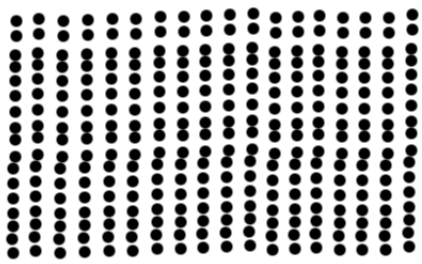
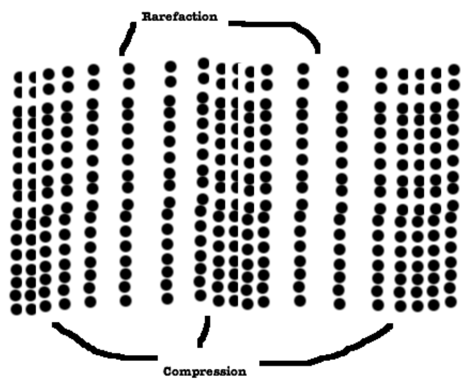
What is Sound?

All material is made of atoms which make up molecules. The phase of the matter (solid, liquid, gas) depends on how tightly the molecules are packed together. In a solid, they are packed so tight that there is not much room to wiggle around. In a gas, however, they are packed loosely enough where they are able to move all around. Sound is created when something moves and is able to vibrate the air molecules. The pictures below show a cartoon of air molecules as well as a vibration moving through air molecules. Notice that the vibration creates *compressions* (where all the air molecules are really close together), and *rarefactions* (where the air molecules are spread apart). You can imagine stretching a slinky and pushing one side forwards and observing the compression of the spring travel along the length of the slinky.





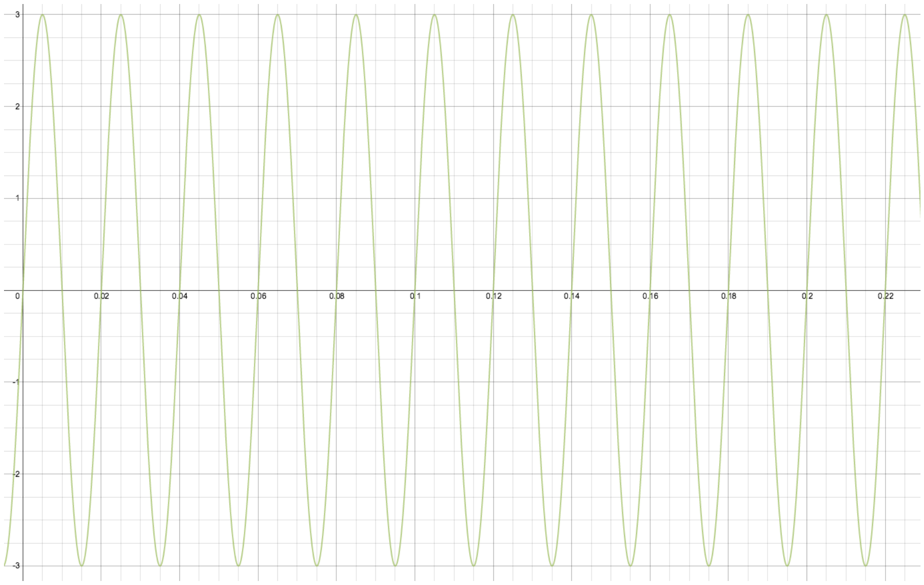
These waves travel through the air and into our ears and vibrate our eardrums. Our ear drums are hooked up to our brains in such a way that our brain can interpret the vibration as sound. Of course, this comes with some limitation. If I wave my hand back and forth in the air, I am creating such vibrations of air molecules, but of course we cannot hear that. A hummingbird, however, flaps its wings really fast (anywhere from 10-80 times per second) and hums–we can hear that. It is correct to guess we cannot hear a waving hand because we cannot wave our hand fast enough.

