とーくばっく

～デジタル・スタジオの話～

第３版

David Shimamoto

Studio Gyokimae

まえがき ～ とーくばっくの歩き方

「とーくばっく～デジタル・スタジオの話 第３版」を手にとってくださり、ありがとうございます。
前書きに代えまして、雑多な話題を含む本書の歩き方について、簡単ながらご案内いたします。

2017 年 1 月の初版リリース後、これが何の本であるか著者自身が説明に困る場面が多くございました。
逆に何の本でないかはわりとはっきりしていますので、皆様の期待を裏切ることのないように先に申し上げておきます。

「デジタル・スタジオの話」とタイトルにある通り、この本は DAW などデジタル音響機器の、内部の挙動や使用時の考え方に関する話が中心です。そのため、アナログ領域の話はほとんどありません。機材の選び方には一切言及しませんし、お読みいただいてもマイキングや特定の機器の取扱いが上達することはありません。ルーム・チューニングや反響の抑制などといった、アコースティックに関する知識も身にはつきません。収録現場での立ち振る舞い（ディレクションやミュージシャンとのやりとり）については…ほんの少しだけ触れる場面もありますが、役に立つほどのものかどうかはわかりません。いずれのトピックもそれなりの資料が現存するため、私があらためて書くまでもないかと思います。

では一体、なんの本なのか？

大枠としては音楽制作、それもミックスやマスタリングにまつわる本であることは間違いありません。
しかし、特定の作業環境は想定していませんし、初心者が音楽制作をはじめる際に必要となる知識を段階的に与えるものでもありません。

この本は、著者自身が細なが～い創作活動を通じて、誰も教えてくれなかったこと、あるいは誤った情報を掴まされやすいと感じたことを中心に、あるときは海外の文献を漁り、またあるときは自身が実験した結果をまとめた、とりとめのない資料集です。しかし、その資料の重要性とレアリティ（希少性）に関しては、あくまで個人による自費出版誌である本書が、音楽大学、専門学校をはじめとする複数の教育機関にて講義資料として採用いただけたことにより、少なからずお墨付きをいただけたものだと考えています。

この本を読めば、多少はミックスやマスタリングの技術が向上するかもしれません。
これまで制作の道々、直感で下していた判断が、より論拠に基づくものになるかもしれません。
また、エンジニア的な視点を得ることで、作編曲の技術も向上するかもしれません。

次に、本書の構成を、各章の生い立ちと併せておおまかに説明します。

まず第２部４章では、日本語の文献が極端に少ないラウドネスの話題について取り上げています。これは音楽制作関係者に限らず、関心のあるすべての方に読んでいただけるよう心がけて書きました。5、6章は、４章を読んでくださった方々からのご意見をもとに、理解が難しいと思われた箇所を補足するため、主に音楽制作経験者に向けて書きました。大手配信プラットフォームでラウドネス・ノーマライゼーションが一般的になりつつある昨今、このあたりの知識は近年いよいよ感覚では済ますことのできない、また知らなかったでは済まされないものになりつつあります。

前後しますが第１部は、これも音楽制作の経験者を対象に、４章の理解を助けると思われるトピックをいくつか選んでいます。知らなくても大ヤケドはしないかもしれませんが、知らないがために損をするよりは、知っておいた方がよいと思われる、録音にまつわる話です。

第３部はやや実践的なトピックを中心に、やはり「誰も教えてくれなかったこと」をまとめました。

第４部以降は、著者の主観がより色濃く反映されており、議論の余地があるため明示的に切り離したトピックや、より断片的な資料の置き場です。固い技術の話から随筆まで、内容は多岐にわたります。筆者としては第３部の延長線上にあるものと考えていますが、あくまでもいち個人のいち意見が中心であることを念頭に置いてご覧いただき、各々が内容の是非をご判断いただければと思います。

このように段階的な話はほとんどなく、トピックの重要性、はては想定読者でさえ章ごとに異なります。したがって、目次をご覧いただき、なんとなく面白そうなところから順に読んでいただいても結構です。

「で、結局この本は役に立つの？」

おそらくここにまとめた資料は、商業音楽で成功する上では必須ではないでしょう。現に、本書の内容の半分も理解せずに業界の各分野をサバイブされている方々が圧倒的多数ではないかと思います。もう少し踏み込んで申しますと、ここに取り上げるトピックの一部は、長らく制作に携わっている方々に対して軽々しく話題にしようものなら talk back（口答え）しているように受け取られるものもあるかもしれません。しかし、以前からなんとなくモヤモヤを感じていた方々には、胸のつかえがスッキリと取れたように感じる場面もご用意できたのではないかと思います。（実際に初版リリース後はそういったご意見を多数頂戴しました。）なにより、Loudness Warについて他者に説明できるだけの理論と語彙を身に付けたい方にとっては、現存する唯一の日本語書籍なのではないかと自負しております。ラウドネスの話ばかりでもありませんので、もうその話題は飽きたという方々も、読み終えたあとになにかしら新しい知見を得て、本書と似たような価格帯のプラグインよりは断然レベルアップに繋がったと思っていただければ幸いです。

目次

まえがき～とーくばっくの歩き方 2

第 1 部　PCM 録音の話

1 章　ビット・デプスの話 6
2 章　ディザリングの話 10
コラム：有効ビット長の測定 16
3 章　サンプル・レートの話 18

第 2 部　収録レベルの話

4 章　収録レベルの話 26
5 章　メータの話 42
コラム：K-System とは 47
6 章　ダイナミック・レンジの話 50
コラム：ラウドネス測定手順 56
コラム：M・S・I
EBU Mode はラウドネス・メータの業界標準となるか？ 58

第 3 部　ノウハウ集

7 章　M/S 処理と留意事項 62
コラム：モノラル互換性について 69
8 章　ネット配信時代 ラウドネス規定導入後のミックス手法 70
コラム：オフ・ボーカルへのリミッティングについて 72
9 章　抑揚をアレンジに活かす 74
コラム：ビット長の現状 77
10 章　補正可能時代の収録ボーカリストが知るべきこと 78
11 章　＋ / －のピークが非対称な波形について 82
コラム：アナログＥＱ＆デジタルＥＱ 87
12 章　次世代音声コーデック "Opus" 概要 88

第 4 部　森の奥深くへ

4-1　YouTube「統計情報」の読み方とラウドネス最適化 92
4-2　トラックの左右配置とステレオ方式の限界 94
4-3　ゲイン・ステージング 98
4-4　音楽性を高めるボーカル・エディット 103
4-5　Vo. フェーダ・オートメーションのススメ 108
4-6　ピーク・レベルの適正値に関する略史 110
4-7　ヘッドルームとコーデックの落とし穴 114
4-8　ダウン・コンバートの手順 118
4-9　英文資料に見る 3 つのキーワード 121
4-10　浮動小数点数のダイナミック・レンジ 124
4-11　CD-R ブランドで音は変わるか 126
4-12　波形とは？海苔波形とは？きざみ海苔波形とは？ 131
4-13　ハイレゾの話 136
4-14　Twitter アンケートに見る Loudness War 139

最終章　～デジタル・スタジオから「音圧」が消える日 142
付録 A　市販曲ラウドネス一覧 148
付録 B　手動ラウドネス・ノーマライズ用 早見表 153

参考文献 154
謝辞 155

あとがきに代えて～表紙改訂の経緯の話 156

第１部
ＰＣＭ録音の話

1章 ビット・デプスの話

DAW やポータブル・レコーダといったデジタル機器に、マイクやラインのアナログ音声を録音するとき、入力された音声信号は A/D コンバータ（Analog to Digital Converter）により数字の列に変換されます。これは早い話、音声信号の瞬間ごとの電圧値を、毎秒何万回と等間隔で測定（サンプリング）することにより実現しています。読者の皆様も、右の図 1 のように信号波形を方眼紙のマス目に落とし込んだグラフをご覧になったことがあるかも知れません。横軸が時間、縦軸がその瞬間における音声信号の電圧値を表します。

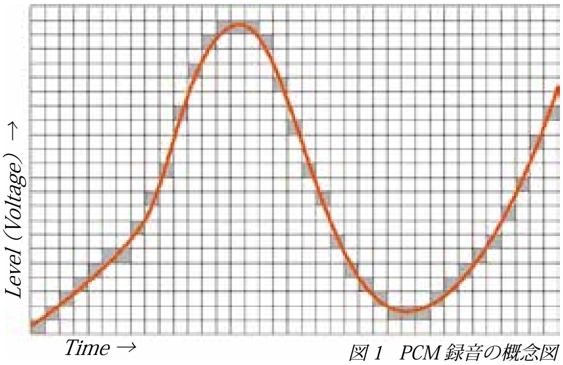


図 1　PCM 録音の概念図

ここで、録音機器やデジタル音声ファイルのスペックなどによく登場する 2 つの語句を確認しましょう。

ビット・デプス（Bit Depth）

グラフの縦軸の解像度で、電圧を測る精度を表します。ビット・デプスが大きいほど方眼紙のタテの目が細かくなります。1 ビット増えるたびに解像度…すなわちマス目の数は 2 倍になります。

サンプル・レート（Sample Rate）

グラフのヨコ軸の解像度で、 1 秒ごとに信号電圧を測定する回数を表します。

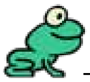
たとえば 16bit/44.1kHz（CD スペック）では、2^{16} 段階での電圧値測定を、毎秒 44,100 回行います。

ここまでの説明は、多くのコンバータや DAW が使用する**パルス符号変調（PCM）方式**に関するものです。
SACD などが使用する **DSD 方式**は、まったく異なる手法で信号レベルをデジタルで表現します。
サンプル・レートについては後の章で考察するとして、本章ではビット・デプスと DAW の内部構造についてお話しします。

ビット・デプス ＝「信号レベル」の解像度

CD に記録された音声は、1 サンプルあたり 16 ビットの解像度で信号レベルを表現します。 また、いわゆるハイレゾ音源は、大抵 24 ビットで配信されています。本書の執筆時点では、業務用 A/D & D/A コンバータの多くは 24 ビットが上限です。

デジタル・オーディオで使用される 16/24 ビットは、正確には**符号付整数**とよばれる数値として扱われます。これはプラスまたはマイナスを表す符合を持つ整数値です。たとえば 16 ビット符号付整数であれば、先頭の 1 ビットは＋／－のいずれかを表し、残りの 15 ビットが整数を表します [*1]。



*1 オーディオでは使用しませんが、コンピュータの世界では「符号なし整数（unsigned integer）」というものも使用されます。ゼロ以上の整数しか表現できない代わりに、同じビット長で表せる最大値は符号付（signed）整数の 2 倍になります。

デジタル・オーディオの信号レベルを表す際に表現可能な範囲は、以下の通りです。

- 16 ビット：-32,768 ～ 32,767
- 24 ビット：-8,388,608 ～ 8,388,607

1 ビット増えるごとに解像度は 2 倍になるので、16 → 24 ビットでは、信号値の表現に使用できる数値は、おおよそ 256 倍になります。ただし、符号付「整数」とあるように、小数は表現できません。また、後述するように「方眼紙の目が細かくなる」だけで、測定できるアナログ信号レベルの上限値が増えるわけではありませんのでご注意ください。

16、24 ビットいずれも、表現できる範囲の下限、上限が、ピーク・メータが示すところの 0dBFS に相当します。

では、32 ビットはどうでしょう？
実はこれは、先ほどの 16/24 ビットとは根本的に性質が異なります。先ほどの 16/24 ビットが符号付整数であったのに対し、DAW の文脈で語られる 32 ビットデータは、**浮動小数点数**、または **32 bit float** といいます。
この「浮動」というのがポイントです。なにが「浮いて」「動いて」いるのかというと、小数点を打つ場所です。
これにより、整数しか扱えない 16 ビットや 24 ビットのオーディオ・ファイルでは不可能な「123.45」や、「0.000012345」といった小数の表現が可能になります。ここで例示した数値は、いずれも小数点を除く部分「12345」は共通ながら、小数点の位置だけが変化している点にご注目ください [*2]。

以降、32 ビット浮動小数点数を「32bit float」と表し、「16/24bit」は暗黙に符号付整数を指すものとします。

32bit float の表現範囲

符号付整数である 16bit、24bit の違いは、グラフのタテの目の解像度だけです。それぞれが表現できる最小／最大値は、同じ信号電圧を表します。これを 32bit float に変換すると、コンバータが入出力できる電圧値の上限、下限がそれぞれ +1.0、-1.0 に対応します。(図 2)

浮動小数点数では -1.0 ～ +1.0 の範囲外にある数値も取り扱えます。例えば、+1.0 より 0.1 だけレベル・オーバーしてしまっても、+1.1 という値として扱えばよいわけです。このため、トラックや AUX バスの出力が 0dBFS（16/24 ビットで表現できる上限）を超えてしまっても、マスター・バスを出る（すなわち D/A コンバータに送られる）までに、再び -1.0 ～ +1.0 の範囲に戻してやれば、音がクリップすることはありません。もうおわかりのように、これは 32bit float の方がデータ量が多いからというわけではなく、数値を表す体系がまったく異なるためです。

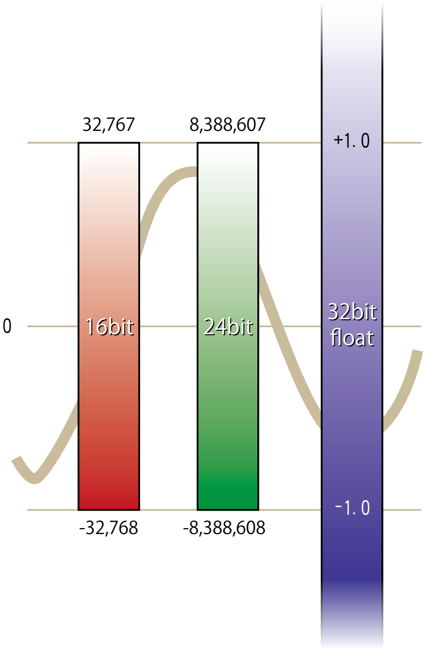


図 2　ビット・デプスごとの表現範囲

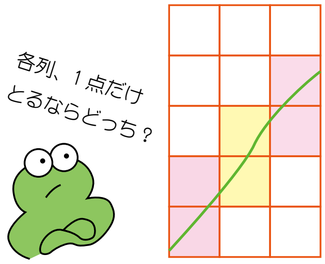
※本章の続きは本誌にて

*2 概念を説明するため例では 10 進数を使用していますが、実際は 2 進数のため若干処理が異なります。

2章 ディザリングの話

PCMデータを取り扱う時に注意すべきことのひとつが、サンプリングする解像度では拾いきれない信号の適切な処理です。 グラフのヨコ軸（時間軸＝サンプル・レート）の話は次章に譲るとして、本章ではタテ軸（信号レベルの記録精度＝ビット・デプス）に話を絞ります。

第1章「ビット・デプスの話」にもあったように、アナログ信号のデジタル化は、瞬間ごとの信号電圧を測定することで可能となりますが、測定した値は整数（方眼紙のマスのいずれか）に落とし込まなければなりません。入力信号のレベルに「ドンピシャ」で合う位置にマス目があればよいのですが、アナログ信号が無限の解像度を持つ以上、それは望めません。やむなく、最も近いマス目に丸め込みながら、電圧値を記録していくことになります。



さて、こうして最も近いマス目に丸め込む（あるいは無理矢理押し込む）という作業を続けるとどうなるかというと、この誤差が累積し、最終的には元々は存在しなかった周波数成分が「捏造」されることになります。この捏造された信号は、レベルこそ低いものの、再生音量を上げると聞き取れますし、肝心の楽音をマスキングしてしまいます。
あとでご紹介しますが、この現象を逆手にとれば、本来そのビット・デプスでは再現できないほど微小な音をよい具合に捏造し、再現してしまうというウルトラCも可能です。

マス目への落とし込みが必要になる場面

オーディオ信号を方眼紙のマス目に丸め込む作業、つまり先に述べた捏造が生じうる場面はいくつかあります。
主なものは、

1．生音を録音するとき
＝ 無限の解像度を持つアナログ信号の電圧を、A/D コンバータが近似する整数値に丸め込むとき

2．デジタル領域でビット・デプスを落とすとき
＝ サンプリング済みの信号レベルを、元よりも解像度の低い方眼紙に落とし込むとき

結論を先にいうと、このように最も近いマス目に数値を丸め込むことでニセモノの周波数成分が生じる問題は、**意図的にノイズを加えることで軽減できます**。一見すると矛盾するようですが、適切なノイズを加えることで、信号精度はむしろ保たれるのです。

上記「1. 生音を録音するとき」については、コンバータ自体が持っているフロア・ノイズ[*1]がこの役割を果たすので、普段は意識する必要がありません。今日、最もハイエンドな A/D コンバータにおいても同様です。
ここからは、「2. デジタル領域でビット・デプスを落とすとき」について説明します。

*1 アンプなどアナログ機器のボリュームを上げると「サーッ」と聴こえるアレです。

ディザリングとは

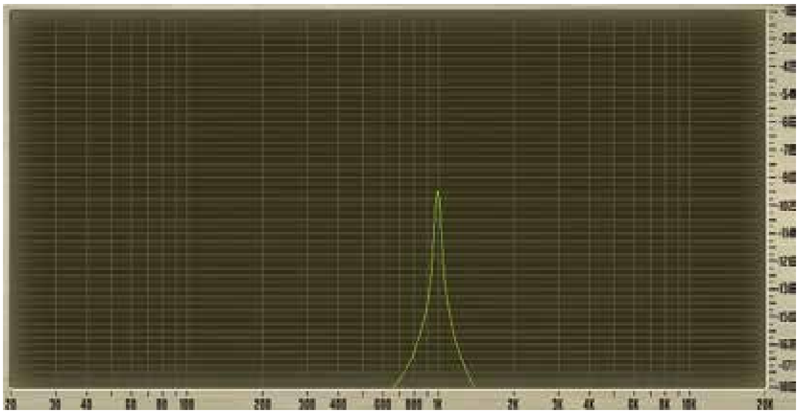
デジタル領域でビット・デプスを落とす（ダウン・コンバートする）と、元々は異なる数値が同じ数値に丸め込まれます。これは音声に限らず、デジタル画像においても色数を落とすときなどに同じような問題に直面します。例えば元々256段階の濃淡で表していたモノクロ画像（下図左）を、白黒の2色に落とし込む場面を考えてみましょう。人物の写真などであれば変換後でもヒトと認識はできるかもしれませんが、細かな陰影は失われ、平面的になってしまいます。

左から、中間階調を含む画像、各ピクセルを単純に白黒の2色に変換した画像、白黒に変換する前にノイズを加えた画像。

（Image: Martin Valigursky / Bigstock.com）

音声信号レベルを近似する整数に丸め込む場合、信号波形が歪む分、元の信号にはなかった倍音が発生してしまいます。この倍音の発生を回避すべく、意図的にノイズを加える作業のことを**ディザリング**（Dithering）や**ディザ**（Dither）といい、その際に加えるノイズを**ディザ・ノイズ**（Dither Noise）といいます。
ディザの適用範囲はコンピュータを使用する場面に限りません。桁数の不足などで計算精度が足りずに誤差が生じ得る状況であれば、紙面で計算を行う場合でも同様の問題に直面する可能性があります[*2]。

