

とーくばっく

～デジタル・スタジオの話～



David Shimamoto
Studio Gyokimae

目次

1部PCM録音の話

1章ビット・レートの話

2章ディザリングの話

コラム：有効ビット長の測定

3章サンプル・レートの話

4

8

14

16

2部収録レベルの話

4章収録レベルの話

5章メータの話

コラム：ビット長の現状

6章ダイナミック・レンジの話

コラム：K-Systemとは

コラム：ラウドネス測定手順

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？

24

38

41

42

48

52

54

3部ノウハウ集

7章M/S処理と留意事項

8章YouTubeラウドネス規定導入後のミックス手法

9章抑揚をアレンジに活かす

10章補正可能時代の収録ボーカリストが知るべきこと

58

66

68

72

付録A 随筆集

A-1 対訳のない制作用語

A-2 CD-Rブランドで音が変わるか？

A-3 波形とは？海苔波形とは？きざみ海苔とは？

A-4 ボーカルへの「返し」

A-5 Vo. フェーダ・オートメーションのススメ

A-6 ハイレゾの話

76

79

84

89

92

94

付録B

B-1 ビットレート変換の実験手順

B-2 エイリアス・ノイズの実験手順

B-3 市販曲ラウドネス一覧

B-4 手動ラウドネス・ノーマライズ用・早見表

97

98

99

102

参考文献

あとがき

103

105



1章 ビット・レートの話

DAW やポータブル・レコーダをはじめとするデジタル機器に、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 (ADC) により数字の列に変換されます。これは早い話、音声信号の瞬間瞬間の電圧値を、毎秒何万回と測定する (サンプリングする) 作業です。読者の皆様も、右の図 1 のような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。

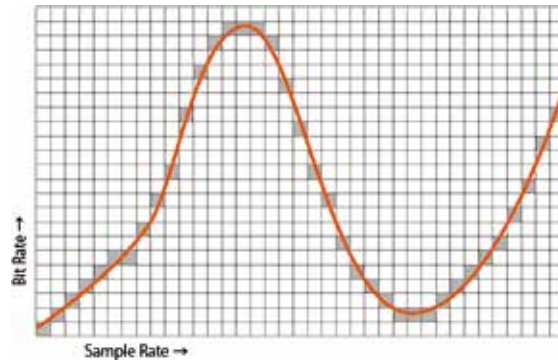


図 1 PCM 録音の概念図

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・レート (Bit Rate)

グラフの縦軸の解像度で、いかに細かい精度で電圧を測ることができるかを表します。ビット・レートが高いほど方眼紙のタテの目が細くなり、1 ビット増えるたびに解像度 (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 (サンプリングする) 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・レートと DAW の内部構造についてお話しします。

ビット・レート = 「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号を持つ整数値です。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。ゼロ以上の数値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「グラフの目が細くなる」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。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

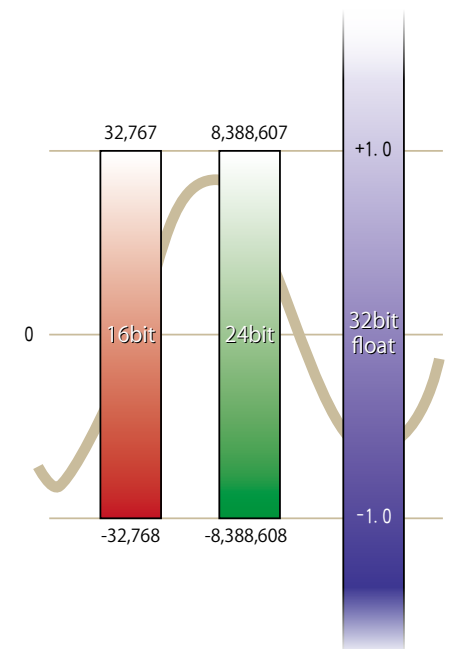


図 2 ビット・レートごとの表現範囲

*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かった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・レート）に話を絞ります。

