fact, finite fields are so useful in cryptology precisely *because* they seem to behave so randomly, ie. they lack structure. In the Harmonic Algebra, we treat phrases as what they really are: pressure waves. These operations, then, are perhaps the *most* sensible... *except*, that is, for the inverse, which we redefined completely (since the more intuitive inverse would almost always produce sounds too high-pitched to hear).

Second, let's examine the constraints on input. In all three algebras, every phrase has a negation, such as it is. In all three algebras, every phrase has an inverse, *except* perhaps in the Galois Algebra with a poorly chosen M. On the other hand, the Galois Algebra has *no* constraints on addition and multiplication. In the Harmonic Algebra, addition and multiplication can *only* be applied to equal-duration phrases. The Linear Algebra *permits* adding and multiplying any two phrases, but I would argue that combining phrases of different duration sacrifices the geometrical structure mentioned in the previous paragraph. The Galois Algebra, never having had that structure in the first place, loses very little by allowing arbitrary operands.

Finally, let's assess the general aesthetic of each algebra: are the results interesting, and do they sound good? Since this is entirely subjective, I'll just provide a table scoring each operation according to my personal aesthetic. The first number rates interestingness and the second rates prettiness. I'm leaving out the Chromatic Harmonics because I think we can all agree they sound awful, as predicted. But perhaps your table will look different!

	$A+B$	$-A$	$A*B$	A
Linear Algebra	2/9	1/10	4/6	9/7
Galois Algebra	7/7	0/10	6/3	1/1
Harmonic Algebra	2/9	1/10	8/7	2/9

Remember, no matter their rating, that these algebras are not intended to stand on their own. If you can really write a good song just by multiplying together a couple phrases in different ways, then good for you. But if you have to tweak things a little (or a lot!), so much the better.