では実際に、32bit floatで生成した1kHzの正弦波を、ディザリングを行わずにCD相当である16bitに変換した場合、どうなるかを見てみましょう。
まず、下図は32bit floatで出力された1kHzの正弦波を、スペクトラム・アナライザで見たところです。



1kHz正弦波を生成しているプラグインの出力（32bit float）

単一の正弦波以外の周波数成分を含まないので、1kHzの1箇所にだけ信号が立っているのが見られます。

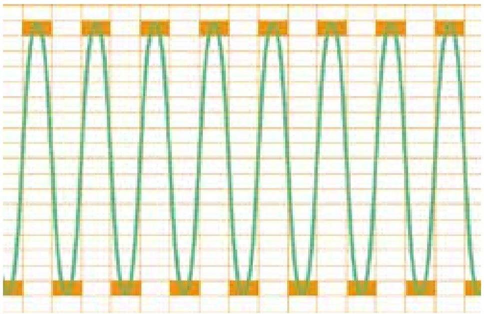
※本章の続きは本誌にて

*2 Wikipediaによると、語句の起源は第二次世界大戦時にまで遡（さかのぼ）るそうです。弾道計算に使用されていた機械式計算機が、不安定で振動の多い爆撃機内で使用したときの方が有用な計算結果をはじき出したとか…

3章 サンプル・レートの話

これまでの章では、PCM サンプリング時に方眼紙のタテ軸にあたるビット・デプスの話をしました。本章では、ヨコ軸にあたる**サンプル・レート**について考察します。

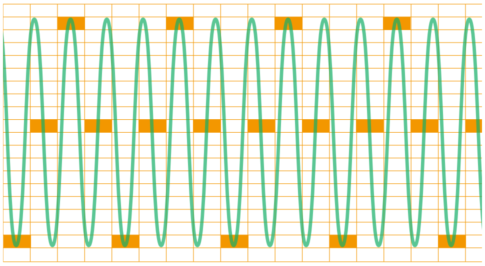
サンプリング可能な最大周波数（ナイキスト周波数）



PCM 方式では、サンプル・レートのちょうど半分の周波数が、記録できる最大周波数になります。また、これを**ナイキスト周波数**といいます。サンプル・レート 44.1kHz で収録される CD は、その半分の 22.05kHz が収録可能な最大周波数です。

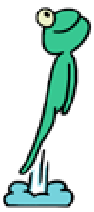
さて、サンプル・レートの半分が収録可能な最大周波数となる、その理由とは一体…百聞は一見にしかず、右の図をご覧ください。緑の線が入力されたアナログ信号波形、オレンジ色に塗ったマスが、各瞬間に信号電圧をサンプリングした値です。正弦波の一周期を表現するには、少なくとも 2 サンプルを使用して上下の往復を描く必要があることがわかります。

エイリアス・ノイズ～録音可能以上の周波数



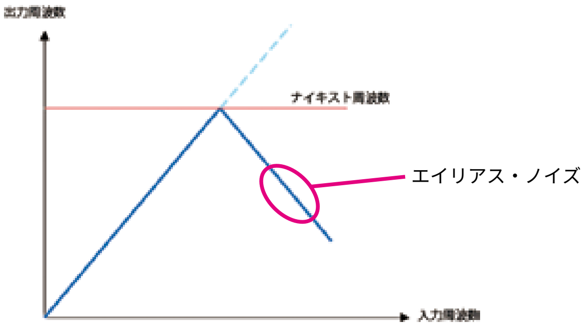
では、ナイキスト周波数以上の周波数をもつ信号を A/D コンバータに入力した場合はどうなるでしょう？ 右の図をご覧ください。

入力信号が「4 サンプルで 3 往復」しているのに対し、サンプリングする頻度が低すぎるために、信号の動きを正しく記録できません。その結果、塗りつぶしたマス目は実際の入力周波数よりもむしろ低い「4 サンプルで 1 往復」の波形を描き、元にはなかった信号がデジタル領域で捏造（ねつぞう）されています。



このように、ナイキスト周波数よりも高い周波数をもつ入力信号は、**エイリアス・ノイズ**と呼ばれる、実際よりも低い周波数の信号をデジタル領域で発生させる元となります。

エイリアス・ノイズは、右のグラフのように元の信号がナイキスト周波数を超えた分だけ低い周波数に発生することから、**折り返しノイズ**とも呼ばれます。折り返して 0Hz まで落ちたエイリアス・ノイズは、再び反射するように 0Hz とナイキスト周波数の間を往復し続けます。



実際の A/D コンバータには、このようなエイリアス・ノイズ発生を防止するためのローパス・フィルタが入っており、ナイキスト周波数以上の信号は PCM に変換される前段階でカットされるため、収録時に問題になることはありません。

しかし、プラグインを使用する場合は、エイリアス・ノイズの取扱いが重要になる場合があります。以下では、この点を説明します。

高サンプル・レートは必要か？

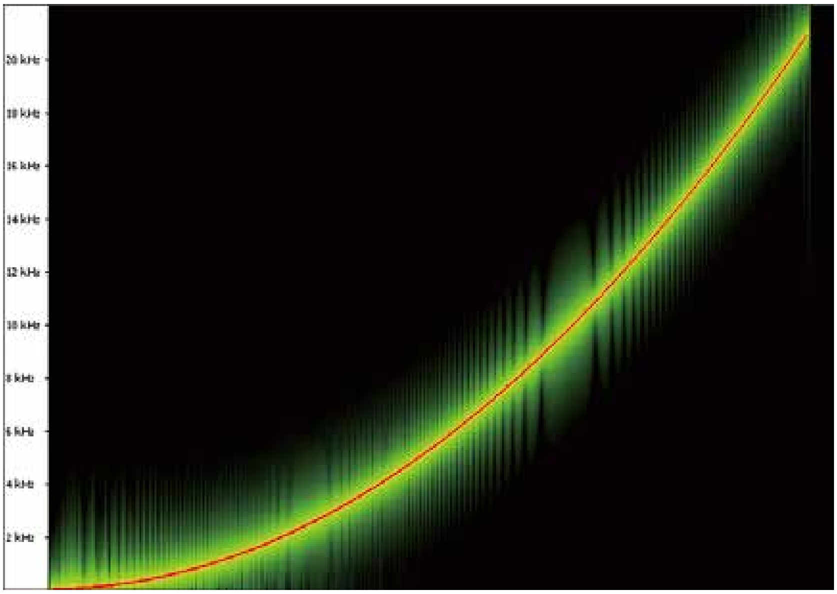
DAW にて録音や作曲を開始するにあたり、プロジェクトのサンプル・レートをどのように設定するかは悩ましいところです。前章までお話ししたビット･デプスとは違い、サンプル･レートは一旦設定すると後からの変更が困難です。一般的にサンプル・レートが高いほど、原音を損なうことなく高音質のマスターが制作できるといわれますが、代わりに作業時にはより高い PC スペックが要求されます。

しかし、最終的な媒体が 44.1kHz の CD の場合でも、高いサンプル･レートで編集を行うメリットはあるのでしょうか？また、扱うジャンルや目指すサウンドの傾向によっては、サンプル・レートの違いは無視できるのでしょうか？

結論から先にいいますと、サチュレータ、コンプレッサなど歪みや倍音が生じるエフェクトを DAW 内で使用する場合、サンプル・レートの設定が大きく影響する可能性があります。

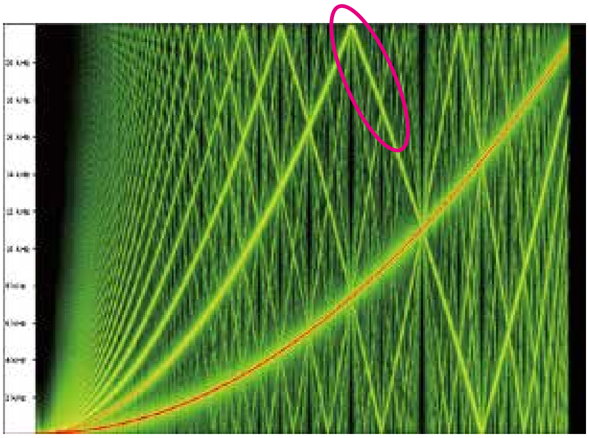
プロジェクトのサンプル・レートによってどのような違いが生じるか、ソノグラムを用いて考察してみます。これはタテ軸に周波数ごとの信号レベル、ヨコ軸に時間を表すことで、音声信号の周波数分布を可視化するものです。

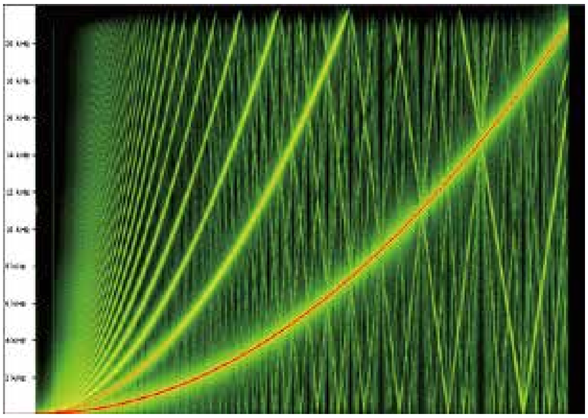
まずは単純なスィープ信号のソノグラムから見てみましょう。正弦波の周波数が、時間とともに 0Hz から 22kHz まで徐々に上がっているのがわかります。正弦波は単一の周波数のみ含む信号ですので、一本の線となって現れています。

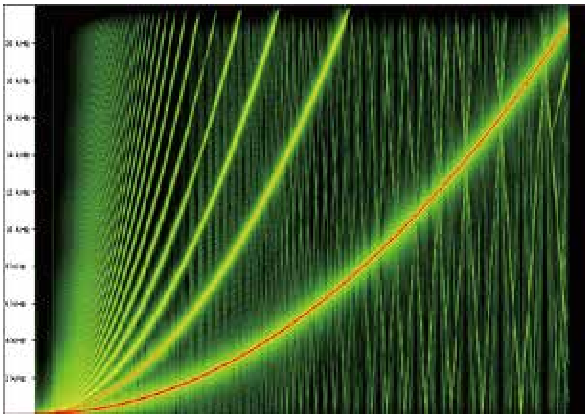


0 ～ 22kHz のスイープ信号

次に、これに歪み系のエフェクトを加えてみます。
ご覧いただくのはいずれも、サンプル･レート設定の異なるプロジェクトでエフェクトを掛けた信号を書き出したのち、ファイルを 44.1kHz に変換したものです。エフェクトは、すべて同じプリセットを使用しています。
比較するファイルは制作工程が違うだけで、それ自体はすべて 44.1kHz であることにご注目ください。


44.1kHz プロジェクトで書き出し


88.2kHz プロジェクトで書き出し


192kHz プロジェクトで書き出し

いずれのケースも元の正弦波に対して、ちょうど整数倍の周波数をもつ信号が付加されているのが見られます。この倍音に注目すると、44.1kHz のプロジェクトから書き出したファイルにおいてはナイキスト周波数である 22.05kHz にて倍音がただちに折り返しているのに対し、プロジェクトが高レートになるほど折り返しが目立たなくなっているのがわかります。ここで生じているエイリアス・ノイズは元音に対して不協和で、アナログ機器では生じえないデジタル特有の響きを持ちます。

ロックなど歪みの多い音楽ジャンルの制作段階において、音のクリーンさは必要ないからサンプル・レートは 44.1kHz や 48kHz で十分、との声をよく聞きます。これは一旦デジタル・レコーダに収録したのち、ミックスをすべてアナログ卓で行う場合は、そのとおりかも知れません。
しかし、デジタル領域（DAW 内）でミックスを完結させる場合、例えば音声に歪み（倍音）を加えるコンソール・シミュレータなどのエフェクトを全トラックに挿入すると、エイリアス・ノイズを相当量、発生させるおそれがあります。これが DAW 内ミックスの音が「デジタル臭い」といわれる原因のひとつかもしれません。この問題は、高いサンプル・レートを使用するほど軽減できるかもしれません。

マスタリング・エンジニアの中には、元の 2 MIX のサンプル・レートに関わらず、すべて一旦 192kHz のプロジェクトに取り込んでから作業を行うという方もいます。これは、サチュレータなど歪みを生じるプラグインを 2MIX に掛ける場合に、エイリアス・ノイズを最小限に抑え、音の明瞭さを保つ手段として有効であると考えられます。

先の例では、単純な正弦波ですら、あれだけのエイリアス・ノイズを発生させていました。マスタリングに際して同じようにサチュレータを使用した場合、より複雑な信号を持つ 2MIX が受ける影響は計り知れません。

続いて、他の例を見てみましょう。

Native Instruments Guitar Rig4 挿入時（プリセット「Jimi's white pleasure」）

ギター用のディストーション・エフェクトをかけた例です。（次ページ、左列）
ノイズも音作りの一部…と解釈することは可能ですが、エイリアス・ノイズを発生させることは、原音に対して周波数の相関がなく不協和な信号をデジタル領域で捏造する行為です。これは、ギター・アンプにみられるようなアナログ・ディストーションとは特性が異なります。

Native Instruments Guitar Rig など、スタンド・アロンで動作し、サンプル・レートを簡単に変更できるツールを使えば、演奏しながらサンプル・レートによる違いを手軽に聞き比べられるかもしれません。

Cakewalk VX64（Preset "Sugary Sweet"）→ IK Multimedia CSR（Preset "Thick Vocal Plate"）

「サチュレータ→リバーブ」という、ボーカル・トラックでは日常的に使用されるエフェクタの接続パターンの例です（次ページ、右列）。リバーブは、いわばディレイの集合ですので、すべての入力信号が右に向かって尾を引いているのがソノグラムにも現れています。デジタル領域で発生したエイリアス・ノイズも例外ではなく、リバーブが演出する空間を不協和音で濁しているのが見られます。

※本章の続きは本誌にて

2CAudio: Aether の設定画面。オーバー・サンプリングの有無を、リアルタイム再生時とオフライン・レンダリング時、それぞれ個別に設定できるようになっています。（詳細は 23 ページ）

4章 収録レベルの話

制作された年代の異なる楽曲を続けて再生するとき…たとえば1990年代の半ばまでにリリースされたタイトルを近年の作品と並べたとき、古いタイトルのほうが音量が小さく感じられることは、多くの方が経験されたことがあるかと思います。これは聴感上の収録レベルを簡単に大きくできる**マキシマイザ**という制作ツールが一般化した末に、ときには音質を犠牲にしてでも収録レベルを高くした方がよい、という考えが広まった結果です。また、音楽を制作する側の皆様は、市販タイトルと並べて聴いた（あるいは聴かれた）時、自分の作品の方が音が小さく感じられることに不安を抱いた経験をお持ちではないかと思います。

しかし、収録レベルを大きくすることは、音楽をリスナーに届ける上で本当に必要なのでしょうか？

日本の音楽業界やクリエイターの間ではあまり話題になっているのを聞きませんが、実はむしろ、マキシマイザで収録レベルを大きくした分だけ、そのまま後悔となって制作者に跳ね返ってくる…かもしれない、状況が整いつつあります。この章では、マキシマイザを使用した収録レベルの向上、いわゆる「音圧上げ」と呼ばれるプロセスがそもそもどんなものか、またそれが仇（あだ）となりかねない理由について説明します。

マスタリングとは

マスタリングは、ミックスと並び音楽制作における工程のひとつです。**ミックス**、あるいは**ミキシング**が、録音された楽器ごとのバランスを調整して1曲を完成させる作業なら、**マスタリング**は**1曲または複数の曲を、収録する媒体に向けて最適化する作業**のことをいいます。たとえば、収録する媒体がCDアルバムであれば、アタマから最後まで通して聴いたときに違和感が無いよう、音量、周波数帯のバランス、押し出し感（パンチ）などに一貫性を持たせるための調整を行います。元のミックスがCDよりも高いビット・デプス、サンプル・レートで収録されている場合は、44.1kHz/16bitに変換します。アルバムに収録された各楽曲が異なる環境でミックスされた、あるいはエンジニアまでもが異なる場合、先ほど挙げたバランスやパンチなどは、大抵曲ごとにバラバラです。これらに一貫性を与える上でマスタリングは重要な作業になります。マスタリングでは主に、押し出し感を調整する**コンプレッサ**や、周波数のバランスを整える**イコライザ**などを使用します。アナログLPのマスターを制作する場合はこれらに加え、再生時に針が飛ぶ原因となる低周波をカットするといった作業が生じます（これも媒体に合わせた最適化です）。近年はインターネットでの楽曲配信も盛んになり、1曲からのマスタリングが行われることも増えました。この場合は、想定される視聴環境に適したバランスにしたり[*1]、mp3などの圧縮形式に変換しても曲のメッセージが損なわれないよう音の調整が行われます。

PCで手軽に音楽制作ができるようになる以前は、制作現場におけるミキシング・エンジニアとマスタリング・エンジニアの役割は明確に分かれていました。制作ツールの入手性が高まり、クリエイターの数だけ制作スタイルが存在する現在では、ミキシングとマスタリングの境界線はぼやけつつあります[*2]。

---

*1 例えば同じ曲でも、SACDとしてハイファイな環境で再生されるか、ネット動画のように携帯電話やノートPCの内蔵スピーカで再生される想定かによって、音の調整方針は異なります。

*2 ミックスやマスタリングといった作業区分は、トータル・リコールの難しかったアナログ時代にルーツがあります。DAW内で完結するワークフローで自身のアルバムを作成するときなどは、マスタリング中にパート間のバランスが気になればミックスに戻ることも検討すべきですし、むしろその方が2ミックスを処理するよりも良好な結果が得られるケースが多いです。

マキシマイザの普及前後

さて、20 世紀も末頃になり、Brick Wall Limitter（レンガ壁のようなリミッタ）…いわゆるマキシマイザが登場すると、マスタリングの工程に**マキシマイズ**という作業を取り入れる慣習が広まりました。これは大雑把に説明すると**「瞬間的に信号が大きくなる部分だけレベルを下げて、上限までの余白が生じた分だけ全体のレベルを上げる」**加工です。
ここで、年代の異なる 2 つの CD を比べてみます。下の図 1 をご覧ください。これは 1973 年に制作されたバンドもののアルバム（以降アルバム「A」）から、冒頭 3 曲の録音信号（波形）を表示したものです。ちなみに 1973 年にはまだ CD は存在しないため、後年 CD 化されたものを使用します。