What is Sound Made of?

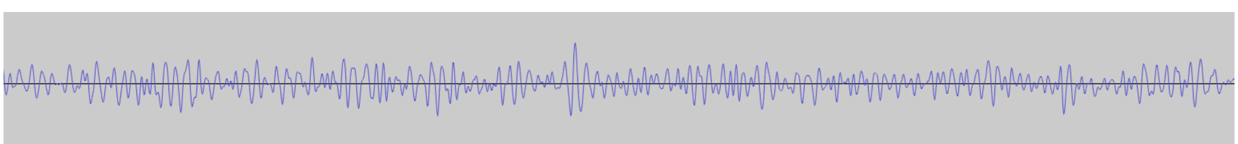
If the pattern of compressions and rarefactions is repeated regularly, that is if the rate of compression and rarefaction is kept constant for even a short length of time, we hear this as a particular *pitch*. We can describe this pitch with some measurement of how fast the wave oscillates between compression and rarefaction. We call this the *frequency* of the sound wave. That is to say, "how frequently does the cycle of compression to rarefaction and back again happen." Typically, we measure frequency in sound in *Hertz*, which is the same as one cycle per second. That is, the frequency of a clock is 1 Hz. A blinking light that flashes 5 times per second would have a frequency of 5 Hz. The lowest frequency common for humans to be able to hear is 20 Hz. Lower sound frequencies correspond to lower pitches (bass notes of a piano) and higher sound frequencies correspond to higher pitches (top notes of a piano, flutes, violins, whining noises from your fridge etc.)

We can use a special math function called the Sinusoidal function to draw pictures of sound waves. This function will be written like $A\sin(bt)$ where A is called the amplitude (which is a measure of how extreme the compressions and rarefactions are, and corresponds to the loudness of the sound) and $b/(2\pi)$ gives the frequency. Therefore, the function $3\sin(200\pi t)$ has amplitude 3 and frequency

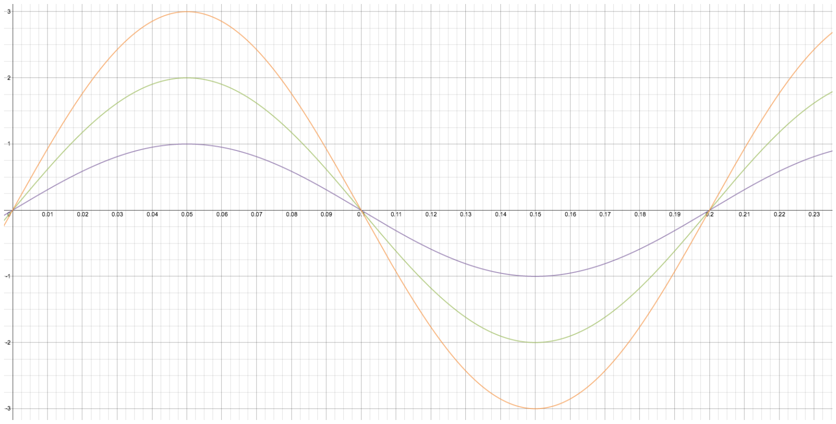
$(200\pi)/(2\pi) = 100$ Hz, or in other words, completes 100 cycles per second. From the point of view of your eardrum, we can draw a picture of this sound. In the picture below, we can think of the x axis as undisturbed air. Above the x axis is compression and below is rarefaction. We see that the air keeps going back and forth between compressed and expanded, and this happens at a regular interval, namely once per 1/100th of a second.