第1章「ビット・レートの話」にもあったように、アナログ信号のデジタル化に際しては瞬間瞬間の信号電圧を測定するわけですが、それぞれの値は整数値（方眼紙のマス目のいずれか）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、それを望むのは難しい話です。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、こうしてもっと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^{ねつぞう}」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。あとで紹介するように、この現象を逆手にとれば、本来そのビット・レートでは再現できないほど微小な音をよい具合に捏造することで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・レートを落とすとき

＝ 元よりも解像度の低い方眼紙に、信号レベルを落とし込むとき

結論を先に言うと、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・レートを落とすとき」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・レートを落とす（ダウン・コンバートする）と、近似する異なる数値が、同じ数値に丸め込まれるという問題が生じます。これは音声に限らず、デジタル画像で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を、4 段階…白、黒、その中間の灰色 2 色に落とし込む場面を考えます。人物の写真などは変換後でもそれと認識できるでしょうが、おそらく肌などは「のぺー」っと平面的になってしまうでしょう。音声信号の場合は、乱暴な言い方をすると信号波形が歪む分、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) といいます。また、その際に加えるノイズが**ディザ・ノイズ** (Dither Noise) です。実はディザの適用範囲はコンピュータを使用する場面に限らず、桁数の不足などで計算精度が確保できないために誤差が生じる、すべての分野に及びます。^{*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。

コラム：有効ビット長の測定

16/24bit の音声ファイルが DAW のバスを流れるとき、信号が一旦 32bit float に変換されることは、1 章「ビット・レートの話」で、またビット・レートを下げるときはディザ処理が必須であることは 2 章「ディザリングの話」でそれぞれ説明しました。

実は、DAW で 16/24bit の音声を出力する際、ディザリングを行わない方がよい場合がひとつだけあります。それは、バスを流れる音声信号がすでに**クオンタイズ**されていることが予めわかっているときです。

無限の解像度を持つアナログ信号を記録する作業を、方眼紙のマス目を埋める作業に例えたとき、信号の電圧値にドンピシャで対応するマス目がなければ、近いマス目に丸め込むことになり、誤差が大なり小なり生じることは先に述べたとおりです。

アナログ信号ではなく、すでに 16/24bit で記録されたデジタル信号を、**32bit float に変換したあとそのまま元のビット・レートに戻した場合**は、ドンピシャで対応するマス目が存在するので、丸め込みや誤差は生じません。

「16/24bit → 32bit float → 元と同じ 16/24bit」の変換については、**信号レベルに変更を加えていない場合に限り**、きちんと元の値に戻ります。このように 32bit float の信号でありながら、すべてのサンプルがより低いデータ・レートにドンピシャで対応できる状態を、信号が**クオンタイズされている**と言います。MIDI シーケンサにおいて、発音のタイミングを拍子ぴったりに合わせる機能の名称としてもお馴染みの語句です。

では、ディザが不要になる状況とは、具体的にどんなものがあるでしょう？

.....

ディザが必要かどうか迷う場面

最後に、ディザ処理を行うべきかどうか、原則が適用されないため判断を迫られる場面をいくつか考えてみましょう。

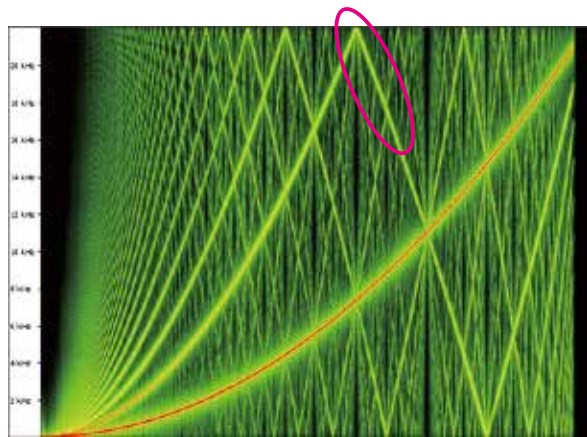
ここまでの話を面倒と思われたかもしれませんが、難しく考える必要はありません。ディザリングが必要かどうかは、ファイルの書き出し前に「目視確認」すればよいのです。