横向きに伸びた上下 2 段の帯は、左右それぞれのチャンネルの信号波形を表しています。左の曲アタマから右に向かって流れ、帯の太さが各時点での信号レベル（ここでは「音量」のことだと思ってください）を表します。この図を見るだけで、なんとなく 1 曲目の最後の方に奏者が一斉に休止する「ブレイク」がありそうなことや、2 曲目がアタマから終わりに向けて徐々に大きくなっていることがわかります。
トラックのピーク値 [*3] を測定すると、左右一方のチャンネルが -1.5dBFS まで上がる瞬間はあるものの、演奏時間の大半は -5 dBFS にすら届きません。デジタル収録可能な上限レベルである 0dBFS まで、わりと余裕があるといえます。また、曲中のセクションごとや、楽曲ごとのレベルに明確な幅があります。（たとえば大サビは一段音量が上がるなど）
次に、同じくバンドもので、2008 年にリリースされたタイトル（以降アルバム「B」）から同様に波形を表示した図 2 をご覧ください。こちらは先ほどのタイトルとは対照的に、ピークはほぼ絶えず 0dBFS に到達しています。また、波形を一見すると、曲中のセクションごとや、楽曲ごとのレベル差は見られません。ご想像のとおり、**再生機器のボリュームを固定する限りにおいては**アルバム「A」よりもアルバム「B」の方が遥かに音が大きく感じられます。

図 1　アルバム「A」より冒頭 3 トラック。1973 年リリース。

図 2　アルバム「B」より冒頭 3 トラック。2008 年リリース。

---

*3 楽曲中の信号レベルの最大値。CD をはじめとする PCM 方式では一般にゼロを収録可能な最大値とし、単位「dBFS」で表します。（詳細は p.42）

Chapter 4　収録レベルの話

音圧戦争 Loudness War の歴史

CD が普及し始めた 1980 年代初頭、マスタリング時に可能な限り高い信号レベルを CD に収めるという概念はなく、注意が払われていたのは、アナログ信号をデジタルに変換する際にレベルがオーバー・ロードしない（0dBFS を超えない）ようにすることだけでした [*4]。また CD の収録レベルについても、これといった基準もありませんでした [*5]。

時は流れて 90 年代後半、ここに**マキシマイザ**という、これまでになかったタイプの制作ツールが誕生します。

マキシマイザとは、入力されたオーディオ信号が、任意に設定したレベル（＝スレッショルド）を絶対に超えることがないよう加工するツールです。スレッショルドを超えた信号はきわめて自然にレベルが下げられ、逆にスレッショルドを超えない信号は、一切劣化することなく通過します。これにより、スレッショルドを 0 dBFS 付近に設定した上で入力信号のレベルを上げていくことで、全体の収録レベルを大きく上げることができるようになりました。この効果は従来のアナログ機器では原理上えられないものでしたが、デジタル信号処理により初めて可能になりました。こういったツールは開発当初、オーバー・ロードを回避するための保険…つまり意図せず 0dBFS を超える信号に対処するために活用されていましたが、やがて収録レベルを上げることを目的に積極的に使用されるようになります。

アルバム「A」が発売された頃、デジタル収録レベルの上限である 0dBFS は、積極的に近付く必要はなく、まれに近付いてしまったときには絶対に超えてはならない『境界線』のようなものとみなされていました。それがやがて、いかに長時間、張り付いていられるかを競う『天井』として認識されるようになったのです。

そもそも、マキシマイザが作用する信号の大部分を占める「速く鋭いアタック」は、アナログ・レコーダでは録音することが難しい部類の信号で、それを再現できることが CD をはじめとするデジタル録音の強みであったはずです。その恩恵を放棄してまで、収録レベルを大きくする慣習が広まったのはなぜでしょう？その辺りの事情はプロジェクトにもよりけりでしょうが、おおよそ以下に分類されるかと思います。

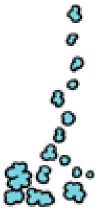
1. 人間の聴覚は、相対的に大きい方の音を『良い音』と感じる性質があるため

同じ音源を DAW の複数トラックにコピーして並べ、一方を 1dB 下げ、2 つを交互に切り替えながらどちらが良く聴こえるか身近な人に尋ねてみてください。おそらく多くの人が、大きい方をよく感じると答えるでしょう。再生機器のボリュームを固定する限りにおいては、マキシマイズ後の音の方が派手に聴こえるのは、大きい方の音が良い音と評価されがちなためです。

2. ラジオの乗りをよくするため

上記のフレーズは、2 つのまったく異なる意味を持ちます。 ひとつ目は先述のとおり、大きい音の方が相対的に派手に聞こえるため、楽曲がオンエアされればザッピングしているリスナーを足止めしたり、CD セールスに結び付ける上で有利と考えられました。

ただし、後述するようにこれは誤った認識であるどころか、多くの場合マイナスに作用します。ふたつ目は、ラジオ番組の選曲者の注意を引きやすいという意味です。選曲者がオンエアする曲を選ぶ目的で CD を順に聴いているとき（再生機のボリュームを変更しないという前提で）大きく聴こえた方が相対的にはインパクトがあるため、「おっ、この CD アガるねぇ、今日の放送ではこれをかけよう」という効果を狙って…という考えです。



3. 他のタイトルより小さく聴こえる（＝派手に聴こえない）と、商売上不利と思われていたため

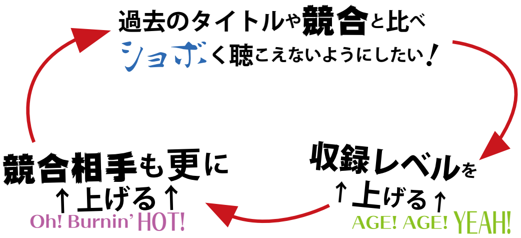
おそらく、マキシマイズが施されたタイトルのほとんどは、この理由が背景にあると思われます。特に似たジャンルで比較対象となる他者の作品が、自分よりも収録音が大きければナーバスになるのは解らなくもありません。

---

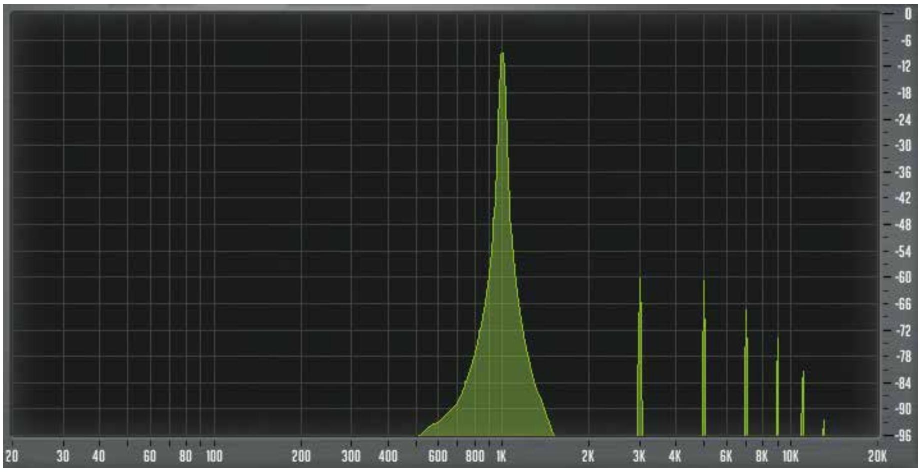
*4 デジタル・レコーダで収録可能な最大レベルを超えて入力された信号は、耳障りな歪み（＝ノイズ）になります。

*5 2019 年現在も CD の収録レベルに基準はありません。歴史的に見てもレベルの管理に厳しい放送業界とは対照的といえます。なお、業界標準や統一規格はないものの、たとえば「インター・サンプル・ピーク」の存在が広く知れ渡った結果、ピークを -0.2 dBFS 前後に設定するといった「慣習」のようなものはいくつかあるといえます。

こうしてクリエイターや音楽関係者が互いに聴感上の音量を競い合った結果、CD の収録レベルは際限なく上昇を続けました。さきほど「マキシマイザが違和感なくピークを抑えられる」と述べましたが、これはあくまでもちょっとしたオーバー・ロードを抑制する程度に使用する場合です。競争が進むほどに市販タイトルの多くはアタック感を失い、歪みを増し続けました。



ピークを抑えるというのは、基本的に正弦波を矩形波に近づける作業でもあります。下図は 1kHz の正弦波をマキシマイザに通したあとのスペクトラムです。矩形波の特徴である奇数倍音が可聴域に現れているのがわかります。



また、体感音量を上げるべく従来のコンプレッサも投入した結果、**ポンピング**という現象により、旧盤では表拍にあったビートが裏拍に回ってしまった著名曲のハイレゾ・リマスター版なども存在が確認されています。

このように音質やアレンジを犠牲にしてまで収録レベルを競う動きは、いつしか**音圧戦争**、**音圧競争**、また海外では **Loudness War** などと呼ばれるようになりました。興味深いことに、このように収録レベルを競う動きは、CD 以外のデジタル媒体においてはさほど顕著ではありませんでした [*6]。

やがて、CD ばかり競争が激化していたことを示す、ある象徴的な出来事が 2008 年に起こります。
Metallicaのアルバム「DeathMagnetc」が、過度に掛けられたマキシマイズのため躍動感に乏しく、録音も歪みだらけの状態でリリースされました。これをきっかけに、著名バンドによる新作がかくも悲惨なことになった経緯…つまり音圧戦争が招いた惨状について、音楽を専門としない多くの一般誌さえもが一斉に取り上げたのです [*7]。
その後、同アルバムの収録曲が、いわゆる音ゲーである「Guitar Hero Ⅲ」用のダウンロード・コンテンツとして提供されるようになりました。ゲーム業界においても収録レベルについての統一基準はありませんが、多くのメーカがそれぞれの社内基準を有していることもあってか、同じ「Death Magnetic」の収録曲も比較的余裕のある（＝過度にマキシマイザをかけない）状態でリリースされたのです。かくして、本来は音楽専用のメディアであったはずの CD が、ゲーム中で使用された同一楽曲に品質で劣るという奇妙な現象が生じてしまったのです。

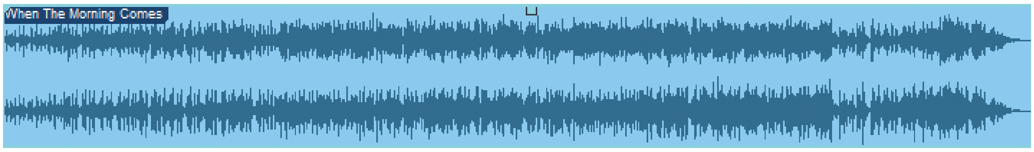
※本章の続きは本誌にて

5章 メータの話

音声のレベルを測定する方法はいくつかあり、いずれも計測手段や目的が異なるため、それぞれに一長一短があります。DAW ユーザにはおなじみのピーク・メータを例にとってみましょう。

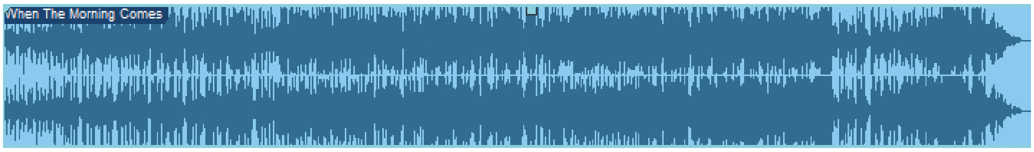
ピーク・メータは、ADC に入力された信号や DAC から出力される信号が、それぞれの扱える最大値までどの程度の余裕（マージン）があるかをバーや数値で表します。これは ADC/DAC の過大な入出力により信号が歪むことを未然に防ぐのに役立ちます。
逆に、**ピーク・メータが示す数値は、聴感上の音量を表す上では一切役に立たない**という短所は、普段あまり語られません。

たとえば、すでに 0 dBFS にピーク・ノーマライズされたファイルがあるとします。



再生すると、ピーク・メータは曲中のどこかで最大値 0 dBFS を指すでしょう。

このトラックにマキシマイザを挿入し、6dB 持ち上げると下図のようになります。



再生中にピーク・メータが表示する最大値は、やはり同じ 0 dBFS ですが、どちらの方がより大きく聞こえるかは、いうまでもありません。

本章では、ピーク・メータやその他各種メータの違いや主な役割を紹介します。

結局、最終的には耳での判断がすべてになりますが、たとえば長時間のミックス作業を行うときや自身のミックスをリファレンス・トラックと比較するときなど、再生音量の差に騙されず [*1] に正しく比較する上でメータは有効なツールとなります。ときには目標とするサウンドに効率よく近づくための武器にもなるでしょう。

*1 他の章でも度々述べたよう、ヒトの聴覚はえてして音量が大きい方を派手で「良い」と感じるものです。EQ やダイナミクス処理を行う場合、それが良い方向に作用しているかどうかを正しく評価するには、処理前後の聴感レベルを揃えて比較する必要があります。ボリューム設定が同じでも、比較する一方が大きく聴こえる状況では正しい判断はできません。

ピーク・メータ

ピーク・メータのおおよその性質は、先の導入部で説明したとおりです。

- **長所：ADC/DAC が扱える最大値（0 dBFS）までのマージンを知ることができる**
- **短所：ほとんどの場合、聴感上の音量を知る上で参考にはならない**

たとえば、ADC の最大入力レベルより 6dB 低いレベルの信号に対しては、-6 dBFS と表示されます。
ピーク・メータは本章で紹介する中では唯一、**1 サンプルの信号レベル**を表示します。とはいえ、瞬間瞬間のレベルを表示していたのではメータの動きが速すぎて、目視では追いかけることができません [*2]。そのため、通常は一定時間内に生じた**最大値のみを表示**します。

ところで先ほど、ピーク・メータは「ADC/DAC が扱える信号レベルの最大値を基準にする」と述べましたが、「ファイルへの記録」時は、それよりも大きな信号も扱えます。ただし、正しく工程を管理すればですが。詳細については 1 章「ビット・デプスの話」をご覧ください。

アナログ・コンソールやテープ・レコーダの時代にも、機器ごとに推奨される基準レベルはありましたが「一定のレベルを 4.41 万分の 1 秒でも超えると、確実に信号が歪む」というような具体的なレベルはありませんでした。ピーク・メータは、レベルの絶対的な上限値という概念をもつデジタル・レコーダの登場とともに生まれました。

※本章の続きは本誌にて

コラム：単位系について

ここで少しだけ、単位系のお話をします。
0 dBFS は、0 dB **フル・スケール**の略で、ADC/DAC が扱える最大値を表します。これに対し「dB」は、数値間の**相対的な差を表す**単位で、デジタル、アナログを問わず、また音楽にとどまらず電波強度の表示などにも広く使用されています。
以下の用例を、○×とあわせてご覧ください。

○：レベルを 3dB 下げる。←基準（元位置）からの相対的な差
○：マキシマイザのスレッショルドを -3 dBFS に設定する。←絶対的な値
×：このファイルのピーク・レベルは -0.5 dB。←絶対的な値なので dBFS とすべき
×：ファイルを -1 dB でノーマライズする。←絶対的な値なので dBFS とすべき

上記に加え dB SPL[*3] など他の単位も「dB」と一括りにすることは慣習上よくありますし、会話の当事者同士で意思の疎通ができていればよい話ですので、細かい使い分けにまで文句をつけるつもりはありません。しかし、まれに単位の混同が、そのまま DAW の仕組みに対する誤った認識に繋がっている場面を目にします。後々誤解が生じないよう、最低限、本コラムの内容は理解しておくとよいかと思います。

---

*2 たとえば 48kHz で収録を行う場合、毎秒 48,000 回、信号レベルを測定します。

*3 端的にいうと、空気振動の物理的な強さです。収録レベルが一定でも、再生機のボリューム位置によって dB SPL は上下します。

6章 ダイナミック・レンジの話

ダイナミック・レンジという語は、文脈によってまったく異なる使われ方をします。 よく耳にするわりには、じつは大変まぎらわしく、混乱や誤った認識の元となる語句でもあります。主に、以下の３つの意味で使用されることが多いかと思います。

① 曲中において、音の大きい箇所と小さい箇所の差

これは観念として用いられることが多く、通常は具体的な数値では表しません。

例：クラシックの録音について「pp から ff までのダイナミック・レンジが広い」など

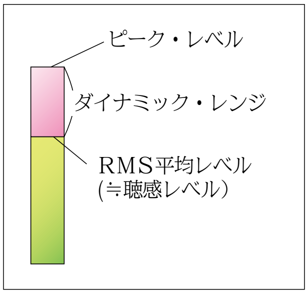
② 機器などが取り扱える信号レベルの、最小値と最大値の比率

主に、機器やシステムのスペックを語るときに使用されます。

例：「16bit 録音のダイナミック・レンジは 96dB」など

③ RMS 平均レベルとピークの差（クレスト・ファクター）

上記③「信号の RMS 平均レベルとピークの差」を図示すると、右のようになります。**ピーク・レベル**は、信号レベルの瞬間的な最大値を表します。ただ、意外なことに瞬間的な最大値は、聴感上のレベルとはあまり関連がありません。では、聴感上のレベルはおおよそ何で決まるかというと、一定時間の信号レベルの平均をとった、**RMS 平均レベル**です。この２つの値の差も「ダイナミック・レンジ」と呼ばれます。



以下、この章では③「RMS 平均レベルとピークの差」に注目し、なぜこれが重要かについて、またミックスやマスタリングを行う上でダイナミック・レンジを考慮すべき場面を説明します。

ダイナミック・レンジが大きいとどうなるか？

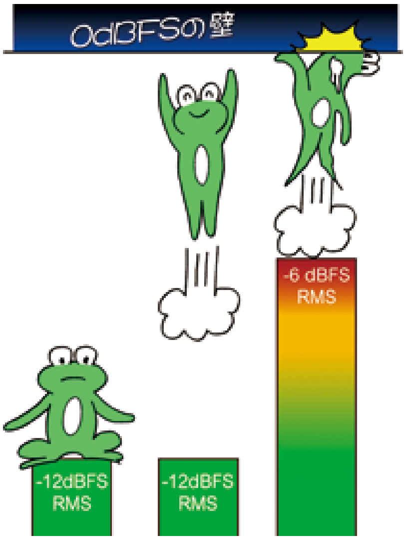
ダイナミック・レンジが大きい楽曲は、それより小さい楽曲に比べ１つ１つのアタックが強く感じられます。また、パートごとのアタックに個性を持たせやすいため、パート間の分離がよくなります。

キック（バス・ドラム）のレベルを例に考えてみましょう。
オケ全体の RMS 平均レベルに対し、アタックが大きい（ダイナミック・レンジが広い）キックは、それだけピークが他のパートよりも突出します。逆に突出しない（ダイナミック・レンジが狭い）キックは押し出し感が弱く、場合によっては壁に阻まれたような、つまったような印象を与えます。しかしダイナミック・レンジに関係なく、RMS 平均レベルが同一であれば、全体としての音量感自体はさほど変わりません。