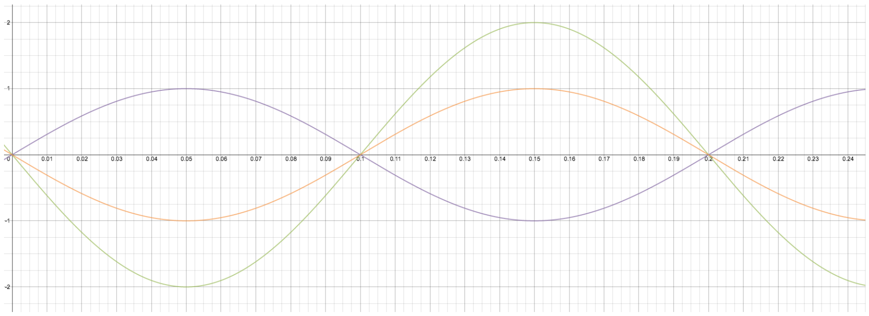
This is a representation of what is called a pure tone. These are rare in nature, but can be generated by a computer. A tuning fork is a good example of something that gives a close version of a pure tone. Most of the time, even the simplest sounds–like the sound of one note on a piano, or the hum of an air conditioner–are made up of combinations of lots of different sine waves. More complicated sounds like a cymbal crash, or speech is made up of a very complicated combination of sine waves. Pictured below is a picture of the sine wave representation of part of an audio recording of me reading that last phrase.



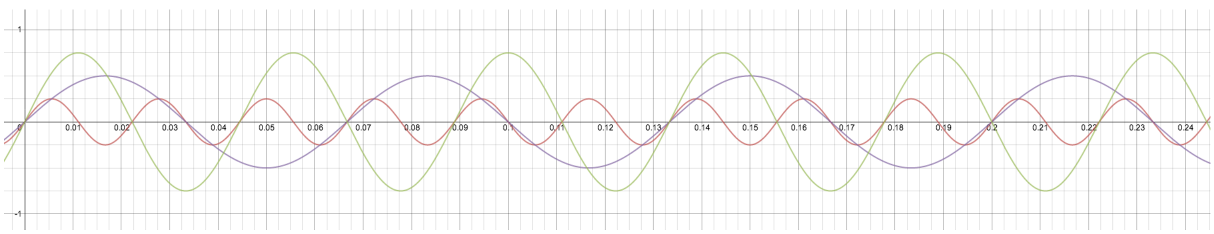
So, lets start simpler than that. Lets first try to combine only a few pure sine waves, and moreover, lets make note of the relationship between the frequencies. First, notice what happens when we add two sine waves of the same frequency. In the picture below, I have graphed $2\sin(10\pi t)$ in green, $1\sin(10\pi t)$ in purple, and the sum of them in orange. Notice that both of the green and purple waves reach the tops and bottoms of their oscillations at the same time. So the sum gets more positive at the tops and more negative at the bottoms. This is called *constructive interference*.

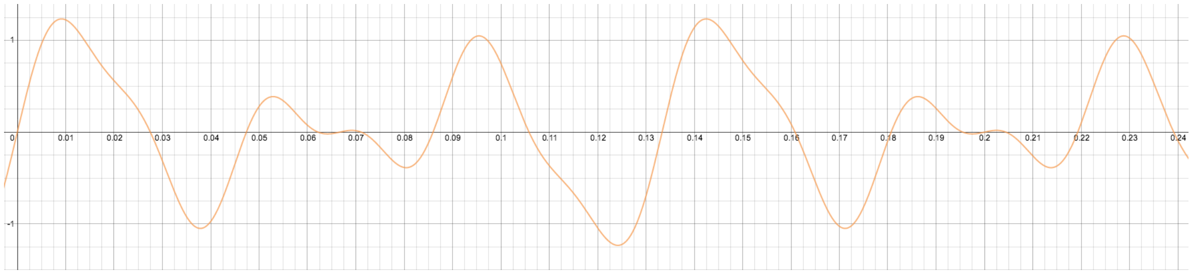


On the other hand, consider what happens when the green wave reaches its bottom when the purple wave reaches its top and visa-versa. Instead of being made bigger than both of these waves, the orange wave is made smaller than both of these waves. This is called *destructive interference*.