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。

ソノグラムを比較しているファイルは制作工程が違っただけで、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

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。高いサンプル・レートを使用するほど、この問題は回避できるといえます。

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（倍音）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。

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

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 を越えない) 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、これといった基準はありませんでした^{*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 (=スレッシュホールド) を絶対に越えることがないように加工するツールです。スレッシュホールドを越えた信号はきわめて自然にレベルが下げられ、逆にスレッシュホールドを越えない信号は、一切劣化することなく通過します。これにより、スレッシュホールドを 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 を越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、積極的に近づく必要はなく、まれに近付いてしまったときには絶対に越えてはならない『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

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、商売上不利とされていたため

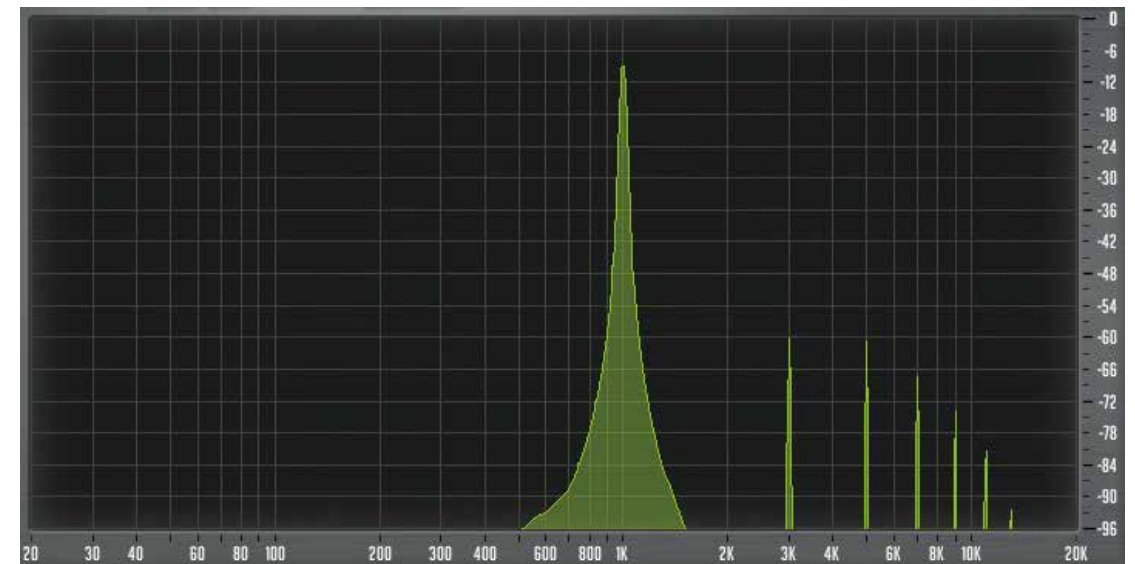
おそらく、マキシマイズが施されたタイトルのほとんどは、この理由が背景にあると思われます。特に似たジャンルで比較対象となる他者の作品が、自分よりも収録音が大きければナーブスになるのは解らなくもありません。

^{*4} デジタル・レコーダで収録可能な最大レベルを越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} 2016 年現在も C D 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。