ダイナミック・レンジをさらに狭めると、キックの音が**耳に聴こえるが、スピーカーやヘッドホンからは押し出し感がまったく感じられない**状況が生じます。

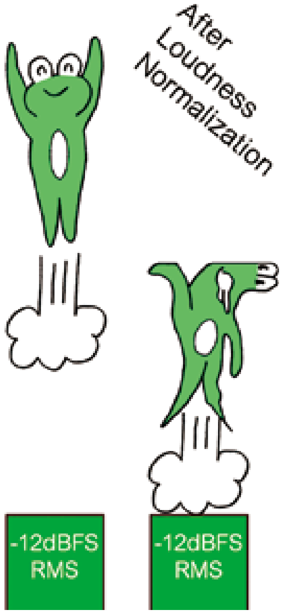
ダイナミック・レンジと RMS 平均レベルの関係

ここで、デジタル媒体に音声を収録する際の RMS 平均レベルと、表現可能な最大ダイナミック・レンジの関係について考えてみましょう。次の図をご覧ください。



DAW や CD などのデジタル機器に録音する際には、0dBFS が収録可能な最大レベルになります。左図、中央のカエル君にご注目ください。このカエルは、収録時に平均レベルを -12dBFS RMS に設定しました。つまり、収録可能な最大レベルである 0dBFS まで、12dB の余裕があることになります。次に、右端のカエル君にご注目ください。彼は中央のカエル君よりも自分の音を大きく聞かせるべく、Wavez 社製の最新マキシマイザ「L5-512」を使って -6dBFS RMS で楽曲を CD に収録しました。この場合、0dBFS まで 6dB の余裕がありますので、ドラムなど、アタックの強いパートの表現に使える余白は 6dB ということになります。なにせ最新のマキシマイザを使っているので、ピークを抑え込んだことによる歪みは皆無です。この 2 つを比較すると大きく聴こえるのは、もちろん RMS 平均レベルの高い右端のカエル君の方ですが、楽曲中の平均レベルに対してドラムなど瞬間的にレベルが突出するパートの表現に使える範囲が狭くならざるをえないことは、この図からも明らかですね。

さて、ここで**ラウドネス・ノーマライゼーション**という概念が登場します。
これは、複数の楽曲やその他のコンテンツを、信号レベルではなく**聴感上のレベルで揃える**仕組みです。先ほど説明したように聴感上のレベルはおおよそ RMS 平均レベルで決まるので、２匹のカエル君の楽曲を、RMS 平均レベルが揃うようそれぞれのレベルを調整します。



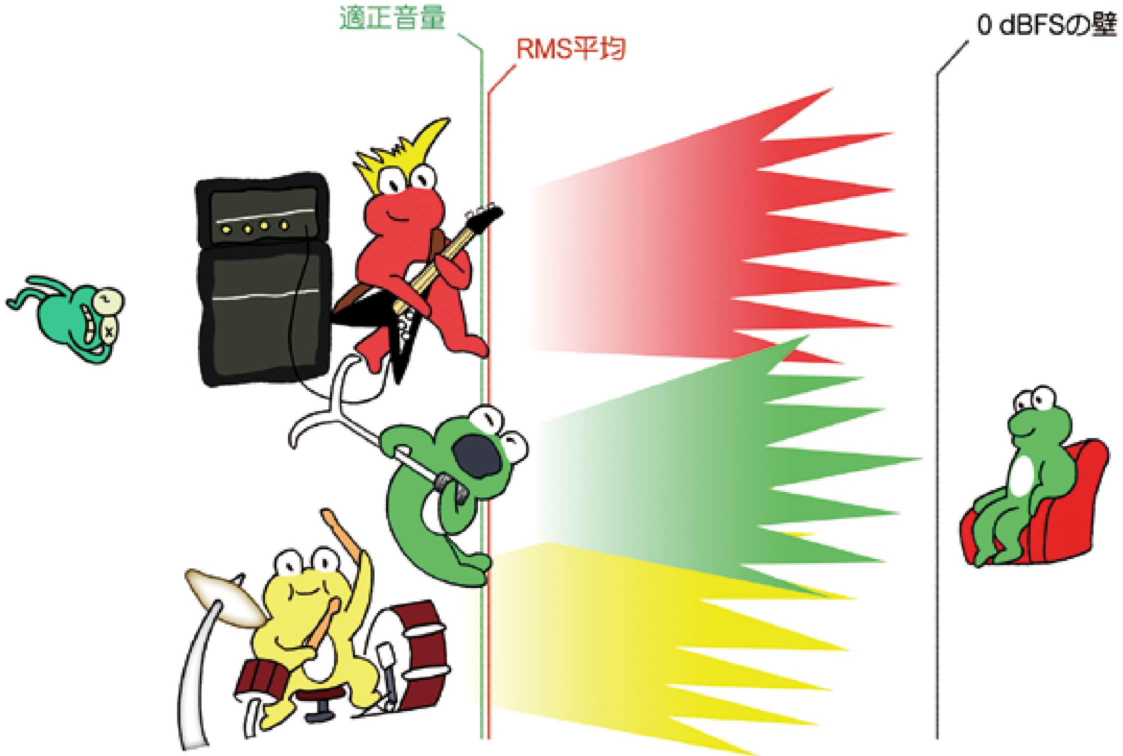
右のカエル君の方はダイナミック・レンジが小さいため、飛び出す勢いを欠いたようなかたちになってしまいました。まるで見えない天井があり、そこから上には進めないかのようです。
４章「収録レベルの話」でも述べたように、再生時の音量を聴感上のレベルで揃えるという動作（まさにラウドネス・ノーマライゼーション）は、再生機を操作するリスナーが各自で、あるいはテレビ放送やネット配信サービスの多くが厳密な基準により機械的に行っています。
「見えない天井にぶつかったカエル」サウンドは、すでに現実の街中に溢れているのです。

マキシマイズのプロセス

先の２匹のカエル君の例と同じ概念を、別のアプローチからもう少し詳細に考察します。

今度は別のカエル…ここではリスナーガエル君（下図、右端）が、自室のオーディオ機器で、バンド演奏の CD を聴いているところを見てみましょう。自慢のハイファイ・システムを構成するのは「ソ○ーは、ハイレゾ」のコピーでおなじみ、ソ○ー社製の最新 DAC。そして「ハイレゾはパナ○ニック」のコピーでおなじみの某社…の系列にあるＲＡ○ＳＡのハイエンド・スピーカ。再生ボタンを押してソファに深く腰掛けると、まるでバンドが目の前で演奏しているような臨場感溢れる音が部屋に響き渡ります。

ここでは、信号レベルは左から右に向かって大きくなります。
前項で述べたように、「RMS 平均レベル」はおおよその聴感レベルに対応します。現在、アンプのボリュームはリスナーガエル君が**最適と感じる位置に自分で合わせた**ため、RMS 平均レベルと、リスナーガエル君が最適と感じる再生ボリューム（適正音量）の２本の線はイコールです。



次に、マキシマイザやコンプレッサでピークを抑え、ダイナミック・レンジを狭め、収録レベルの RMS 平均レベルを上げます [*1]。聴感上のレベルが増し、演奏しているバンドがグイっとリスナーに近付きました。少なくとも再生機のボリューム位置を変えない限りにおいては、先ほどより随分とラウドになりましたね。マキシマイザでレベルを持ち上げた状態がこれに相当します。

しかし、ここで 1 つ問題が発生しました。
先に述べたように、**聴感上の音量は、ピーク・レベルではなく、おおよそ RMS 平均レベルによって決まります。**
おそらく多くの制作現場で 2 ミクスがより ラウドに聴こえるようにと、マスタリング時にマキシマイザが掛けられていると思われますが、RMS 平均レベルはリスナーガエル君が最適と感じる視聴レベルを大きく超えてしまいました。
当然、リスナーガエル君はボリュームを下げます。

すると…

※本章の続きは本誌にて

*1 リスナーカエル君が自慢のリスニング・ルームにマキシマイザを常備している変わり者…ではなく、制作の過程においてマキシマイズ処理される前後の楽曲を比較している、という設定でご勘弁ください。

7章 M/S 処理の話

近年、主にマスタリングの文脈で **M/S 処理**という技法をよく耳にするようになりました。技術的には特に目新しいものではなく、万能なテクニックでもないのですが、各社が相次いでプラグイン･エフェクトに実装した結果、副作用があまり意識されないまま使用されることが増えたのではないかと筆者は考えます。効果的な使用方法については有用な資料がたくさん世に出ていますのでそちらに譲り、本章では以下に要点を絞って解説します。

前半「M/S 処理とは何か？」

まず M/S 処理の概要を説明します。M/S 処理が別段魔法のような処理ではないことをご理解いただきたく、制作プロセスに興味をお持ちのリスナーを主な対象に書きました。せっかくなのでクリエイターにとっても再考の機会になればと、他の文献とは異なるアプローチでの解説を試みました。

後半「M/S 処理時の注意点」

M/S 処理による副作用、特にステレオ音像を広げた場合の留意点について、主にクリエイターを対象に説明します。ご自身では制作を行わないリスナーの方がお読みになっても、疑問に思っていた「アノ 1 枚」の事情を理解するヒントが見つかるかも知れません。

M/S 処理とは何か？

音楽収録やトータル・エフェクトの主流は 2 トラック

カセット・テープ、CD、LP、etc.. これらはすべて、基本的に 2 つの収録トラックを持つ（複数の独立した音声を同時収録できる）媒体です。各音声トラックには通し番号が振られ、一般にトラック 1 には左（L=Left）のスピーカから再生したい信号を、トラック 2 には右（R=Right）のスピーカから再生したい信号を、互いに独立して収録します。

また、上記のような L/R から成るステレオ 2ch への処理を目的とするエフェクタ、あるいは家庭用の再生機に内蔵されているイコライザなどは、多くの場合 L/R の 2 つのチャンネルに等しく作用します。
さて、ここまでは当たり前のような話ですが…

L/R だけが 2 チャンネルの音源を収録／加工する方法ではない

録音機やエフェクタの入出力が 2 トラック分あるからといって、必ずしもステレオ音声を L/R に分けて取り扱う必要はありません。たとえば、制作現場で使われるエフェクタには、L/R のチャンネルを個別に加工できるものもあります。これを **L/R の個別処理に用途を限定しない、2 チャンネルの汎用プロセッサ**と考えると、可能性が広がります。

エフェクターの入出力が 2 チャンネル分あるからといって、L/R の個別処理に用途を限定する必要はありません。

2 チャンネルあれば、L/R の信号を「センター寄りの成分」「両サイド寄りの成分」の 2 つに変換して収録したり、加工することもできます。
「あれ？それって L + R + センターの 3 トラックが必要でないの？」と思われるかも知れませんが､そうでもありません。その理由を説明する前に、そもそもセンターから聴こえる音とはなにかについて考えてみましょう。

「センターから聴こえる」と感じる音

同一の信号が L/R 両チャンネルから等しい音量で再生されるとき、ヒトの聴覚はそれを正面から来る音と認識します。逆に左右の耳に入ってくる音量に偏りがあると、音が大きい方向に音源があるように知覚します。いわゆる PAN は、同一の信号が左右のチャンネルから出力される音量の割合を変化させることで、正面以外の方向から来る音を擬似的に演出する仕組みであるといえます。

また、ステレオ･スピーカの正面に座ったとき、センターの音が発生していると知覚する方向にスピーカはありません。このように、スピーカが存在しない正面中央から聴こえる音声、またはその現象を**ファンタム・センター**と呼びます（Phantom= 幻、錯覚、実態のない、etc..)。
考えてみれば、音楽にとって最も重要なパートの多くが位置するセンターの再現が、このような錯覚に委ねられているのは不思議な話ではないでしょうか。

Left/Right の信号を Mid/Side に変換する

さて、このことを考えると、**センター寄りの信号**（Mid）と **L/R のいずれかに寄った信号**（Side）は、それぞれ以下の式で取り出せそうです。

Mid = L + R　→ L/R に共通する信号を強調
Side = L - R　→ L/Rに共通する信号をキャンセルし､共通しない信号を1トラックにまとめる(ただしRは逆相)

たとえば、トラックの PAN を時計の 2 〜 3 時の方向に回すと、同じ信号が L よりも R の方に若干高いレベルで含まれることになります。このことからもわかるように、Mid といえども「センターど真ん中」の音声しか含まないわけではありません。

上記で Mid、Side に変換した信号を再び L、R に戻すには、先ほどとは逆の計算を行います。

L = (Mid + Side) ÷ 2
R = (Mid - Side) ÷ 2

R チャンネルに見られる信号の引き算は、ミキサーや DAW において φ で表される**フェーズ（位相反転）スイッチ**で行います。

なお､仮に信号劣化がない回路があった場合､この L/R ⇔ M/S 間の相互変換は､何度繰り返しても元の信号に戻ります。また、特に M/S 処理に対応したエフェクタでなくても、チャンネルごとの操作が可能なあらゆるステレオ・プラグインで、あるいは十分に精度が高ければアナログ機器でも M/S 処理は可能です。

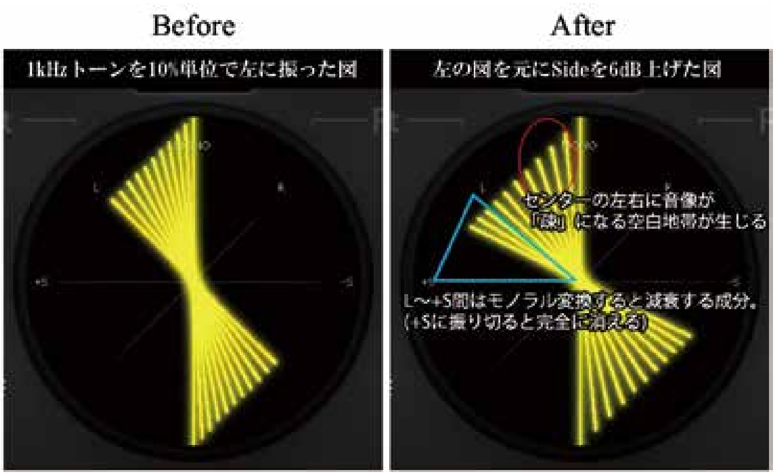
M/S 処理 = 変換して取り出した Mid/Side に対して個別に処理を行う

遠回りになりましたが、こうして取り出した「Mid 寄りの信号」「Side 寄りの信号」それぞれ個別に EQ やダイナミクス処理を行おうというのが、M/S 処理の基本です。
ただし、M/S 処理には副作用が伴います。この副作用は、M/S 処理を用いる代わりにミックス・ダウン中に各パートに対して別の手段で対処するか、あるいは場合によってはアレンジにまで立ち返ることで回避できます。
一方マスタリングの場合、すでに L/R の 2 チャンネルに整理された素材、あるいはよくてもステム（演奏セクションごとにまとめられた素材）しかエンジニアの手元にはありません。必然的に M/S 処理は、個別パートの操作が難しいマスタリングの文脈で語られることが多くなります。

M/S 処理時の注意点

ステレオ音像拡張の副作用

2MIX に対し、特に左右の広がりを強調するために Side を持ち上げる目的で M/S 処理を行う場合、留意すべき点があります。まずは下の図をご覧ください。
図中のフェーズ･メータ（リサージュ、リサジューともいう）は、ステレオ音声の左右の広がり具合を可視化するものです。



正面から音声が聴こえる、つまり左右チャンネルの信号が同一の場合、メータの中央に垂直な線が表示されます。
左右チャンネルの信号が同一ながら、レベルのみが違う場合…たとえば前述の PAN で左右に振った場合などは、L/R の音量差に応じてメータの線は傾き、L/R いずれかに振り切ると 45 度になります。この仕組みによって、**フェーズ・メータに描かれる図形を見ることで、ある程度は音像の広がり具合を目視確認**したり、曲間で比較することができます。
なお、ステレオ収録された楽曲では左右チャンネルの収録内容が均一でない場合がほとんどであるため、上図のように直線が描かれることは滅多にありません。フリーのプラグインもたくさんありますので、制作環境をお持ちの方は、まずは適当なものを入手し、いろいろな音源を通して動作を体感してみるのが理解への近道になると思います。

さて、左ページの図「Before」は、1kHz のサイン波をフェーズ・メータに入力し、PAN をセンターから L に向かって 10% ごとに振り、計 11 回の測定結果を画像編集ソフトで重ねたものです。垂直の線がセンターに位置するサイン波、45 度に傾きメータの L を指しているのが、PAN を左に振り切ったサイン波です。11 回の計測結果は等間隔です。

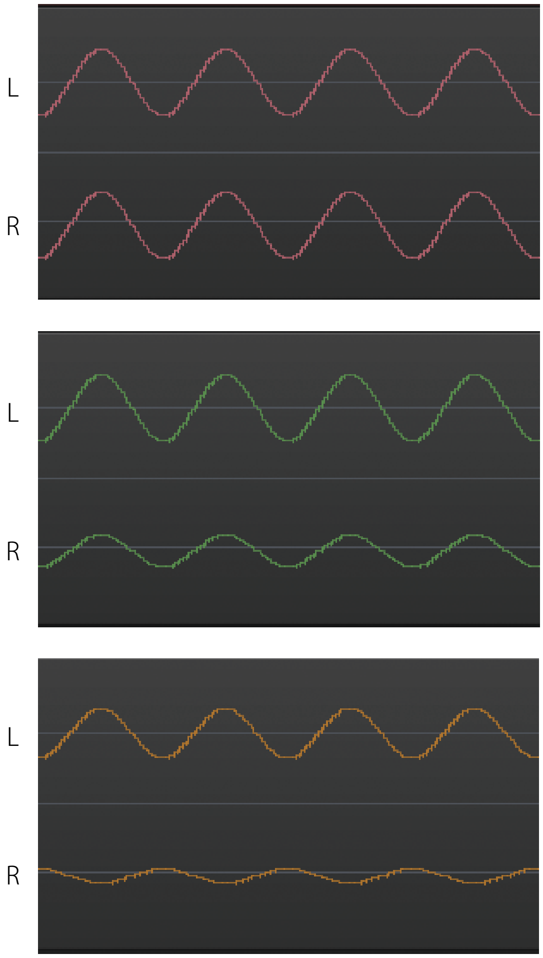
次に、この 11 本のサイン波を同様に PAN で 10% 間隔に配置し、M/S 処理により Side を 6dB 持ち上げた結果が「After」です。この 2 つの図を比べると、以下の 2 点に気が付きます。

1．**M/S 処理前は信号が存在しなかった「+S ～ L」の範囲に信号が生じる。**（水色の三角形で囲った範囲）
2．**処理後のセンター付近の分布が、処理前よりも「まばら」になっている。**（赤色の円で囲った範囲）

それぞれの現象について、以下に説明します。

副作用 1：L/R 両端に近いパートほど、逆相成分が生じる

ステレオ・メータの「+S ～ L」の範囲にある信号は、左右チャンネル間に**逆相成分**があることを意味します。
次の各信号波形をご覧ください。



1 段目（ピンク）：PAN 0（センター）設定時の正弦波の信号波形です。L/R（上下）チャンネルの波形が等しいです。

2 段目（緑）：PAN を左に 50% まで振ったときの波形です。L/R のレベル差はありますが、位相（山 or 谷の向き）は、揃っています。

3 段目（オレンジ）：ステレオ・メータの「+S ～ L」の範囲に位置する信号の波形です。左右チャンネルが逆相になっている…つまり、波形の形状は同じでありながら、波形の山と谷が反転していることがわかります。

