

4 音の表現

4.1 音のデジタル化

一般に、空気中で物が振動すると音が発生します。例えば、太鼓を叩いた場合、太鼓の膜が振動して「ドン」という音がします。このような音は、物理的に言えば、空気が振動する波です。

音を電気信号へと変換するのに用いられる機器としてはマイクがあります。紙コップに向かって「あー」と発声すると、紙コップの底が振動しますが、その紙コップの底にあたるのが振動板です。そしてこの振動板には、コイルが取り付けられています。コイルが磁石の中で振動すると、コイルに起電力が生じ、この原理により、音を電気信号へと変換することができます。

マイクで電気信号へと変換した音は、そのままでは、連続的な値をとるアナログ情報です。一方、コンピュータ上で扱うことができるのは、飛び飛びの値をとるデジタル情報です。このため、音をコンピュータ上で扱うためには、アナログ情報をデジタル情報へと変換するA/D変換を行う必要があります。A/D変換は、Analog Digital変換の略です。A/D変換は、標本化や量子化といった処理により行われます。

まず、標本化ですが、標本化はサンプリングとも呼ばれ、一定間隔で値を取り出す処理です。標本化を行う場合、もとの信号の再現性を高めるためには、サンプリングを行う時間間隔を短くする必要があります。1秒間にサンプリングを行う回数をサンプリング周波数と呼びます。一般に、サンプリング周波数を上げれば上げるほどデータの再現性は高まります。しかし、データ量はサンプリング周波数に比例して増加します。

通常、電話回線はサンプリング周波数が8 kHzです。すなわち、1秒間に8000回のサンプリングが行われます。一方、音楽CDでは、サンプリング周波数が44.1kHzです。すなわち、1秒間に44,100回のサンプリングが行われます。

次に、量子化が行われます。量子化は、サンプリングにより得られた値を、飛び飛びの値へと変換する処理です。一般に、量子化の際に取りうる値が多いほど、もとの信号の再現性が高まります。

例えば、電話回線では、量子化ビット数が8 bitです。すなわち、256段階の値でもとの情報を表現します。また、CDでは、量子化ビット数が16bitです。すなわち、約6万5千段階の値でもとの情報を表現します。

量子化して得られた値を、コンピュータが処理できる0と1の系列に変換する処理は、符号化と呼ばれます。符号化された音声データを記録するための形式（フォーマット）としては、MP3、WAVE、AIFFなどがあります。

4.2 音の3属性

音を特徴づける属性には、音の大きさ、高さ、音色があり、これらは、音の3属性と呼ばれます。

4.2.1 音の大きさ

音の大きさは、もとの空気の波で考えれば、密度が高い部分と密度が低い部分の圧力差に相当します。小さな音では、密度が高い部分と密度が低い部分の圧力差が小さくなります。すなわち、横軸が時間、縦軸が音圧のグラフを考えると、小さい音では、振幅が小さいということになります。一方、大きな音ほど、密度が高い部分と密度が低い部分の圧力差が大きくなります。グラフで考えると、大きな音では、振幅が大きいということになります。では、圧力差が大きい場合、何が起こるのでしょうか。音の振動の中に物体が置かれている場合、その物体は、音の振動により、圧力が高い方向から圧力が低い方向へ力を受けます。そして、この力は、圧力差が大きいほど大きな力になります。ですから、大きな音ほど、それを受けた物体は、より大きく振動します。

音の大きさの指標としては、フォンとデシベルがあります。フォンは、人の感覚に基づく音の大きさの指標です。人にさまざまな大きさの音を聞かせて、それを数値化することで音の大きさを決めたものです。一方、デシベルは、物理的に計算できる音の大きさの指標です。音の大きさの計算式で対数が使われますが、これは、人の感じる音の大きさは、物理的な音の大きさそのものではなく、ほぼ、その対数で感じているためです。物理量を対数的に感じることは、感知する物理量の幅が大きい場合に都合が良くなっています。