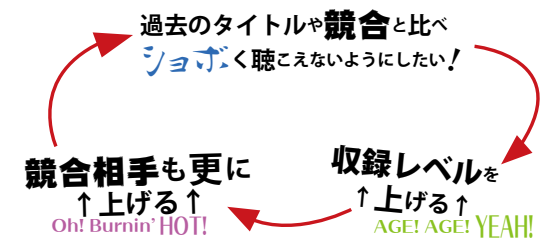
こうしてクリエイターや音楽関係者が互いに聴感上の音量を競い合った結果、CD の収録レベルは際限なく上昇を続けました。さきほど「マキシマイザが違和感なくピークを抑えられる」と述べましたが、これはあくまでもちょっとしたオーバー・ロードを抑制する程度に使用する場合です。競争が進むほどに市販タイトルの多くはアタック感を失い、歪みを増し続けました。

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。



このように音質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。



音圧戦争に抗うムーブメント

2016年の本稿執筆時点、相変わらず日本では音圧競争に関するまとまった情報はなかなか見つかりません。一方、英語圏の書籍やインターネット上のフォーラムをあたれば、音圧戦争の弊害に関する資料や議論がわんさか出てきます。以下に、音圧戦争に終止符を打つべく、世界で起きたアクションの例をご紹介します。

一方、日本では

国内の某巨大掲示板でも過去に高音圧の是非に関する話題が出ていますが、私が見た限りでは「音圧稼げない負け犬の遠吠え」と一蹴されて議論が終わるか、「そうはいつでもアーティストが…」「わかっちゃいるけどディレクターが…」という非建設的な話に終始するものばかりでした。クライアントとのしがらみなどなく、本来自由に表現活動ができるはずのアマチュア界限でも、聴感レベルを上げるためだけのマキシマイズが「百害あって一利無し」だという考え方はまだまだ浸透していないように見えます。

収録レベルの未来

映画産業や放送（映像）業界には、音声の収録レベルについて厳格な基準が存在します。一方、音楽に特化した媒体であるはずのCDには、そういった強制力のある規定もなければ、共通の推奨値すらありません。では適切な収録レベルなどあるのでしょうか？

しばらく前であれば、個々の好みの問題として片付けられた話でしょう。しかし2012年、業界にある動きが起きたのを機に、状況が一変しました。

『ピーク・ノーマライゼーション』は、制作に携わる読者はご存知のとおり、曲中のピーク値が0 dBFS、あるいは任意に指定したレベルに合うように全体のレベルを増減する処理です。リミッターなどでピークを抑えるほど上限(0dBFS)までの余裕が生じますので、あとのピーク・ノーマライゼーション時に全体の平均レベル…いわゆる音圧が上がります。代わりにダイナミック・レンジは狭くなります。

よく似た語句ですが**ラウドネス・ノーマライゼーション**という概念があります。こちらではピーク（聴感上の音量と実はあまり関係がない）を重視せず、代わりに聴感上のレベル（＝**ラウドネス**）が目標とする基準値に合うよう、曲、または番組全体のレベルを調整します。

そもそも近年まで、ラウドネスを定量的に測る信頼できる手段、ましてや統一基準といったものは存在しませんでした。2012年にラウドネスの測定方法が規格化され、これに基づくレベル管理が欧米の放送局、そして少し遅れて日本のテレビ局でも運用されるようになりました。ラウドネス・ノーマライゼーションによりコンテンツの聴感上のレベルを統一する慣習は、現在多くの音楽配信サービスにまで広がりつつあります。

さて、ここで疑問に思われるかも知れません。

「これって放送の話で、**オンエア**予定のないCDには関係ないよね？」と。

これが実は、そうでもないのです。その理由は、次ページのキーワードにあります。



ある曲のラウドネス評価結果。単位LKFSで表される数値（右上）は、コンテンツや曲全体のラウドネスをIntegralと呼ばれる単一の数値で表す。画像はStudio One2内蔵機能による表示。



EBU R128 のロゴ・マーク



DR12 認定マーク（前ページ）

最後に

ここまで読んでくださった制作サイドの皆様が「収録レベルアップのためだけのマキシマイズ」を控える、あるいはマキシマイズを一切使わないことを少しでもご検討いただければ、私としても書いた甲斐があったというものです。