先述のようにミキサーの「PAN」は個別のパートを左右に配置する仕組みですが、動作自体は単純な音量操作ですので、これにより L/R チャンネル間で位相が反転することはありません。
では、同一の動きをする信号が＋ / －反転した状態で L/R チャンネルに含まれると、なにか問題があるのでしょうか？たとえば作品を飲食店の天井スピーカで再生、または AM ラジオで放送することを目的にモノラル変換する場合、通常は単純な L/R 両チャンネルの足し算が行われます。すると両チャンネルの逆相成分はたがいに打ち消しあい、結果として他のパートよりもレベルが下がることになります。さらに、信号がステレオ・メータの「+S/-S」のいずれかに振り切っている場合、そのパートは完全に消滅します。

※本章の続きは本誌にて

１０章 補正可能時代の収録ボーカリストが知るべきこと

今世紀初頭より、ボーカル演奏のアラを修正するためのツールが音楽制作の現場で常用されるようになりました。以降、これらのツールは進化を続け、いまや多くの DAW に標準で付属されています。

演奏に関わる要素の中には、これらツールで修正できるものと、そうでないものがあります。どうせスキル不足や収録にかけられる時間の制約から完璧な演奏が望めないのであれば、編集では補正できないポイントを重点的に練習し、また演奏時には確実に押えて、編集で直せるものは直す…と開きなおるのも、よりよい録音物を作るための有効な手段かもしれません。

また、これは演奏者だけでなく編集サイドにも重要なことですが、逆にそれらのポイントを考慮せずに正確なピッチやタイミングに固執した補正を行った結果、後述するように現実には相容れない要素がトラック内に混在し、聴く人が聴けば（悪い意味で）人間離れした演奏に聴こえることもあります。実際、有線放送でもそのような歌をよく耳にします。以下、このように聴く人に違和感を感じさせかねないポイントを、**不気味の谷への入り口**として紹介します。

「不気味の谷」は元々はロボット工学の分野に用いられた用語で、要するにロボットが人間に近付けば近付くほど、本物の人間とのわずかな違いが見る者に違和感を与える、というものです。近年テレビでもよく紹介される女性を模したロボットや、CG で描かれた仮想空間のキャラクターにこのような印象を抱いたことのある方もおられるかと思います。これと同じことが歌声にも起こると、私は考えます。極端に人間離れした演奏は、ギミックと割り切って聴くことができますが、むしろちょっとだけ不自然な方が奇妙に聴こえるものです。個人的には、メジャー・デビュー直後の Perfume くらい思い切った加工は「ロボット・ボイス」として楽しむことができましたが、その後に出てきた二番煎じや三番煎じの中には、本家に遠慮したのか、あるいは様式だけ模してポイントを押さえられなかったのか、加工が中途半端であったために単なる奇妙な声として不気味の谷に転落したケースもあったように思います。
以下は、編集する側の方にとっては当たり前、あるいは「言われてみれば…」といった話ばかりかもしれません。エンジニアリングは解らないけど演奏や収録を楽しんでいるという方には新鮮な話もあるのではないかと思います。

編集で補正できるもの

音程

いじる要素としてはド定番です。どの程度の範囲まで自然に音程を変えられるかは、声質や歌唱法によって異なるため一概にはいえませんが、意外と大きな音程移動が可能なケースもあります。一般に、澄んだ声ほど自然に動かせる範囲が広く、逆にノイズ成分の多いウィスパー・ボイスやしゃがれ声などは可動域が狭いように思います（非倍音成分の扱いに困るのは、どの DSP 技術も同じなのかなぁという印象を受けます）。

タイミング

発音タイミングだけでなく、グルーヴ / ノリもある程度は修正可能です。例えば、モッサリした演奏を、母音と子音、それぞれの長さを個別に操作することで軽快に聴かせることができます。自然に行える範囲はけっこう限られますが…

音量の大小

単純な音量の大小です。これはボーカル補正専用のツールで行う修正だけでなく、ミックス時のフェーダによるレベル操作も含みます。後述する「声量感の強弱」とはまったく異なるので注意が必要です。

ビブラートの振幅

すでに収録されたビブラートの振幅は、ある程度自然に増減できます。
極端に広いビブラートを自然なまま平らに均（なら）したり、逆に無から作り出すことは困難です。

編集で補正できないもの

滑舌（かつぜつ）の悪さ

滑舌の悪さは、どうにもなりません。
EQ、ボリューム調整などで埋もれた音を前面に出すことはできますが、これは明瞭さや力強さとは異なります。
ヘニャヘニャした声はどう料理してもパーカッシヴにならないことを思えば、発音が悪いために失われるグルーヴ感はタイミングをかっちり合わせるなどしたところで取り戻せないことは想像に難くないかと思います。

不気味の谷への入り口

滑舌の悪さは、それだけで歌手の技量の低さを聴者に知覚させます。その一方でタイミングが正確だったりすると矛盾する要素が同居することになり、無意識のレベルではリスナーに対して奇妙さを感じさせ、また意識的なレベルで「あぁ編集しているなァ」と思わせます。

声量感の強弱

音量ではなく、声の張り具合やノドの開き具合など、いわば音が持つ「テンション」に繋がるニュアンスの強弱です。
ボソボソした声は、煮ても焼いてもシャウトにはなりません。その逆もしかり、です。

また、これが不自然に変化する（必然性なく変わる）と旋律がスムーズに流れず、場合によっては一本調子の歌よりも下手に聴こえます。

不気味の谷への入り口

悪い滑舌と正確なタイミングが同居する先の例と同様、一本調子、あるいは必然性なく声の張りが変化する演奏は、それ自体が「そもそも歌が達者ではない」ことを聞き手に知覚させるヒントになってしまいます。そのわりに音程や発音のタイミングだけは補正の結果バッチリ合っていたりすると、いかにも修正したような歌になります。
個人的にはこれを「生歌がボカロ以上にボカロっぽくなる瞬間」と呼んでいます。

ここの理解が足りず損をする例として、たびたび耳にしているのが以下のようなケースです。

1．遥か遠くで歌っているコーラスを演出しようとして、近接するマイクの前で小声で歌う。
2．バックコーラス（「アー」とか「ウー」）がフェードインする様子を演出しようとして、アタックが弱々しくなる。

いずれのケースも、仮歌デモの雰囲気に歌手が従った（つもり）の結果、収録されたのは声に芯のないタダのひょろい歌声でした。おそらく声に芯を残したままマイクとの距離を変えるなり、ミックス時のレベル／EQ 操作に委ねれば結果は何倍もよくなったと思います。また、それを指示できるエンジニアやディレクターがいなかったのも残念です。

声量感の「フレーズ中の変化」、これもまた補正は困難です。
上手な人は、声の張り具合や喉の開き具合のバリエーションが豊富にあり、旋律や歌詞に沿って適宜流れるように変化し、細かいニュアンスを与えるため聴者を飽きさせません。元々、日本語のポップスでは比較的求められることのない要素ではありますが、日本人による洋楽カバーなどではこれがゴッソリ削ぎ落とされることがあります。よく英語の発音の悪さが言及されますが、こういったニュアンスの平坦さの方が、むしろ発音以上に音楽的に残念な結果につながることが多いように感じます。

不気味の谷への入り口

上手な人がフレーズの頭をしゃくり上げる場合、音程だけでなくアタックや張りも自然に変化させます。これを理解していないがために、修正の結果「音程はしゃくり上げているのに声の張り具合は一本調子」になるというのも、奇妙な印象を聴者に与える「不気味の谷」ポイントです。

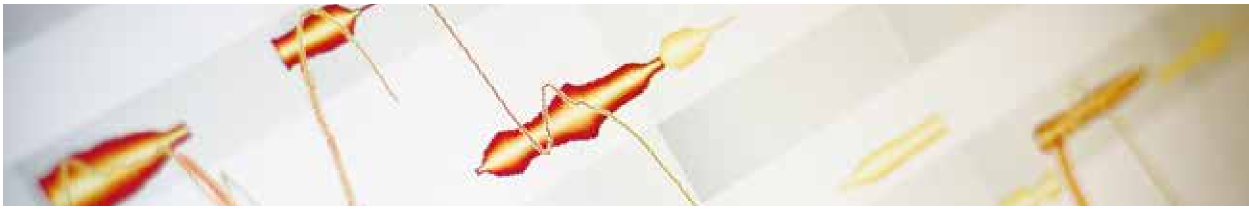
また、大抵の歌手は声域が変わると体の使い方や声の倍音構成が変化します。声に限らず多くの楽器は、高い音を出す上でより大きなエネルギーを必要としますので、より張った音色になりがちです。ハイトーンは決まって張り上げるべきものとは思いませんが、オクターブあるいはそれ以上音程が移動するのにまったく声質が変わらない歌や、跳躍前のタメを感じ取れない歌を耳にすると「この歌手は一体どんな表情で歌っているのだろう？」と収録風景を想像してしまいます。たとえば走り幅跳びの選手が、軽いジョグ程度の助走からいきなり 10m ほど飛んだら、感心する以前に、まず物理法則に反する動きに違和感を覚えないでしょうか。

ビブラートの周期

音節のどの辺りからビブラートを開始するか、またその周期（揺れる速さ）は、ある程度までは修正ができます。音の長さを伸縮させるだけでよいわけですので。しかし、ロングトーンになるほど自然な補正は困難です。

ビブラートは、振幅や周期を変化させることで、そのままであれば単調になりがちなロングトーンに表情を加えたり、旋律の大きな流れを演出する上で欠かせません。上手な歌手は、先述の声量感の強弱と同様、ビブラートを開始するタイミング（あるいは「しない」という判断)、振幅、周期、またその変化に必然性があり、メロディに流れを作ります。
逆に必然性のないビブラートは旋律の流れを止め、最悪の場合は曲の世界観を壊します。

以前に編集させていただいたボーカリストで、シャウトはとてもパワフルで魅力的なのに、２、３拍以上持続する音になると決まって、曲のテンポと相関のない深〜いビブラートを掛けるクセをお持ちの方がいらっしゃいました。あらゆる現代的なジャンルの曲を「演歌化」させていたのが残念と思わざるをえませんでした。



お勧めの練習方法

僭越ながら、一歩踏み込んで「補正可能時代」のボーカリストに求められるポイントを重点的に鍛えるための練習法を提案します。

録音予定の曲に取り組む際、まず、喉から胸部、できれば頭部までが最も共鳴する音域、音程を探します。次に、その音域、音程を中心に、旋律から音程を取り去った状態で歌ってみます（つまり「お経」状態です）。子音や張りの強弱、タメなどにのみ注目します。このとき、最初に決めた音程にこだわる必要はなく、若干上下してもかまいません。楽器たる歌手自身の身体が、一音一音をどの程度まで響かせられるか、そのポテンシャルを測ることが目的です。
上手な歌は、音程という要素を取り去ったあとに残る、声色の起伏やグルーヴ感だけでも聞かせどころや流れを持っているものです。どのぐらい豊かな声色をフレーズの各部に充てられるかを知ることは、ひいては旋律が持つポテンシャルも知ることにつながります。この作業は、必要以上に声を絞ってしまったときなどに、響きがポテンシャルに至っていないことを察知し、旋律の魅力を意図せず削いでしまうことを未然に防ぐ上で役立つと思います。

旋律の各部をどう響かせるか、筋道が見えてきたところで、今度は音程に乗せてみます。そうした途端、理想とした響きが得られない音に直面するかもしれません。さらに、そこからどう身体の使い方を変えても、先ほどのお経状態ほどの響きが得られない場面が生じるかもしれません。そのようなときの対処方法はいくつかあります。
例えば状況が許すならキー変更や旋律の書き直し、細かいフレーズに区切って部分ごとに分けてレコーディングしてみる、あるいはいずれも不可能なら修行に邁進…（投げやり）。とにかく、そのフレーズを最大限活かせる理想的な声質のイメージがあるならば、それを諦めてまで収録時に正しい音程を出す必要のない場合もあります。音程への集中が過ぎて、しっかりハッキリ歌うことへの注意がなおざりになってしまうのは残念です。少なくとも現代のレコーディングにおいて、ハイトーンが 2、3 度及ばないということであれば、その音だけ低く、その代わりに理想とする声色で歌うことをためらう必要はありません。

忘れないでいただきたのは、歌に説得力を持たせる上で

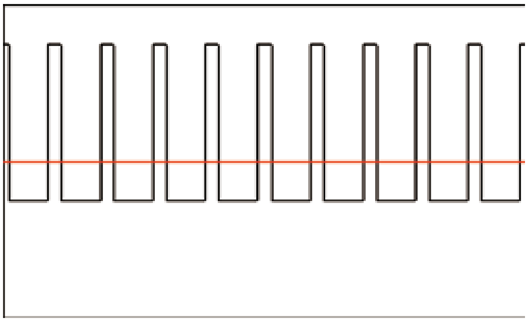
自然に響かせて 2、3 度補正した声　>>>　越えられない壁　>>>　無理に絞り出して少し音程補正した声

という事実です。

以上、著者がボーカル編集担当として 1,000 曲以上に携わった経験則に基づく私見ばかり好き勝手にお話ししましたが、少しでも皆さんの作品作りのお役に立てば幸いです。

11 章　＋/－のピークが非対称な波形について

この章では、下図のように＋ / －のピーク・レベルが非対称な収録音について考察します。

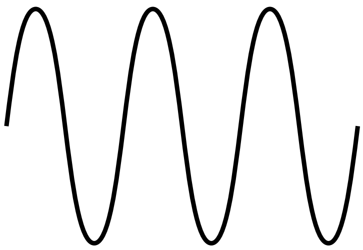


このような信号波形は、ボーカルや管楽器、あるいはシンセサイザーといった、一部の楽器群を収録した際にたびたび見られるものです。普段、ゼロを中心に＋ / －両側の形状がほぼ対称的な波形に慣れた方は不安に思われるかもしれません [*1]。このように波形のピークに偏りがある場合、**対処を必要とするかどうかは状況次第です**。

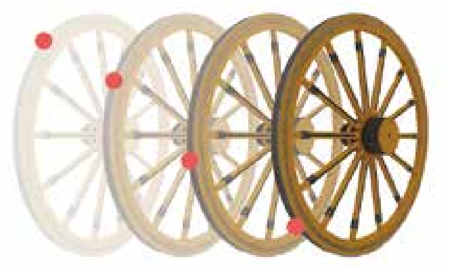
以下では、このような波形がなぜ生じるのか、またどのようなときに問題になりえるのかについて説明します。
同時に、この現象をよりよく理解するため、単一パートを構成する倍音間の位相についても取り上げます。

あらゆる信号は正弦波に分解できる

音楽制作に携わる方であれば、下図のような正弦波（サイン波）が、音声信号の最も基本的な単位であることはご存じだと思います。



正弦波



正弦波は、等速円運動をする車輪上の任意の一点の Y 軸の変化をグラフにとったもの。（Image: 3dmi / Bigstock.com）

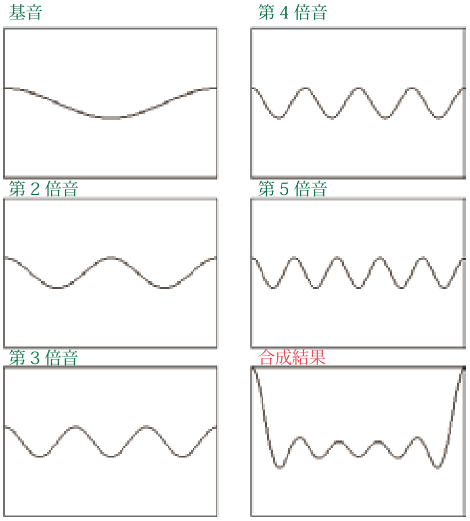
周期的に繰り返すあらゆる音声信号は、1 つ以上の正弦波の組み合わせに分解できます [*2]。
また、モノフォニックな（一度に単音しか演奏しない）楽器の波形は多くの場合、比較的少ない数の正弦波を組み合わせて表現できます。

---

*1 筆者自身、大昔に矩形波のデューティー比を変えながら収録した波形を見たとき、はじめは機材の不調を疑ったものです。

*2 たとえば 3 分間におよぶ曲でも、全体が 3 分間をかけて周期的に繰り返す信号とみなせば、正弦波の組み合わせに分解することは理論上は可能です。実用上、意義があるかどうかはさておき…

楽音内における成分（正弦波）の位相関係



図：基音〜第 5 倍音まで、ピーク・レベルの等しい 5 つの正弦波の信号波形と、その合成結果

楽音の音色をつかさどる基音や倍音…これら複数の正弦波が重なるとき、山や谷の関係性によっては、はじめに見たようにピークが＋ / －のいずれか一方に偏ることがあります。

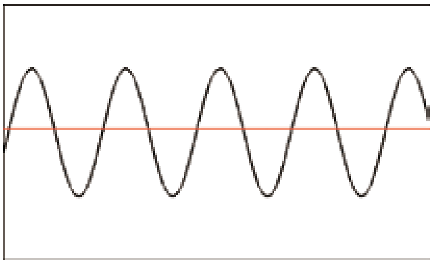
ここで実際に、等しいピーク・レベルをもつ 5 つの正弦波…440Hz の基音と、第 2 〜第 5 倍音を重ねてみましょう。
信号の生成には、Pure Data（Pd）というフリーの音響生成ツールを使用します [*3]。

コサイン波 [*4] を重ねた結果、ご覧のように＋側の方がピーク値が高い信号が生成されました。

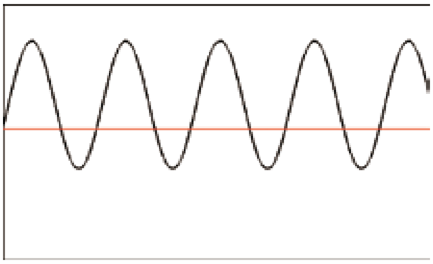
後ほどまた詳細を述べますが、とりあえずここでは、合成結果を構成する **5 つの正弦波が、いずれもゼロを中心に＋ / －側ともに対称的なカーブを描いている**ことにご注目ください。

ピークが偏ることの問題点

さて、ピークが＋ / －の一方に偏ると、なにか不都合があるのでしょうか？
まず、ヘッドルームが消費されます。たとえば、次のように単一の正弦波からなる信号があるとします。



ヘッドルームが均一に消費される。



＋側だけヘッドルームが消費される。

信号が上図左のようにゼロを中心としていれば、消費するヘッドルームは、＋ /—ともに均一で済みます。
しかし、上図右のように全体的に＋ / －のいずれかに偏ると、一方のヘッドルームを消費しはじめます。
これが、ピークが＋ / －の一方に偏ることの代表的なデメリットです。

※本章の続きは本誌にて

---

*3 Pure Data の公式サイト https://puredata.info/
*4 正弦波（Sin 波）と同じ軌跡をもつが、角度 0 において +1 となる。

4-2　トラックの左右配置とステレオ方式の限界

ステレオ方式とは、独立した２つの音声信号を同時に収録、および再生する記録方式です。この２つの信号を左右のスピーカ、またはヘッドホンのL/Rからそれぞれ再生することで、広がりのある音像を再現することができます。
この方式において、ミックス段階で信号に手を加え、空間の左右にパートを配置し、広がりを演出する手段はいくつかありますが､それぞれにメリット､デメリットがあります。本章では､それらの違いについて考察してみたいと思います。