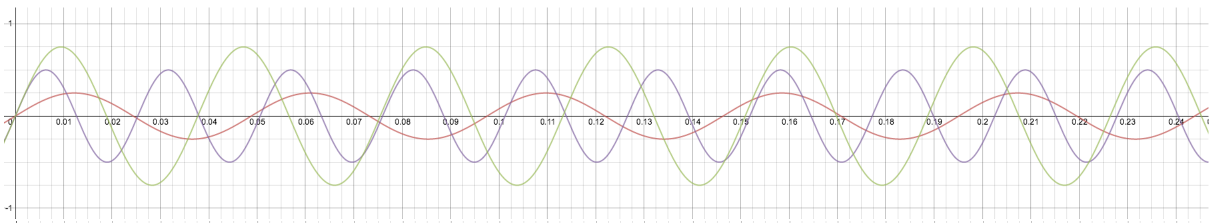


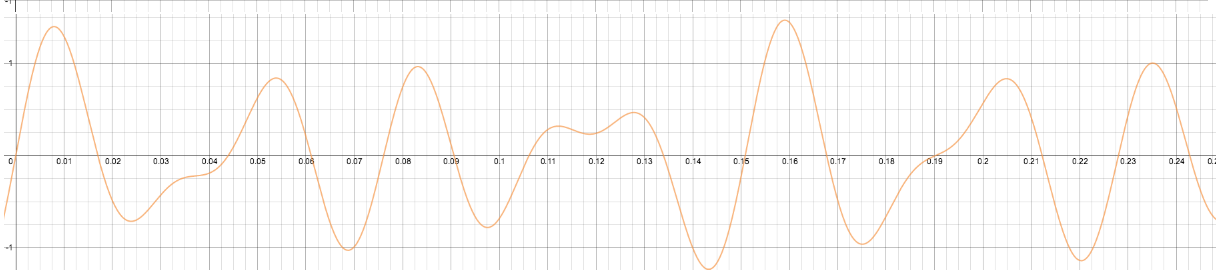
Now if we change the frequencies, we see that we get a combination of both constructive and destructive interference. First, consider the sine waves $\frac{1}{4}\sin(90\pi t)$ (red), $\frac{1}{2}\sin(30\pi t)$ (purple) and $\frac{3}{4}\sin(60\pi t)$ (green). The sum is shown in orange. Notice that each one of the numbers inside the function are closely related to each other, they are all divisible by 15. We see that $90 = 6 * 15$, $30 = 2 * 15$ and $60 = 4 * 15$. The corresponding frequencies are $(90\pi)/(2\pi) = 45$ Hz, $(30\pi)/(2\pi) = 15$ Hz, and $(60\pi)/(2\pi) = 30$ Hz, respectively. We see that these numbers are related in a similar way and that the ratio of each is a 'nice' fraction, for example 45 Hz/15 Hz = 3, or 15 Hz/30 Hz = 1/2. The resulting wave (orange) is more scattered than each of the three on their own, but does show an organized pattern.





Now, watch what happens if we pick frequencies that are unrelated. Consider the sine waves $\frac{1}{4}\sin(41\pi t)$ (red), $\frac{1}{2}\sin(79\pi t)$ (purple) and $\frac{3}{4}\sin(53\pi t)$ (green). We see that the sum (in orange) has a much more erratic behavior, and does not really admit any sort of nice pattern. Similarly, the relationship between the frequencies, 20.5 Hz, 39.5 Hz and 26.5 Hz is not nearly as 'nice' as the relationship of those above. For example, the nicest proper representation of the fraction 20.5 Hz/39.5 Hz is 41/79.





Which one of the above scenarios, when translated into sound, seems like it would sound nicer–the one with the regular pattern, or the one with the irregular pattern?

A couple of interesting side notes related to language:

(a) Each one of us has uniquely shaped vocal chords that output a unique combination of sine waves. This unique combination is what allows us to tell each other apart on the telephone.