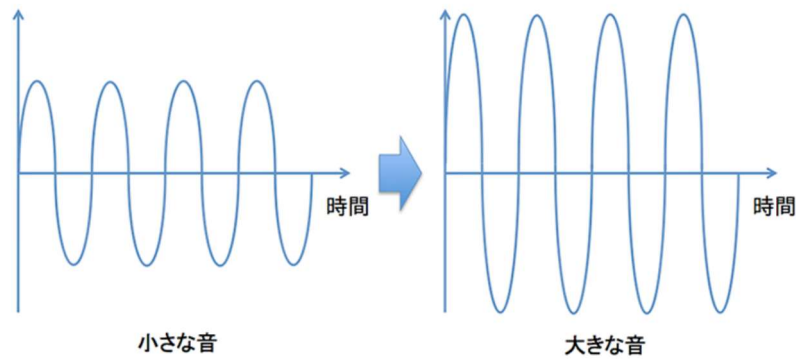


図4-1 音の大きさ

4.2.2 音の高さ

音の波形を見ると、通常、同じような波形が繰り返し現れます。このような波が1秒間に繰り返し現れる回数は周波数と呼ばれ、単位は、Hzが使われます。周波数が高いと高い音に、周波数が低いと低い音になります。

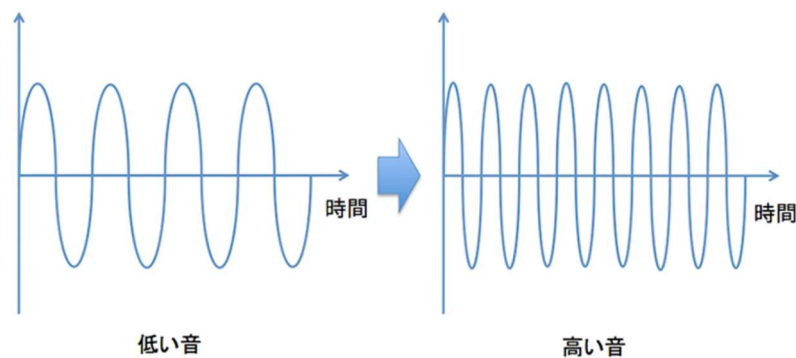


図4-2 音の高さ

音の高さについて、音楽の立場から見てみましょう。オーケストラで演奏前に音合わせをする際の音は、ラの音で、この音の周波数としては、通常、440Hzが利用されます。一般に、音は、周波数の比が整数倍であれば音が響きやすくなり、特に、ある音とその2倍の周波数の音は、波が周期的に丁度揃うことから、響きやすくなります。このような周波数が2倍違う音は音楽で言うところの1オクターブ違う音になります。一般に、ある音の周波数を1とした場合、 n オクターブ上の音の周波数は、 n 回2倍にしていくことから、 2^n 倍となります。具体的には、例えば、3オクターブ高い音は周波数が2の3乗で8倍になります。また、2オクターブ低い音の周波数は、 2^{-2} 倍となるので、周波数は1/4倍となります。

次に、半音について考えてみましょう。ピアノの鍵盤を見ると、ドから1オクターブ上のドまでの間で、白鍵と黒鍵が12個並んでいるのが分かります。これらの鍵盤の音は、音の高さの差が同じに感じるように並んでおり、これらの音の間隔は半音に相当します。すなわち、半音12回分音が上がると1オクターブ分、音の高さが上がることになります。ここで、人間の音の高さの感覚が対数的であるとすると、これらの音は、周波数の差ではなく、その比が一定ということになります。この半音の周波数の比がいくつになるか実際に求めてみましょう。

半音の周波数の比を X とします。そして、半音が12回分上がると1オクターブの違い、すなわち周波数が2倍になります。すると、周波数の比について、 X の12乗=2が成り立ちます。この式を解くことで、 X の値は、1.05946..と計算できます。すなわち、半音違う音とは周波数が1.0595倍違う音ということになります。