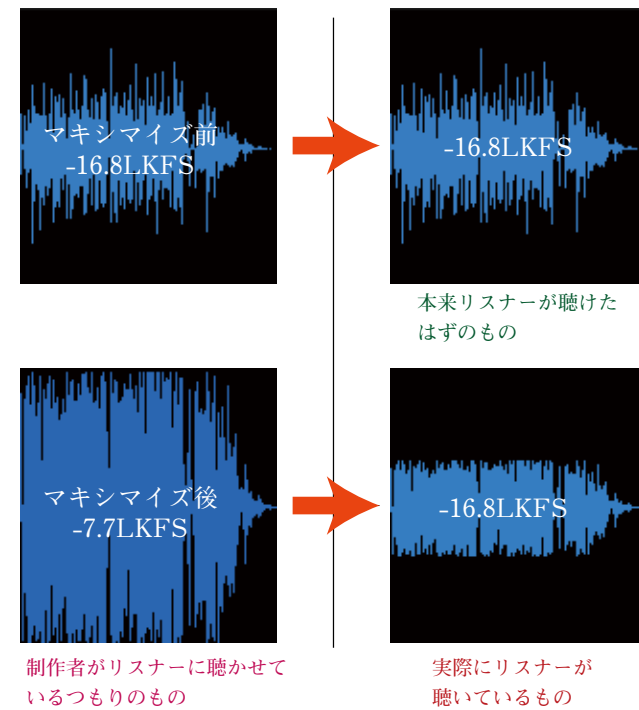
方法論としては比較的新しいものに聴こえるかもしれませんが、以下の基本的事実は、蓄音機が発明されたその昔から今日に至るまで、一切変わっていません。

残念ながら政治的な事情から、あるいは立場上、そういった説明をクライアントやアーティストにするのが難しいこともあるでしょう。

しかし、

- ・もしあなたがエンジニアで、他者の意見に耳を傾けられる賢明なアーティストやディレクターに恵まれたとき、
- ・もしあなたがアーティストで、心無いディレクターやエンジニアから自分の作品を守る必要に迫られたとき、
- ・もしあなたがアマチュアのクリエイターで、収録レベルに迷ったとき、

ここにまとめたお話が、その素晴らしい楽曲を、ありのままにリスナーに届けるヒントになれば幸いです。



5章 メータの話

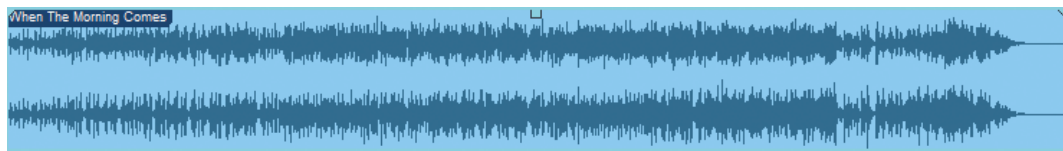
音 声のレベルを測定する方法はいくつかあり、いずれも計測手段や目的が異なるため、それぞれに一長一短があります。DAW ユーザにはおなじみのピーク・メータを例にとってみましょう。



ピーク・メータは、ADC に入力された信号、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たします。



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ができません。

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参考にはならない

たとえば、ADC の最大入力値より 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・レートの話」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「このレベルを 4.41 万分の 1 秒でも越えると、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、絶対的なレベルの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主流になったもののなのです。