(b) As we speak, we manipulate these sine waves with our vocal chords. As we change the frequencies, in English, we interpret this as inflection–that is, the difference between a question, or an exclamation, or a declaration or an affection.

(c) In some eastern languages, the pitch shape a word changes the meaning of the word. In Chinese, the same word can be said four different ways and mean four different things. Interestingly enough, native speakers of these languages are much more likely to have 'perfect pitch,' the ability to absolutely determine a musical pitch, without having to hear it in context or in relation to another pitch.

Math and Music

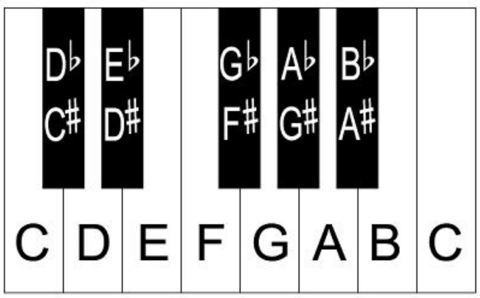
Pitch and Frequency

Now lets talk about the relationship to pitch and frequency and music. First, When we look at the sine waves, we notice a few things, first how tall are they? This corresponds to the loudness of the sound. The taller the waves, the louder the pitch. Second, we can look at how fast the oscillate. The faster the oscillation, the higher the note. The lower the pitch, the less frequent the oscillation. In terms of our formula $A\sin(f\pi t)$, if A is bigger, then the pitch is louder. If f is bigger, than the pitch is higher.

What would the sine wave look like if a pitch was simultaneously getting louder and lower? Higher and softer? Lower and softer?

Some terminology

Our western musical system is organized into twelve distinct pitches, arranged on a keyboard as follows. After the cycle of 12 notes finishes, it repeats, as can be seen with the left and rightmost notes both being 'C'.



Also, notice that some notes have two different names. The pound sign (pronounced 'sharp') means move up one key from the one before–which is why $D\sharp$ is one key above D. The little b (pronounced 'flat') means move down one key, which is why $E\flat$ is one key below E. Further, notice that the black keys have two names, simply because they are in between two different white keys. Lastly, notice that there is no $E\sharp/F\flat$ or $B\sharp/C\flat$.

We measure the distance between notes in 'steps.' The distance between two adjacent keys on the piano is called a 'half-step'. We assign names to specific numbers of half-steps.

Name of interval	Number of Half Steps
minor second	1
major second	2
minor third	3
major third	4
Perfect fourth	5
Tritone	6
Perfect fifth	7
minor sixth	8
major sixth	9
minor seventh	10
major seventh	11
octave	12

So, for example, if we wish to find the note that is, say a minor sixth above $E\flat$, we count 8 notes above this and find the note B is a minor sixth above $E\flat$. If we wish to find the note that is a Major third below F, we count down 4 half steps to $D\flat$.

Overtone Series and Timbre and resonance

An interesting physical phenomenon arises when a tone is played by an actual instrument or sung by a human, or created by some natural process. The tone itself is not the only one present. There are actually more tones that happen beyond this one. We call the tone created the 'fundamental tone' and the extra ones created are called 'overtones'. The next surprising thing is that these overtones always occur in the same pattern, but with various volumes depending on the sound source. When a tone with frequency f is created, tones with frequencies $2 \times f$, $3 \times f$, $4 \times f$ etc. are also created. So, if you create a fundamental tone with frequency 100, the frequencies $200, 300, 400, 500, 600, ...$ etc. also happen. The overtones happen with different volumes according to the source of the sound. This sequence of frequencies corresponds to a set sequence of intervals. The first overtone is an octave above the fundamental tone. The second is a perfect fifth above the first overtone and so on. The table below shows the first handful of overtones.