これまでの知識をもとに和音を考えてみましょう。ドミソの和音を例に考えてみます。ドの音の周波数を1とします。ミの音は、半音4つ分高いので、周波数は X の4乗となります。また、ソの音は、半音7つ分高いので、周波数は X の7乗となります。ここで、先ほど求めた $X = 1.05946..$ を用いると、ドミソの和音の周波数比は、1 : 1.2599 : 1.4983 となります。ここで、全体を4倍すると、この周波数の比は、4 : 5.0397 : 5.9932 となり、この周波数比は、ほぼ4 : 5 : 6 という簡単な整数比になっています。これが、ドミソの和音は美しく響く理由になります。

このように、半音の周波数比を一定とした音階は平均律と呼ばれます。一方で、音の響きをより良くするために、和音の周波数比が完全に整数になるように調律する方法もあります。これが純正律です。しかしながら、純正律は、転調した場合に和音の周波数の比が整数比から大きくずれるという問題があります。このため、転調する曲や、異なる調の複数の曲を演奏する場合などには向いていません。このため、通常は平均律が広く利用されます。

4.2.3 音色

次に、音色について考えてみましょう。例えば、同じ大きさで、同じ高さのベルの音と、バイオリンの音があったとします。この2つの音は、明らかに違う音として認識されます。この音の違いが音色の違いになります。このような音色を決める要素は2つあります。音色を決める要素の1つは、エンベロープです。例えば、図4-3(a)はベルを叩いた場合の音の波形を示した例ですが、このベルの波形をみると、振幅すなわち音の大きさが最初は大きく、すぐに減衰しているのが分かります。また、図4-3(b)はバイオリンの弾き始めの音の波形ですが、こちらでは、弾き始めから振幅は徐々に大きくなっています。このような振幅を囲む包絡線はエンベロープと呼ばれます。このエンベロープは、音色を決める1つの要素となります。

音色を決めるもう1つの重要な要素として、音を構成する波形の違いがあります。図4-4(a)は、短い時間範囲でのベルの音を表示したものです。一方、図4-4(b)は、バイオリンの弾き始めの音です。これらの音は、周波数がほぼ同じであるため、音の高さはほぼ同じで、また、振幅の大きさもほぼ同じ程度であり、音の大きさもほぼ同じです。しかし、繰り返し現れる波の形が異なっているため、ベルの音とバイオリンの音はかなり違って聞こえることとなります。

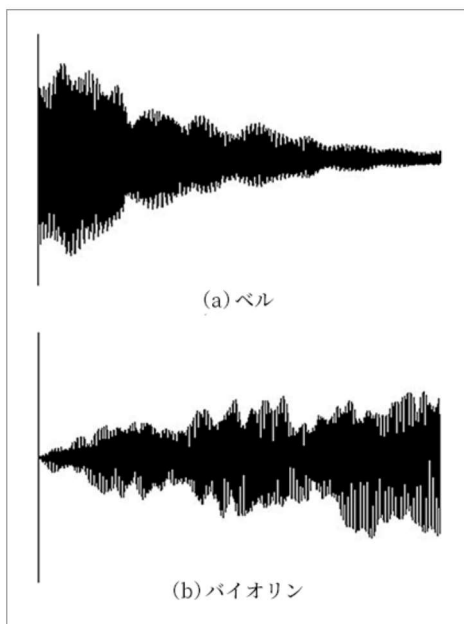


図4-3 エンベロープ

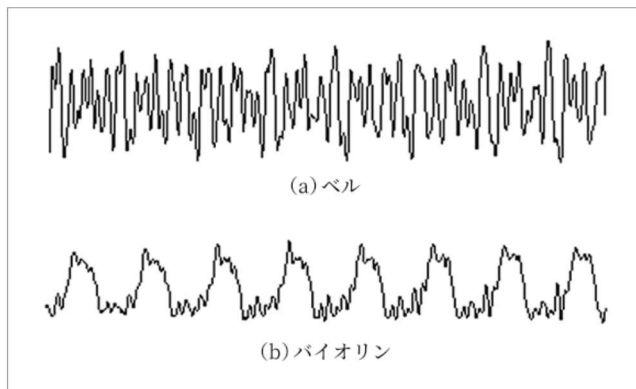


図4-4 波形の違い