なお、このセクションは第７章「M/S処理と留意事項」の応用編ともいえる内容を含みますので、未見の方はそちらを先にご覧いただくことをお勧めします。

奥行を表現する方法

人間の脳は優秀にできており、耳にする音の特性とその発信源の位置との関係性を日々学習しています。また、その記憶に基づき音の諸条件から発信源の距離や方位を推測する能力を備えています。ミックスにおける空間作りは、いうなればリスナーの脳にアクセスし､音の特性と方位にまつわる記憶領域を呼び覚ますよう細工を行うことでもあります。

本題である「左右の広がりを演出する方法」に入る前に、まずはウォームアップがてら前後の奥行を知覚させる音の特性についても考えてみましょう。

アタックやピークの大きさ

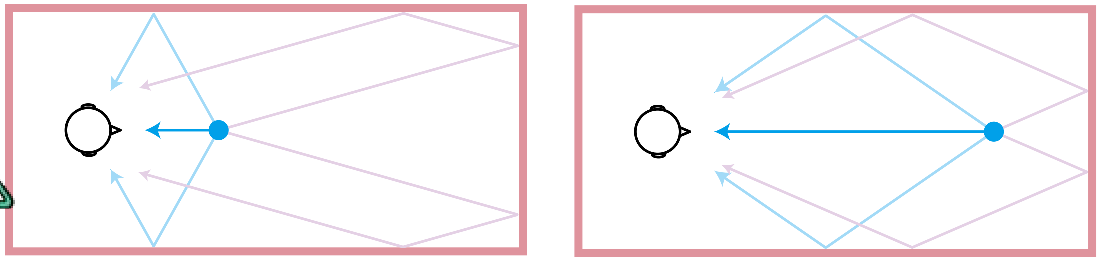
一般には、アタックやピークが大きいパートほど前面に感じられます。
逆に、アタック・タイムの早いコンプレッサなどでピークをなくせば、そのパートは奥に配置されます。

周波数バランス

大気中では高い周波数ほど大きく減衰します。このため、奥に配置したいパートは他よりも高域を相対的に減らすことで、前後関係を演出できます。生楽器の場合、リスナーが記憶する「本来あるべきスペクトラム」をリファレンスに無意識に距離を判断している可能性もありますので、増減の範囲は超現実的にならない程度が妥当かもしれません。

リバーブ

リバーブでも、ある程度の奥行は表現可能です。
たとえば初期反射（Early Reflection）が多いパートの方が、それを施さないパートよりも一歩前に出て聴こえることがあります。また、現実の空間をイメージするとわかりやすいように、リスナーと音源との距離によってリバーブの遅延（プリ・ディレイ）や「生音：リバーブ」の比率だけでなく、「初期反射：後期反射」の比率も変わってきます。



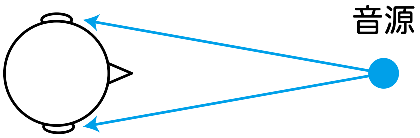
初期反射と一部の背面からの二次反射のみを記した略式の図ではありますが、音源との距離によって直接音を含むそれらの到達時間やバランスがどのように変化するか、おおよそイメージはしていただけるかと思います。

個人的な意見ですが、前述のピークや周波数バランスに比べると、一般に思われているほどリバーブの量（センドレベル）自体は、前後関係の認識に寄与していないように感じます。なお、これらの手法で得られる奥行き感は、えてして相対的なものであり、性質が対称的なパートが比較対象として音像内に配置されない場合、その効果は限定的です。

左右の広がりを表現する方法

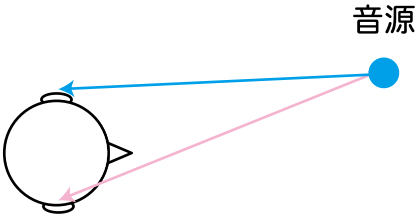
いよいよ本題に入ります。まず、実際の空間で放たれた音声が、どのようにリスナーの左右の耳に到達するかを考察してみましょう。

正面から聴こえる音



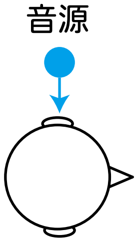
正面方向にある音源から発せられた音は、左右の耳に到達する音量、位相ともに均一です。よって、左右の耳が等しく受け取る音声を、ヒトは経験則から正面（あるいは水平方位が 0 となる真上、真後ろなど）より飛来しているものと知覚します。

左右方向から聴こえる音



次に、左右いずれかの方向から聴こえてくる音を見てみましょう。まず、音源からの距離が左右の耳で異なることによる音量差が生じます。2つ目に、リスナーの頭部自体がフィルタとなり、音源から遠い方の耳には、特定の周波数が減衰して聴こえます。３つ目に、両耳に音声が到達するタイミングの時間差、ひいては周波数別の位相差が生じます。このように、少なくとも３つの点において、先ほどの例に比べると左右の耳が受け取る情報に差が生じることがおわかりいただけるかと思います。

近接する音



音源がいずれかの耳に極端に近い場合は、左右の音量差が最も大きくなります。このため、一方の耳にしか到達しない音は、その大小を問わず、発信源も近くであるかのように知覚されるものと思われます。筆者私見ではありますが、ステレオ音像を大きく聴こえさせようと M/S 処理などによりサイド・チャンネルのレベルを上げると、場合によってはあっというまにセンターの主役を「食って」聴こえることがあるのは、このためであると考えられます。

ところで、バイノーラル録音やディレイ / フィルタを駆使し、リスナーの全方位から飛来する音を再現すべく数々の研究がなされてきましたが、ステレオ方式で実現する確実な手法はいまだに存在しません。これが難しい理由のひとつが、前述した「頭部のフィルタ特性」が体格など個人差がある要素に依存し、一般化しづらいためです。よって、最新の技法においてもフィルタなどに使用する定数は、多数の被験者からサンプリングした平均で代用する以外になく、それとて音像としては不自然になりがちであるのが実情です。

それでは、これらを踏まえた上で左右に広がるステレオ音像中に楽器を配置する方法を考えてみましょう。

注：次ページ以降で取り上げる以外に、空間を演出する手法としてはディレイなどを使ったものもありますが、これは PAN ポットの使用に加えて疑似的な初期反射を組み合わせるものと解釈し、PAN そのものの代替手段に注目する本稿では割愛します。

※本章の続きは本誌にて

4-3　ゲイン・ステージング

本書では、これまで主にマスター（最終媒体）収録時の音声レベルに焦点を当ててきました。本章ではそのマスターが完成するまでの途中過程…たとえばマイクでの収音時、オケやミックスのファイルへの書き出し時、その他 DAW 内の信号経路などにおける、適切な信号レベルについて考察します。

アナログ機器においては、回路設計などの技術的な観点から、適切とされる信号レベルが存在します。これは「ある範囲よりも信号レベルが大きければ、機器の回路、あるいは録音媒体の物理特性により信号が歪んでしまう。しかしながら過度に小さければノイズに埋もれてしまう」という両端を秤にかけて設定されています。

一方、デジタル領域で音声信号を扱う場合、上記のような制約の大部分から解放されます。特に近年の DAW やデジタル・ミキサーがそうであるように、信号処理に float や double といった浮動小数点数が使用されていれば、信号レベルが極端に大きかったり小さかったりする場合でも、信号品質に影響はありません。よって、音声をアナログ信号に変換するまでの段階で一定の範囲内にレベルが戻っていれば問題はありません。

そういった DAW 内の信号レベルに関しては、厳密なルールこそないものの「ベスト・プラクティス（合理的、あるいは能率のよい手法）」は存在します。本項タイトルの「ゲイン・ステージング（Gain Staging）」とは、経路や工程ごとの信号レベルが、基準となる一定の範囲内に収まるよう心掛けながら作業を行うことを意味します。

後に詳説するように、最終的にデジタル領域で基準とする信号レベルは各々が任意に設定できますし、ある程度仕組みを理解して運用すれば、基準値の異なるスタジオ間でデータをやり取りしても、あるいは基準レベルを一切無視しても、さほど大きな問題にはならないでしょう。

しかし、このようなワークフローやメソッドが不在であったがために、各工程を担う作業者自身も気付かないうちに音楽性を損ねてきた可能性があることは、近年、複数の著名エンジニア方も話題にしています[*1]。

使用するメータ

以降、本章に登場するレベルはすべて音声信号の実効値、すなわち信号のエネルギーに関するものです。これはおおよそ、DAW のディスプレイに現れる信号波形の面積、あるいは体感音量に相当します。この測定には、VU メータ、RMS メータなどを使用します。

一方、ピーク・メータは一切使用しません。瞬間的なレベルに関しては基本的にピーク・インジケータが点灯さえしなければよく、それが点灯して初めて問題対処のためにピーク・レベルを参照します。

---

*1　下記は一例

・Ian Shephard "The Mastering Show #39 – Digital Gain Staging – What's the point ?"
　http://themasteringshow.com/episode-39/
・Bob Katz "How To Make Better Recordings in the 21st Century – An Integrated Approach to Metering, Monitoring, and Leveling Practices."　https://www.digido.com/portfolio-item/level-practices-part-2/

すでに K-System（K-20）をワークフローに取り入れている方々にとっては自明でしょうが、以下で説明するメソッドもまた十分にヘッドルームを設けることを前提としますので、よほど特殊なケースでない限り信号レベルが 0dBFS Peak を超える心配はありません。

共通する規定レベルは 0VU

のちほどアナログとデジタルを個別に詳説しますが、結論からいいますとあらゆる場面で**基本となる信号レベルは適正レベル（Nominal level）、または 0VU** となります。

ただし、少しばかり面倒なことに 0VU は「適正」であることを示すに過ぎない相対的な値であり、それが指す具体的なレベルや物理エネルギー量は、ところ変われば異なります。これはシチュエーションごとに異なるという意味ではなく、文字どおり国が変われば推奨される基準レベルも異なる場合があります。

では、なにをもって適正レベルとするかを、アナログ、デジタルと順にみていきましょう。

アナログ機器の適正レベル

「業務機器は +4dBu、民生機は -10dBv」という話を聞いたことがあるかもしれません。
一般にアナログ機器では、これらのレベルが 0VU に相当し、適正レベルとなります。

さて、たいていの音声信号は RMS/VU メータを絶えず変動させて定まりませんし、さらにややこしいことに、よく見ると業務機器と民生機器、それぞれの適正レベルは単位まで、dBu、dBv と異なります。

このように、まるで動く的を射るような話にきこえるかもしれませんが、実はあまり難しく考える必要はありません。

業務用、民生用を問わず、おおよそ制作に使用するであろう音響機器には、通常なにかしらレベルメータが付いています。「受け側となる機器の内蔵メータが 0VU 振る程度のレベルが基本」とだけ理解しておけば、運用上は問題ないかと思います。※本章の続きは本誌にて

使用するアナログ機材に VU メータがあれば、0VU を指すレベルがその機材が設計時に想定する適正レベルであると考えて差し支えありません。（メータが正しく調整されていることが前提です。）

（※中高吹奏楽部に在籍した経験のある方は、たかだか 1 フレーズを磨き上げるために膨大な時間をかけた、あの夏を思い出していただければよろしいかと思います。）

ボーカル・エディットにおいては、ともすれば、切り出したミクロな一部分が自然か不自然か、正確か不正確かに目がいきがちです。一旦この視点に捕らわれてしまうと、旋律全体が躍動するひとつのフレーズとしてリスナーの感情に訴えかける力を持っているかどうかといった、よりマクロな視点から評価をすることが困難になります。

音程、タイミング、子音や母音の音価（長さ）のそれぞれにおいて、旋律の流れを妨げる要素がフレーズ中やその前後にないか、全体を俯瞰するように注意深く耳を傾けることが重要です。

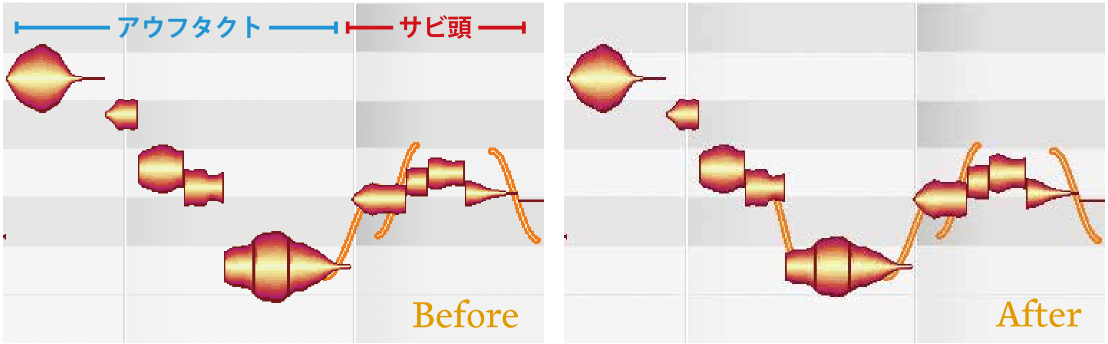
研摩ポイント事例

以下では、旋律の流れを止め、一度掴んだリスナーの意識を手放す原因になりかねないポイントをいくつか例示します。

研摩ポイント：ダウンビートやアウフタクト、その前後の強弱差

丁寧に書かれたメロディや歌詞は、一音一音をどのぐらいの重みで歌うべきかを自ら語りますし、また、奏者がそれに従うことで力強い流れを作ります。やや極端な例ですが、行進曲風のアレンジで 8 分音符主体のメロディを歌う場合は裏拍よりも表拍の方がわずかでも強い方が、少なくともその逆よりは自然に流れるでしょう。また、サビ前のアウフタクトよりもサビ頭の方が強くなるのが自然です。（アウフタクトの和訳が「弱起」とされたのは、それなりの理由があってのことです。）
しかし、歌手の技量や、掛け録りされたコンプレッサの設定などによっては、収録された演奏が、旋律が示唆する音量変化とは逆行することもあります。こういった場合、各音の強弱関係をより自然なかたちに寄せるように、音量をわずかに調整するだけで流れが大きく改善します。

先ほどあえて「自然なかたちに**寄せる**」としたのは、必ずしもダウンビートの方がアウフタクトよりも波形やメータ上で大きくある必要がないためです。本来小さくあるべき方を若干小さく、大きくあるべき方を若干大きく寄せるだけで、最終的には大小関係が入れ替わらなかったとしても、何もしないよりは十分に効果的です。



アウフタクトの末尾（右から二番目の最も低い音）と、それに続くサビ頭（右端の音）の例です。ここではアウフタクトの最後を少し小さく、サビ頭のダウンビートを少し大きくしています。目視で明確な差は見つけにくいかもしれませんが、それだけ操作が微小であるということです。

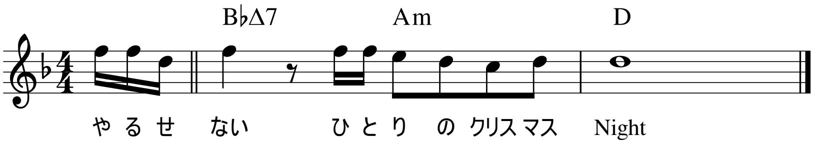
ところで、この作業はフレーズの一部に対して音量のみを調整するわけですので、極論するとミックス時にトラック・オートメーションでも同様の効果を得ることは可能です。しかし、あえてボーカル・エディットの段階で対応する方が合理的といえる理由がいくつかあります。

理由１：音量調整の結果、他の要素も再調整を要する場合がある
旋律を構成する要素は、それぞれが複雑なバランスの上に成り立っています。たとえば上記の例であれば、音量というひとつの要素を操作した結果、タイミングや音程といった他の要素についても「落としどころ」が変わって聴こえる場合があります。よって、ボーカル・エディットの段階でまとめて調整を行う方が合理的であると考えられます。

理由２：曲全体が最終形に近い状態からミキシングを開始できる
ボーカル・エディットの段階で音節の音量調整を終えると、編集者とミキシング・エンジニアが同一でない場合、後者の手元には最終形に近い素材を渡すことができます。また、私見ではありますが、ディレクターがボーカル・エディットやミックスを兼任する場合でも、それぞれの役割を果たすべくアタマを切り替える頻度が少ない方が、何かと作業がスムーズに進むように思います。今回の例に当てはめると、ミックスに取り組む段階においては、ディレクター的なマインドで演奏のブラッシュ・アップにフォーカスしてしまうと、本来ミキシング・エンジニアが注視すべき全体のバランスを見失いやすくなります。前もってボリュームの抑揚を演奏の一部として完成させておくことでミックス中に役割の切り替えを減らすことができるのであれば、それに越したことはありません。

他にも些細なことではありますが、Melodyne といった専用ツールが提供するインターフェースを通して音節単位のレベル調整を行うことに比べると、複数のノードを管理しながらオートメーションを描く作業は、やや煩雑に思われます。

研摩ポイント：子音の音価



上図をご覧ください。冒頭「や」の子音をどれだけの時間をかけて展開するか（いきなり「や」と口を開くか、あるいは「ぃやー」とゆっくり目に開くか）は、それに続くフレーズのスピード感に影響を与えます。

また、「やるせない」の「な」の子音Ｎをどの程度タメるかについて、２つの極端な例として
Ｎが短い：やるせなーい
Ｎが長い：やるせンなーい
があり、その中間には無数のバリエーションがあります。このタメ具合も、昨今のツールを用いてある程度の操作が可能です。

こういったマ行やナ行の子音をどれだけタメるか、あるいは軽く短く流すかが、旋律の流れに大きく影響することがあります。これは当該部分だけを切り出して聴く分には、好みによる判断で片付けることもできますが、ときに、その前後の子音との比重が大きく異なったために浮いて聞こえるというケースがあります。
なお、こういった「タメ」の長さがもたらす変化に関してサ行の歯擦音にも同様のことがいえますが、ナ行やマ行ほど奏者がなおざりにすることがないためか、流れを止めるほどの影響があるケースは少ないように思います。

※本章の続きは本誌にて

4-6　ピーク・レベルの適正値に関する略史

本書では、これまで主に RMS やラウドネスのレベルについて複数の角度から取り上げてきた一方、ピーク・レベルに関しては「意識すべきでないもの」「ヘッドルームは楽曲が必要とするだけ確保すべきもの」というスタンスを基本としてきました。これは K-System をワークフローに取り入れたり、浮動小数点数形式のファイルでデータを管理することにより楽曲のミックス段階あたりまでは比較的容易に実践できますが、最終工程であるマスタリングにおいては、いまだに徹底するのは難しいのが実情ではないかと思います。たとえば、筆者自身の経験では、K-20 を基準にミックスした楽曲を現在 YouTube が採用している規定ラウドネスである -13LUFS 程度にまで持ち上げると、サウンドには大きく影響しない頻度ではありますが、楽曲中に何ヵ所かはピークが 0dBFS を超える箇所が生じ、デジタル・リミッタの使用を完全に避けて通ることはなかなかできません。