Overtone	frequency	Distance above last overtone (# of half steps)
fundamental	f	0
first	2f	octave (12)
second	3f	perfect fifth (7)
third	4f	perfect fourth (5)
fourth	5f	major third (4)
fifth	6f	minor third (3)
sixth	7f	minor third (3)
seventh	8f	major second (2)
eighth	9f	major second (2)
ninth	10f	major second (2)
tenth	11f	major second (2)
eleventh	12f	major second (2)
twelfth	13f	minor second (1)

The table, of course continues, but for the sake of simplicity, we will leave it there. On a keyboard, you can reproduce this overtone series as follows. Beginning on a note C towards the bottom of the piano, we can play this sequence of notes in ascending order: $\{C, C, G, C, E, G, B\flat, C, D, E, F, G, A\flat, B\flat, B\natural, C\}$.

Timbre is the thing that makes your voice sound different from your friend's voice. It is the thing that makes a clarinet sound different than a piano, even if they play the same note. As we said above, when you produce a natural tone, along with it come the overtones in various volumes. This combination of volumes gives everything its unique timbre. The physical nature of the object producing the sound is what determines the strength of the overtones. That is, the shape of your vocal chords, or the shape of the body of the violin determines how loud each of the overtones are.

The tables below show the fundamental tone and the volumes of the overtones for three different instruments. The first instrument playing a note at frequency 128 Hz will create an overtone series of the frequencies listed below 128. Notice that these are just $2 \times 128 = 256, 3 \times 128 = 384, 4 \times 128 = 512$, etc. The first and second overtone are twice as loud as the fundamental tone. The third overtone is slightly softer than the fundamental. All of the higher overtones are very soft (with exception of the seventh overtone).

f	volume %	f	volume %	f	volume %
128	50	73	30	349	100
256	100	146	90	698	10
384	100	219	10	1047	50
512	30	292	5	1396	10
640	5	365	2	1745	3
768	5	438	2	2094	2
896	0	511	2	2443	1
1024	30	584	1	2792	1
1152	5	657	2	3141	1
1280	5	730	1	3490	1

Finally, resonance in a vibrational system is a phenomenon where small vibrations at a certain frequency are 'echoed' and turned into big vibrations. Sometimes the results can be destructive, like singing really loud at the right pitch towards a wine glass. Sometimes the result can be catastrophic like wind rocking the Tacoma narrows back and forth at the right frequency. In music, the results are beautiful. A well-designed instrument is made so that the right overtones are made louder and 'resonate' in the body of the instrument. Similarly, concert halls should be designed so that the music echoes in the right way to amplify the right overtones.

Pythagorean Tuning

Pythagoras didn't just study triangles, he studied music. As we've talked about, he loved ratios of whole numbers. Legend has it he was passing by an anvil shop and noticed a relationship between the pitch produced when the blacksmith hit two anvils of different sizes.

He began to explore the relationships with pipes of different sizes. He found that when he hit two pipes that were the same length, they produced the same tone. When he hit two pipes, one of which was half as long as the other, he got the same tone, except the shorter pipe was 'higher'. Further exploration showed that when he hit a pipe that was 3/2 as long as another pipe, the two notes were not the same, but sounded good together. But when he hit pipes that had a length ratio of 1024/729, the tones sounded bad. He concluded that when you take a ratio of the frequencies of the vibrations, and the fraction in lowest terms involves small whole numbers, the pitches sound good. When they involve large whole numbers, the pitches sound bad.

Below is a chart of the ratios between the frequencies of the pitches in each corresponding interval. For example, we have that the frequency of the note A near the middle of the piano is 440 Hz. The frequency of the A above that is 880 Hz. So the ratio of their frequencies is $880/440 = 2/1$. The frequency of the note E (a perfect fifth) above that A is 660 Hz. So the ratios of these frequencies is $660/440 = 3/2$. The rest of the table is computed similarly.

Building Chords

Finally, we turn our attention to 'chords,' the building block of music. A chord is a collection of three or more notes played all at the same time. We will pay attention to a couple of three-note and four-note chords.

interval	ratio (higher frequency / lower frequency)
minor second	256/243
major second	9/8
minor third	32/27
major third	5/4
Perfect fourth	4/3
Tritone	1024/729
Perfect fifth	3/2
minor sixth	128/81
major sixth	27/16
minor seventh	16/9
major seventh	243/128
octave	2/1

Three Note Chords

First, consider the major chord. A major chord is a tone (called the root), a major third (4 half steps) above the root and a minor third (3 half steps) above the second tone. For example, if our root is C, then the rest of our chord is E and G. Notice that the third note is a perfect fifth above the root. According to our rule of small whole-number ratios, all of these pitches sound good together, and so the chord sounds good.

Next, consider the minor chord. A minor chord is a root note, with a minor third on top (3 half steps), with a major third (4 half steps) on top of that, ($C, E\flat, G$, for example). Again, notice that the top note is a perfect fifth (7 half steps) above the root note. All of these notes sound good with one another, and so the chord sounds good.

Penultimately, consider a chord that is a root note, a major third and another major third ($C, E\flat, A\flat$). Again the first and second notes sound good, the second and third sound good, but the first and third sound less good. The chord overall does not sound as good as the first two. This is called an augmented chord.

Lastly, consider the chord that is a root note, a minor third (3 half-steps) and another minor third (3 half-steps) ($C, E\flat, G\flat$, for example). The first and second notes sound good, and the second and third notes sound good, but notice that there is 6 half-steps between the first and the third note, which we decided above sounds bad. So this chord sounds bad–or to use a better word, dissonant. We call this a diminished chord, and we will discuss why dissonant chords are important in music in a second.

Four Note Chords

There are lots and lots of different four note chords, but I only want to pay attention to two of them.

First, consider the chord C, E, G, B. The intervals between these notes are a major third, minor third and major third. Between C and G, there is a perfect fifth and between E and B, there is a perfect fifth. In between C and B there is a major seventh. The major seventh is a fairly dissonant interval, but this chord is so full of consonant intervals that this chord ends up sounding pretty consonant. This is called a 'major seventh' chord.

Now consider the chord $C, E, G, B\flat$. This is the same as the major seventh chord but the top note is lowered by one half-step. Now, the interval make up is a major third, minor third and minor third. There is still a perfect fifth between C and G, but now there is a tritone between E and $B\flat$. Furthermore, there is a minor seventh between C and $B\flat$. This chord sounds dissonant.

Chords in Music

Most of the music we normally listen to is based off of one scale. We should all be familiar with what we call the Major scale, which is do-re-mi-fa-so-la-ti-do. If not, go watch The Sound of Music. What's more, the music only uses chords built from the notes from this scale. In musical terms, we say that the music has a 'tonal center' and makes us think of one note as a home note. Even though the chords may not contain that note, they are related in such a way to always return home. The interesting bit is how the chords interact with our perception of home. Most musicians look back to J.S. Bach (1685-1750) for being the guy who figured out some basic rules about how to use these chords to make the music we are used to today.

Returning to the idea of chords above, if we are 'in the key of C major,' then a C major chord sounds like we are at home. Now consider a B diminished chord. Recall this is a B then a minor third above that (D) and a minor third above that (F). This chord does not contain C and it contains two notes that are dissonant with C. So, if C sounds like home to us, then this chord sounds unstable. It feels like it needs to become stable, and one way to do that is to 'resolve' to a consonant chord. Now consider the chord G, B, D, F. Notice that this is the same as a B diminished chord, except with a G in the bottom. The effect of this G in the bass is to strongly tie this chord with the home key of C, since C and G are a perfect fifth away, but the dissonance between the home key and the other three notes remain. The chord now feels unstable, but the G gives it a direction in which to resolve, in particular, back to C.