このデジタル・リミッタのスレッショルドに設定すべき値、すなわち楽曲中のピーク・レベルの最適値をインターネットで検索すると、-0.1dBFS とするものから -2dB TP（True Peak の略。詳細は後述）程度を勧めるものまで情報に幅があり、確たる正解らしきものはなかなか見つかりません。本稿では、このようにピーク・レベルの最適値について諸説生まれた歴史と背景について考察してみます。

なお、執筆にあたり、いくつか古い資料を紐解いた結果、物事の時系列や数値の根拠については確信を持って述べますが、それぞれのアプローチが主流であったといえるほどに実践されていたかどうか、あるいは本当に注意を払うほどの事象であるかについては筆者にとっても未知であり、様々な考え方があるかと思います。以下はあくまで筆者の視点によるものであり、また特定のタイトルの信号レベルを計測しても、その時代のベスト・プラクティスとされた手法に添った痕跡が見られるとは限らない点については、あらかじめご了承ください。

1. 0dBFS がボーダーとされた時代（デジタル黎明期、1980 年代半ば〜）

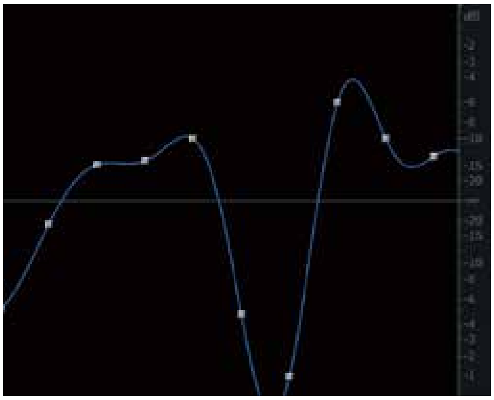
デジタル方式としては初の普及媒体である CD が誕生して間もない頃は、ただただ信号レベルが A/D、D/A 変換時に 0dBFS を超えることがないようにだけ注意が向けられていました。ただし後年のように体感音量を競う動きもありませんでしたので、数 dB のマージンがあるようなタイトルも珍しくありませんでした。この頃は、現在のように 0dBFS が連続する CD マスターをプレス工場に持ち込もうものなら、製造時のミスによる欠陥品と間違われることを懸念した工場に突き返されるか、さもなくばマスタリング・エンジニアが即日解雇されるか、あるいはその両方が起こってもおかしくないほどの蛮行であったといわれています。

2. 0dBFS 超えのピーク・レベルが意識されはじめた時代（音圧戦争、2000 年頃〜）

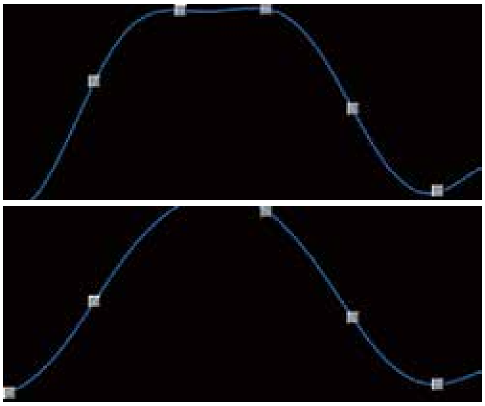
2000 年前後より、デジタル・ピーク・メータが示すサンプルの最大値よりも D/A 変換後のピークが高くなる場合があり、ときにはこれが歪みとなって現れるケースがあることが話題になりはじめました。

このような現象が生じる原因は、主に 2 つあります。
ひとつは、インター・サンプル・ピーク（Inter Sample Peak=ISP）、またはトゥルー・ピーク（True Peak=TP）と呼ばれるものの存在です。これは、音声データ中のサンプルがすべて 0dBFS 以下であっても、サンプル・レートの変換時にサンプル間を補間した結果、0dBFS を超えるレベルで現れるピーク成分です。（次ページ、左上の図）

D/A 変換中にピーク・レベルが 0dBFS を超えてしまうもうひとつの主な原因は、ローパス・フィルタを通した結果、元々は矩形波に近かった信号が変形される現象によるものです。（次ページ、右上の図）



インター・サンプル・ピーク（トゥルー・ピーク）の例。サンプル間を補間した結果、0dBFS を示す画面下端を超える +0.8dB のピークが生じていることがわかります。



ローパス・フィルタによりオーバーロードが生じた例。マキシマイズ直後、サンプル・ピーク、トゥルー・ピークともに -0.1dBFS 程度であった信号（上）が、ローパス・フィルタを経た結果、サンプル・ピーク +0.1dBFS、トゥルー・ピーク =+0.4dBFS になっています（下）。

さて、D/A 変換後のピーク・レベルが 0dBFS を超える原因として、サンプル間の補間、およびローパス・フィルタ処理を取り上げましたが、これらはいずれも制作過程でエンジニアが行わなくても、多くの場合、モダンな D/A コンバータの内部でなされています。

このようにして 0dBFS を超過した信号が適切にアナログ変換されるかどうかは、D/A コンバータの設計次第です。仮に、+3dBFS の信号が発生し、DAC チップからは正しくアナログ信号が出力されたとしても、これが正しく再生されるためにはすべての周辺回路もまた、本来想定されるよりも 2 倍近い電力をクリーンなまま取り扱えるように設計されている必要があります。実際、同じ DAC チップを採用した CD プレイヤー間でも、0dBFS を超えるテスト信号を難なく再生するものから、ピークを扱いきれずに矩形波に近いアナログ信号を出力するものまで、幅広く存在することが確認されています [*1]。さらに加えるなら、民生機クラスの再生機でのみ生じる歪みによるサウンド変化が、制作段階では見落とされる可能性も考えられます。

さて、0dBFS を超えうる信号の取扱いに関して D/A コンバータが潜在的に抱える問題は、これが話題になり始めた 2000 年前後よりも以前から認識はされていたかもしれませんが、問題視されはじめたのは音圧戦争の火種がくすぶりはじめ、正しく再現することが難しい信号を含む CD タイトルが市場を席捲し始めたことと無関係ではないように思います [*2]。前述のように、0dBFS が連続する CD は「規格外」であったはずが、いつしかこれが一般的になり、技術的なレイヤーにおいても媒体の扱い方や制約に関する常識が一変してしまったのです。

ピーク・レベルの最適値としての一論である「-0.1dBFS 以下に抑えることが望ましい」とされる考えが起こったのは、この頃であったように覚えています。テスト信号ならいざ知らず、一般的な楽曲においては、この程度のマージンがあれば大抵のケースで D/A 変換時のオーバーロードは避けられるだろうとする、一部エンジニアの経験則を根拠とするものではないかと思います。

※本章の続きは本誌にて

4-7　ヘッドルームとコーデックの落とし穴

前章では、音源制作時に -1dB（TP）はマージンを設けるべきとされる理由をいくつか紹介しました。そのひとつが、音声を AAC/mp3 といった非可逆圧縮形式にエンコードする際、これにともないピーク・レベルが上昇し、0dBFS を超えて歪みが生じるおそれがあるというものでした。とはいえ、市販 CD でさえ 0dBFS に張り付くタイトルがほとんどである中 [*1]、こうした歪みの発生には、実際どのぐらい気を遣うべきなのでしょうか？

結論からいいますと、技術的な側面からみても、本来であればあまり気を遣う必要はないはずなのです。しかし、なんとも理解しがたいことではありますが、現実にはコーデックの多くが非合理的な設計となっており、これが災いして甚大ともいえる影響が生じているといえます。

本稿では、非可逆圧縮形式にエンコードする前の音源に十分なヘッドルームを設けない場合に楽曲が受ける影響について、ゆるく考察してみます。

AAC/mp3 の特性

デジタル・フルスケールに近い音源をエンコードする危険性について論じる前に、まずは前提として、AAC/mp3 形式にまつわるいくつかの揺るぎない事実を確認します。

事実 1：AAC/mp3 はエンコード時に、サンプル・ピークが必ず上がる

前章のおさらいになりますが、音声の圧縮に際しては、フィルタリングや量子化誤差によるノイズが加わり、その結果**必ずレベルの上がったサンプルが生じます**。
ここで問題にしているのはトゥルー・ピーク（サンプル間に生じるピーク）といったような、測定方法によって出力結果が異なるあやふやなものではなく、サンプル自体のレベルであることにご注意ください。
エンコードによりサンプル・ピークのレベルが上昇することは、WAV ファイルを mp3 に変換後、iZotope RX7 に読み込むことで簡単に実証できます。これは、ヘッドルームが十分にあるような音源でも同様の現象が観察できます。

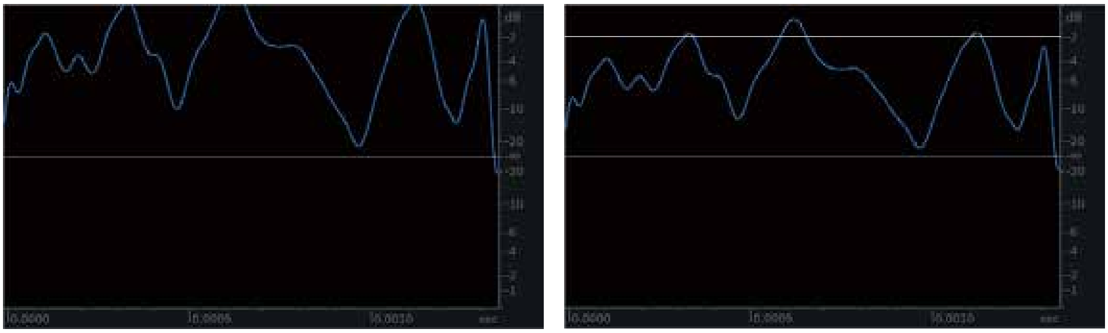
2018 年リリースのある作品より大サビを抽出し、mp3 エンコード前後のレベルを比較したものです。
(上記中央がエンコード前、右がエンコード後)
一見すると、トゥルー・ピークが 0dBFS を超えない CD マスタリングに成功しているように見えますが、このファイルを mp3 に変換すると、サンプル・ピークまでもが 0dBFS を超えるファイルが出来上がりました。

---

*1　本書第 3 版の制作にあたり、巻末の「市販楽曲ラウドネス一覧」をより充実させるため、近隣のレンタルショップで 2018 年 6 月～ 2019 年 5 月までにリリースされたシングルばかり 40 枚を借りてきて調査しました。これは店頭で見つけられたうち、上記条件を満たすほぼ全タイトルにあたります。そのうち、本稿で取り上げる mp3 デコード時の影響を受けないと判断できる収録レベルのタイトルは、ひとつもありませんでした。

事実 2：AAC/mp3 は 0dBFS を超えるレベルの信号を収録できる

mp3 は端的にいうと浮動小数点数に似た性質を有しており、0dBFS オーバーの信号も記録することができます。よって、**エンコード時に 0dBFS を越えるサンプルが生じること自体は問題にはなりません**。ただし、そのまま D/A 変換すると、当然そのようなサンプルは歪んでしまいます。



左：mp3 をデコードした状態、右：同 mp3 ファイルをデコード後に 2dB 下げたもの。
2dB 下げたのちも、-2dBFS 以上のサンプルが個々のレベルを保持していることから、mp3 ファイル内の 0dBFS を超えるサンプルも問題なく記録されていることがわかります。

■世のデコーダにひそむ不条理

さて、勘の鋭い方は疑問に感じていらっしゃるでしょうが、先ほど述べたように AAC/mp3 が 32bit float などといった浮動小数点数のような性質を有しているのならば、D/A 変換前にデジタル領域でレベルを下げることで、0dBFS を超えるサンプルは特に問題にはならないように思われます。

実は、理屈の上ではまったくもってそのとおりです。しかし、ここで本稿の冒頭で触れたように、デコーダにおける「非合理的な設計」という不条理な現実に直面することになります。

※本章の続きは本誌にて

4-11　CD-R ブランドで音は変わるか

いまさらですが…CD-R はブランドによって音が異なるのでしょうか？私自身はといいますと、以前は絵に描いたような懐疑派でした。その後の経験を通じて CD-R のブランドで音が異なる「場合もある」と理解しています。
以下は、CD-R に対する考えを改めるに至った個人的な体験や、そこから得た考察についてのお話です。すでにご自身の結論をお持ちの方も、今後も懐疑派を貫かれる方も、リラックスしてこの小噺をお楽しみいただければ幸いです。

スタジオでの比較実験

いまから 10 年ほど前、私はあるスタジオにアシスタントとして一年半ほど在籍していました。このスタジオは一般向けに開放されたものではなく、とある放送局が自社コンテンツの制作用に擁していたものですが、48ch の SSL9000J を備え、ちょっとしたアンサンブルから小規模なオーケストラまで生楽器中心の収録を日常的にこなす、立派なスタジオでした。

その局では、全社で使用する CD-R を特定のブランドに統一し、自社ロゴを入れたものを大量調達していました。商品名は失念してしまいましたが、記録面がいまや見かけなくなった上品な金色で、確か三菱化学製だったように覚えています。

その社内標準品が、ちょうど私の在籍中にディスコン（製造中止）になることが決まったとの報を受け、急遽スタジオ・スタッフが代わりになる CD-R ブランドを選定する必要が生じました。

まず、納入業者に各ブランドのサンプルを数枚ずつ収めてもらいました。十数種あったかと思います。
次に、それぞれに同一条件でデータを書込みました。書き込む際には CD ライタの温度条件などが等しくなるよう、一枚書き込むごとに PC をシャットダウンして一定時間放置し、起動からン分後に書込みを開始するという徹底ぶりでした。

こうして同じ条件でデータを記録した各種ブランドの CD-R が一通り揃うと、スタジオのコントロール・ルームにスタッフを集め、備え付けられた Genelec のラージ・モニタでの大試聴会が行われました。
CD-R をランダムな順で再生機に入れるのは、下っ端だった私の役割でした。
再生するブランド名は評価を行うスタッフには告げず、コミュニケーション時には、それぞれの盤に割り振った通し番号を使用します。時折、比較のために「○番の盤、もう一度かけて」といったリクエストに応えることもありました。

実のところ私自身は、この時点でもまだ有意な差があるかどうか半信半疑でした。が、これだけの設備、これだけの大掛かりな品評会となれば、もしや差がわかるかも？あるいは自分だけ判らなかったらどうしよう？と、期待と不安の入り混じる複雑な心境でした。

そうして CD-R を入れ替えて再生するたび、親方、先輩、兄弟子方がウーンと唸ったり、クスっと笑ったり、ときには「これはないわー」と口にするものですから、純粋なブラインド・テストの体をなしていたかは今も疑問が残りますが、ともかく各々、手元の評価シートになにやら書き込んでいくのが見えました。
一方、CD 交換係の私は、スピーカの軸から大きくはずれ、再生機の設置されたコントロール・ルームの隅っこの方で聴いていたので…もとい、これが理由と願いたいのですが、なんとなく差異が聴こえるような、ないような、はっきり

しない気分で聴いていました。当時プラシーボの力を非常に強く信じていたので（それ自体は現在も変わりませんが）おそらくそれがかえって違いを素直に受け入れることの妨げにもなっていたのでしょう。「そんなに一喜一憂できるほどの差があるかなぁ？」と不思議に思いながら CD を入れ替える作業を続けました。

ちなみに、再生に使用していたデッキが古く、CD-R 特有の色素から反射するレーザを認識しづらいこともあり、半数とまではいかないまでも結構な数のブランドでは TOC（CD-R 内周に書かれたトラック情報）が読めず、スタートラインにすら立てないうちに次期・自社標準品の候補リストから消えました。局内には古い再生機も多数ありましたので、音質以前に確実に再生できることも重要なポイントだったのです。

そうこうするうちに、１枚のなんとも摩訶不思議な盤に直面します。
テストに用いた音源は、冒頭に完全にドライな（＝一切リバーブのない）小編成のストリングスが奏でるフレーズで始まるのですが、このブランドだけが再生開始から早々に、ピッチが大きくヨレるのです。
高級なストリングス音源を評してよく使われる「Lush」がまさに元音を形容するにふさわしいなら、この CD-R の手にかかれば名演奏もヘニョヘニョ〜ン。

テープに例えると１桁 Hz の周期で揺れるワウ・フラッター、あるいは FM ／ウェーブ・テーブル・シンセシスで遊んだことのある方は、再生速度を低周波で軽く揺らしたような音をご想像ください。

この盤がかかった途端に一同爆笑。このブランドは、それまでにない速さで候補リストからはずれました。

さて、気になったので後日再検証したところ、このブランドの CD-R を特定の再生機と組み合わせた場合に限り、何枚焼いてもまったく同じようにピッチがヨレるのです。
さらに、リッピングすれば書き込み前のデータとバイナリ一致したのもまた、何枚焼いても同様でした。

以上、この一件をきっかけに、私は CD-R のブランドが音質…というより、もっと広義に再生機の出力に及ぼす影響について、考えを改めるに至りました。

余談１：品評会の結論は、ディスコンになったものは仕方ないとはいえ、選定した盤のいずれもが、それまで使用していた例のゴールド盤には遠く及ばない、というものでした（消去法で選んだ盤がどのブランドであったかも、残念ながら失念しました）。

余談２：お詳しい方からは「テストした CD-R メディアに対応するライティング・ストラテジが、ライタ側のファームウェアに登録されていたのか？」などと突っ込みを頂戴しそうですが、他の粗ともども、若気の至りということでご容赦ください…

なぜメディアで音質が変わるか？

さて、なぜそもそもデジタル媒体であるはずの CD-R のブランドが、再生音に影響するのでしょうか？ 0/1 の羅列を盤から読み取るだけであれば、音が聴こえるか、盤面を読み取れずに聴こえないかの ２ 択となり、劣化や変化はありえないはずです。

※本章の続きは本誌にて

4-13　ハイレゾの話

96kHz/24bit といった高サンプル・レート、高ビットは制作環境においては特に目新しくありません。しかし、「ハイレゾ（High Resolution）オーディオ」の名の下に、これが流通媒体として消費者市場を賑わせ始めたのは、比較的近年のことです。

ハイレゾの定義

ハイレゾには明確な定義が存在しません。基本的にはビット・デプスまたはサンプル・レートのいずれかが CD より高く、もう一方が同等以上であれば、ハイレゾと呼んでも後ろ指をさされることはないでしょう。日本国内の業界団体は、一般に 96kHz/24bit 以上と定義しているようです。執筆時点では、Technics 社の Web サイトをはじめ、英語サイトの多くが単純に「CD の 44.1kHz/16bit よりも高いレート」が該当すると記しており、国外の方がハードルが低いようです。同じ高精細を指す語でも、映像業界の High Definition の定義が明確であったのに対し、ここでも音楽媒体のアバウトさのようなものを感じてしまうのは、収録レベルの基準が存在するか否かと同様であるように思います。このあやふやさが、次に見るような諸問題の生じる隙をつくり、ハイレゾ商品全般の信用を損ねている感は否めません。

ニセレゾ問題

ハイレゾとして販売されているコンテンツの中には、それっぽい器にこそ入ってはいるものの、中身がそれに見合うかどうか判断の難しいものもあり、それらは**ニセレゾ**と呼ばれることがあります。その多くが、CD 相当レートのデジタル・マスターしか残っていないタイトルの再販にまつわるもので、おおよそ以下のケースに分類されます。

1) CD 相当、あるいはハイレゾ・コンテンツよりも低レートのデジタル・マスターをアップ・コンバートしただけ

単純なアップ・コンバートは再生機側でも行えますし、極論すると一般リスナーでも CD をリッピングし、フリーソフトを使用して自前で変換することができます。少なくともハイレゾ版に、CD よりも高額なプレミアム価格の値札を付ける意義はありません。

これの変則パターンとして、マスターは 48kHz（一応ハイレゾ）なのに、96kHz にアップ・コンバートされた上で必要以上に高レートで販売されていたケースがあったそうです。所属レーベルより求められたため、理不尽さを認識しながらも、やむなくアップ・コンバートしただけのトラックを納品したことがあるという、あるアーティスト本人から直接伺った話です。このエピソードの本書収録にあたり、匿名とはいえ当該アーティストに掲載の可否を確認したところ「いいですよー、よくあることでしょうし」とのお返事をいただきました。

2) CD 相当のレートで収録されたデジタル・マスターに、メーカ独自技術で高周波を付加したもの

CD 相当のレートでは、ナイキスト周波数である 22.05kHz 以上の周波数は収録できません。
そこで、アップ・コンバート時にメーカ独自のアルゴリズムで、変換前のソースにはない高周波を付加したようなタイトルも存在します。
この手の商品が問題になるのは、その変換時のアルゴリズムが、変換作業を行う技術者の主観が入ることなく、一律に作用する場合です。
一律に作用するのであれば、これも原理上、同様の変換は再生機側に実装することができるはずです。既存の CD コンテンツとは違う値札を付けて販売する、あるいは既に CD を所有しているリスナーに売り込むのは筋が通らないように思われます。

3）CD 相当のレートのデジタル・マスターを、現代的な高レート環境でリマスタリングしたもの

この辺りから、批判する側の正当性も、ややグレーになってきます。
CD 相当のレートを持つマスターが元とはいえ、リマスタリング時にアナログ機材を通し、現代技術の粋を活かした高レート A/D コンバータで取り込んだのであれば、たしかに新規に生じた高周波など、ハイレゾの器にしか収まらない信号も含む商品になるでしょう。しかし、デジタル・マスターが CD 相当というのは言い換えれば、制作環境の整った制作者寄りのリスナーにしてみれば「自分で作りえた」ものでもあります。こういった商品の対極にあるのが、CD 登場初期の A/D コンバータでは取りこぼしがあったため、アナログ・マスターから再度デジタル可し、ハイレゾで販売される商品です。レーベル秘蔵のアナログ・マスターを、数十年の時を経てママ取り込みました…といったタイトルに比べると、マスターの出発点が CD と同クオリティのものはプレミアム感も限定的です。それが近代的なマスタリングを施され、当時のフィーリングを損ねていればなおさらです。

さて、ニセレゾと呼ばれる音源の是非はともかく、これらの例だけでも、収録、編集、マスタリングまでを高レートで行った作品とは随分と出自の違うものが、同じ「ハイレゾ」の名の下に流通している混沌ぶりがわかるかと思います。

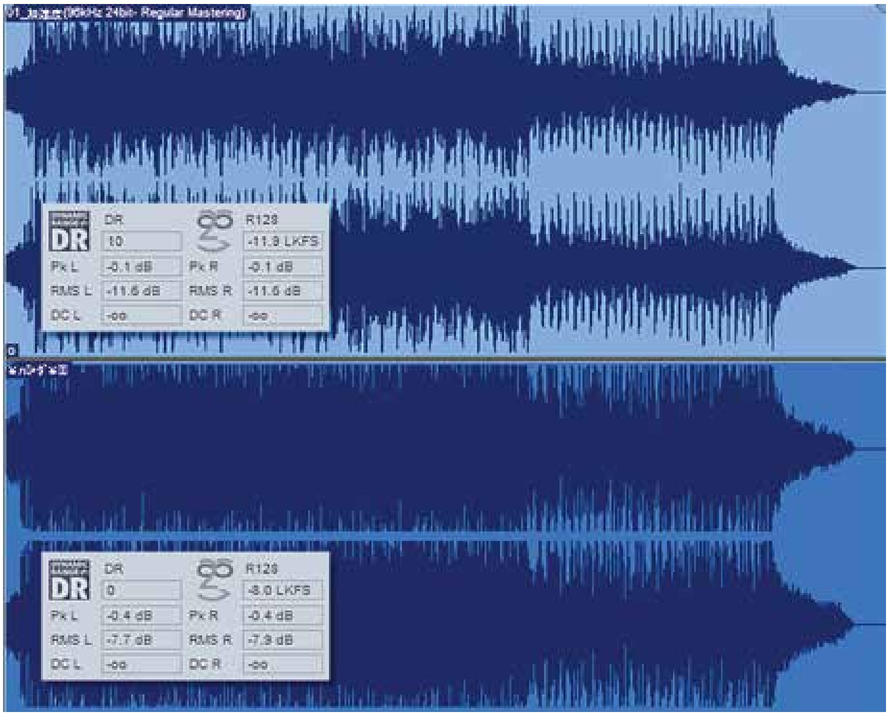
余談ですが、「せっかくアナログ・マスターを元にしたハイレゾ版なのに、現代的なマスタリングでしか聴くことができない」という話、なんとなく映画「スター・ウォーズ」を想起させます。よく知られた話ですが、公開順では初期にあたり、後に旧３部作と呼ばれるエピソード４〜６は、家庭用ビデオ版が刷新されるたび、監督自身の手によって少しずつ編集が加えられました。後には、劇場公開当時はまだ存在しなかった CG 技術をふんだんに使用した新規カットまで挿入され、ついに新三部作との整合性をとるため、合成により役者の差し替えまで行われてしまいました。
（これに対して、旧作を汚されたと感じた多くのファンが激怒しました。）さて、しばらくは最新の編集バージョン、劇場公開時バージョンともに DVD が入手可能でしたが、ちょうど Blu-ray が普及し始めた頃、最新のバージョンのみが正統であると監督が宣言し、以降劇場公開時バージョンは家庭向けには一切販売されなくなりました。今日でも HD 媒体（Blu-ray）では公開当時のバージョンは存在せず、高精細では見ることができません。

マーケティングの問題

上記まではハイレゾ・コンテンツの、いわば中身に関するお話でした。もちろん、流通しているコンテンツのすべてが例示したようなものではなく、中には正しく工程管理され、高レートで制作された良質な作品もたくさんあります。
よって、フォーマット自体を否定するつもりはありません。一方で、ハイレゾのハードやソフトを提供する側の売り文句を見ていると、はたしてその意義を正しく理解しているか、ひいては無理解から正しく工程管理されていない商品もあるのではないかとさえ疑わしくなることがあります。ハイレゾ関連のハード、ソフトを売りたいメーカの気持ちは解りますが、アプローチが概しておかしい点は指摘せざるをえません。
右図は、ある大型家電量販店のディスプレイです。なぜかハイレゾ音源の比較対象が、非圧縮フォーマットである CD ですらなく、96kbps AAC 配信となっています。

また、各社の試聴体験記事では、同一タイトルのハイレゾ版と CD 版を聴き比べる際に、そもそもマスタリングが異なるものが選ばれているのを何度となく目にしました。その上で、ハイレゾの方を評して「広がりがある」「音がクリア」「ボーカルが生々しい」などと書き並べます。いずれもマキシマイズさえしなければ、CD でもそこそこ表現可能なポイントばかりです。特に「目の前で歌っているよう」という感想については、本書 6 章で図解したとおり、CD 版はマキシマイズの結果、演者がリスナーから遠のいて聞こえているだけである可能性が高いことが想像できます。このような収録フォーマットの比較を正しく行うには、同一マスタリングでなければ意味がないことは、制作サイドの人間でなくともわかるかと思います。仮にハイレゾ版のデータ・レートが 96kHz/24bit であったならば、それを正しいプロセスで 44.1kHz/16bit に変換した上で双方を聴き比べる必要があります。もし同一ソースのダウン・コンバート前後を比較した上で違いが認識できれば、ハイレゾ版が優位だと言えるでしょう。逆に違いがわからなければ、そのリスナーにとっては､あるいはそのタイトルにとってハイレゾである意義は小さいということになります（実際､ジャンルやミックスによって、どの程度はっきりと違いが現れるかは異なります）。
さらに、ここからが最も私が伝えたい事実ですが､「ハイレゾを 44.1kHz/16bit 変換した版」を「マスタリング違いの CD」と比べたとき、前者の方が明らかに良好と感じることがあれば、それはハイレゾ云々以前に、むしろ CD のマスタリングに問題があるといえます。これはたとえ話ではなく、少なくとも筆者自身は、単純にハイレゾをダウン・コンバートしたものを CD として出した方がよかったのではないかと思うタイトルに複数の心当たりがあります。

こういった正当な比較を行わず、CD と比べたハイレゾ版の優位性を語る体験記事は無意味であり、まったく信用に値しません。仮にも真面目に技術を扱っている体の下で書かれた記事であればなおのこと、記事の企画から公開まで誰もこの点に気付かなかったとは考えにくいです。故意であれば読者や消費者に対して不誠実ですし、CD マスタリングの問題を指摘すべきところを看過することで、CD セールスを窮地に追いやる行為であるとすら考えます。



とあるタイトルの､ハイレゾ版（上段）と CD 版（下段）。ともに 2015 年リリース。あるサイトの比較試聴記事では解説中、CD マスタリングの際にはリミッティングが不可避であるために差が生じるともとれる表現がありましたが、これは見方によっては、マキシマイザ登場以前に CD フォーマットでリリースされたあらゆる名盤を否定する発言とも受け取られます。

4-14　Twitter アンケートに見る Loudness War

筆者の（悪い？）クセで、ラウドネスにまつわるトピックについて皆様のお考えが気になったり、「えっ、本当にそうなの？」と思う言説を目の当たりにしたりすると、すぐにTwitterの投票機能を使って質問を投げかけてしまうのですが、そうしていると、ときどき思いもよらぬ結果になることがあります。このセクションでは、その中でも特に興味深かったものを選び、問いかけに至った経緯と、集計結果から得られた個人的な見解をご紹介します。

なお、いずれも私をフォローしてくださっている 2,000 前後のアカウントを起点にリツイート等で広まり、お答えいただきましたので、サンプルに偏りがある感は否めません。しかし、非常に無責任な私見ではあるものの、拡散される様子を眺める限りでは、回答数が 200 〜 300 を超えるあたりからそこそこ意義のある数字が出ているのではないかと思っています。皆様にも、ゆるりとお楽しみいただければ幸いです。

まず、ひとつめはこちら。
ある日 Twitter を眺めていると、2 つの異なる方面からほぼ同時刻に、相反する意見が流れてきました。

・音圧戦争に対して文句を言っているのは、音楽制作に携わる一部の人間だけ。一般リスナーが「この CD の音がデカすぎて聴く気にならない」なんて言っているところを見たことがあるか？

・文句を言うのは、聴くことを専門にしている一部のオーディオ・マニアだけ。制作サイドは理解して作っているはず。

さらに別の機会には、ラウドネス・ノーマライゼーションが広まったのは一部の消費者がクレーマー的に振る舞った結果であると批難する意見さえ度々目にしました。

筆者自身は、現状を憂いている制作者とリスナーのどちらをもたくさん知っていますし、制作側に限っても編曲者からエンジニアまでのあらゆる立場で、いまのサウンドの在り方に問題意識をお持ちの方々に心当たりがあります。ですが、それぞれの割合については考えたことはありませんでした。

そこで、皆様にお尋ねした結果がこちらです。

バックグラウンド別にみるとほぼ半々に票が分かれたことからも、リミッティングによりダイナミクスを損なう慣習に不満を抱いているのは制作者だけ、あるいは一部リスナーだけ、と割り切ることはできないように見えますが、いかがでしょうか？

続いてはこちらです。

先日、次のようなツイートを目にしました。
「高音圧は良くないって言うけど、第一印象でインパクトを与えるにはラウドにするしかないよね」

※本章の続きは本誌にて

最終章 ～デジタル・スタジオから「音圧」が消える日

ここまで読み進めてくださった読者の中には、筆者が本書においてラウドネスや、いわゆる「音圧」に関する様々な話題を取り上げてきたにもかかわらず、「音圧」という語句をほとんど使用していないことにお気付きの方もいらっしゃるかもしれません。これは、いかようにも受け取ることのできる「音圧」よりも明示的な表現を心がけたところ、「音圧」という語句にお出ましいただく機会がなかった…というのが実情です。

「音圧」なる具体性を欠く語句が広範に使用されるようになった結果、解釈の相違から不毛な議論が起こる場面を、筆者はこれまでに何度となく目にしてきました。本書のコラムで「dB という単位系の誤認が、そのまま機器の内部動作についても誤った解釈をすることにつながる」ケースについて述べましたが、それとは比べ物にならないほど「音圧」という語句のあやふやさに起因する社会的、文化的な損失は大きいものと考えています。

もう一歩踏み込んでいうと、混乱の温床でしかない「音圧」という言葉には、レコーディングの現場から即刻ご退場いただき、より具体的な表現に置き換えられるべきであると筆者は真面目に考えます。

本稿では、一般に「高音圧」と形容される音源がどのようなものかを再定義することにより、録音芸術の文脈における「音圧」という語句への依存から脱却すると同時に、海苔波形の是非に関するあらゆる議論に終止符を打つことを試みます。その最終的な目的は、特定の傾向をもつサウンドを賛美、あるいは批難することではなく、あくまで現在流布する誤認を正し、さらなる混乱を防ぐことにあります。

以降、本稿では「音圧」の指すところのあやふやさを陽の下に晒すべく、あえてこの語句を用いつつも、ほぼ例外なくカギ括弧付きでご登場いただくことにします。

なお、これまで本書では 4 章を中心に「マキシマイザ」と「ブリック・ウォール・リミッタ」を同義として取り扱ってきましたが、近年はバス・コンプレッサやエキサイタのような効果を併せもつ製品もまたマキシマイザを名乗るケースが増えてきたように見受けられます。これらを区別するため、本稿では以降、旧来のブリック・ウォール・リミッタに相当するものにのみ注目し、「リミッタ」と表現することとします。

コラム：配信サービスが音圧を下げる？

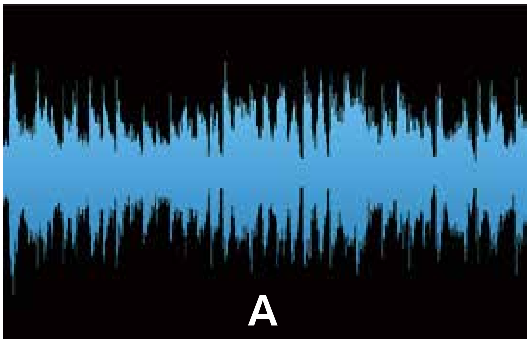
音楽配信サービス各社がこぞってラウドネス・ノーマライゼーションを導入して以降、「配信サイトが音圧を下げる」という表現を多く目にするようになりました。同時期より、配信サービスが音源を「De-compress する」…すなわち音楽表現に干渉すると誤解されるケースも筆者は度々目にしてきました。

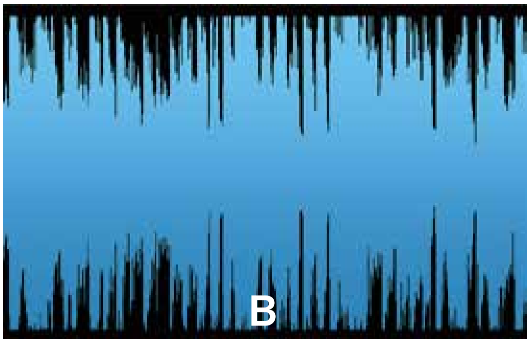
本稿に至った方々には説明不要でしょうが、ラウドネス・ノーマライゼーションを行う配信システムは、音源の再生開始時に一度、全体のボリュームを上下させるだけです。ダイナミクスの演出意図に手が加えられることはありません。ですので、制作者が表現として強いコンプレッションをかけクレスト・ファクタを小さくしたければそのようにするのは自由ですし、リミッタについても同様にサウンドが好みであれば使用すればよいのです。ラウドネス規制下では、同調圧力などを気にせず、好みのサウンドを追及する自由がクリエイターに与えられます。

「音圧」のあやふやさ、あやうさ

まず、「音圧」という語句を使用することのあやうさを説明するために、簡単なテストをしてみましょう。

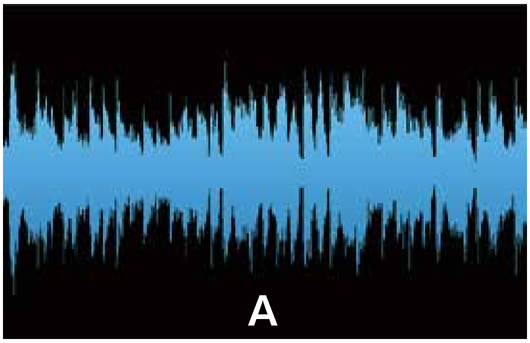
Q1. 次の 2 つの波形 A, B をご覧ください。どちらがより「高音圧」だと思いますか？

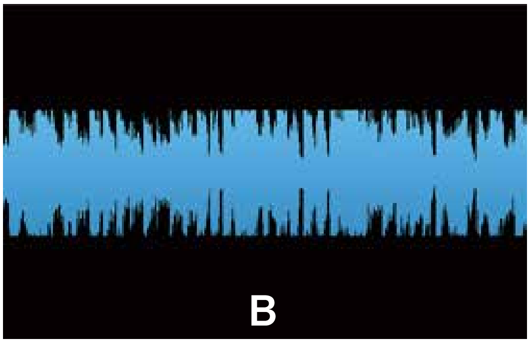




設問が理解できる方であれば、おそらく大半が B を選ぶでしょう。

Q2. では、次の 2 つの波形 A, B の場合、どちらがより「高音圧」でしょうか？





上記 2 つの音源は、いずれも -14LUFS に揃えられています。

<u>A…抑揚がある波形の方を「高音圧」だと思ったあなた</u>

RMS 平均の最大値を基準とするあなたは、これまで「EDM は音圧がキモ」といった表現に違和感を感じたことはありますか？

<u>B…リミッティング後にゲイン・リダクションされた波形の方を「高音圧」だと思ったあなた</u>

クレスト・ファクタの低さを基準とするあなたは、「配信サイトが楽曲再生時に音圧を操作する」という表現に違和感を感じたことはありますか？

A、B のどちらを選んだにせよ、もし上記のような違和感を感じたことがないのであれば、残念ながらあなたもまた、シチュエーションごとに「音圧」の意味が変容することを許してきたことで、「音圧」という言葉が機能不全に陥ったことへの責任の一端を担う…かもしれません。

※本章の続きは本誌にて