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、最小値から最大値までの幅

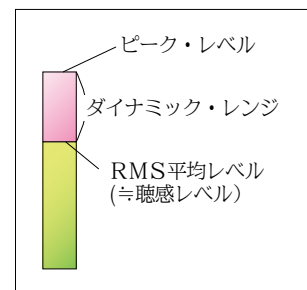
主に、機器やシステムのスペックを語るときに使用されます。

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る余白があるため、パート間の分離がよくなります。

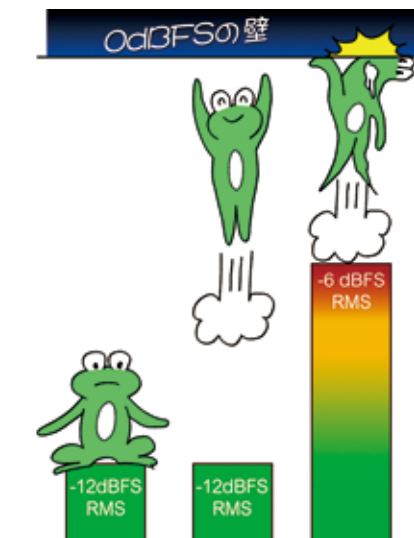
キック（バス・ドラム）のレベルを考えてみましょう。

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

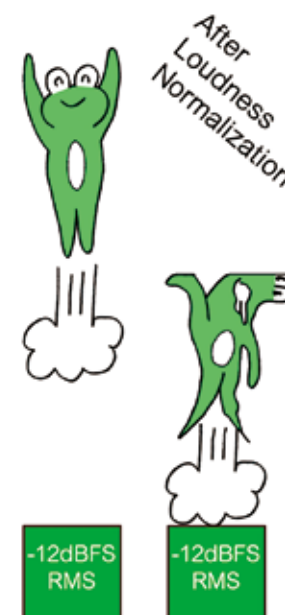
ここで、デジタル媒体に音声を収録する際の RMS 平均レベルと、表現可能な最大ダイナミック・レンジの関係について考えてみましょう。次の図をご覧ください。



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「L4-256」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す、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**作業です。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。

さて、平均レベルが -20dBFS 前後では、編集時はともかく、さすがにパッケージングに際しては現行のポップスとのレベル差が大きすぎます。このような場合、用途に合わせて **K-14**、**K-12** 設定のメータを使用します。

K-14 メータは、ポップスなどダイナミック・レンジが比較的狭い音楽の収録に使用します。

K-12 メータは、放送素材など、さらにダイナミクスの少ない素材を収録する際に使用します。

これら 3 つのメータの違いは、0dBFS までのヘッドルームだけです。

メータを切り替えた場合は、それに応じてスピーカの出力音量も同時に調整します。例えば、使用するメータを K-20 から K-14 に変更する場合、これに伴いスピーカの出力も 6dB 下げます。いずれのメータを使用する場合でも、ピンク・ノイズをメータの 0 位置に合わせて再生したとき、リスニングポジションでの音量が各スピーカとも 83dB SPL になるという原則は変わらないということです。

以上の手法を実践すれば「収録レベルをピークに近付けなければ…」という脅迫観念から解放されるため、**サウンドにとって何がベストか**がミックス時の絶対的な基準となり、聴感上の音量を上げるため「だけ」に広がりやダイナミック・レンジを犠牲にする必要がなくなります。



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？

K-System がすでに歴史的役割を終えたと提唱者 Bob Katz 氏自身によって宣言されたことは以前に述べたとおりですが、その核ともいえる「メータは統一仕様のものを使用すべき」とする思想は引き継がれており、ラウドネス測定仕様 R128 を策定した EBU（欧州放送連合）が、**EBU Mode** という、ラウドネス・メータの仕様を公開しています。

究極的には一個人による提案に過ぎなかった K-System とは異なり、こちらは EBU の支持があること、また既に放送の現場で広く用いられている R128 と親和性が高いことから、少なくとも K-Meter よりは今後普及するものと思われます。

2012 年より各国で運用されている R128 や ITU-R BS.1770 は、ラウドネス（聴感上のレベル）の測定方法を定めた仕様ですが、その測定結果をどのようにユーザに表示するかまでは厳密に定めていません。R128/BS.1770 に準拠するラウドネス・メータであることを名乗る上で必須になるわけではありませんが、より踏み込んで**メータの統一仕様**である**"EBU Mode"** 準拠とするには、いくつかの条件を満たす必要があります。



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

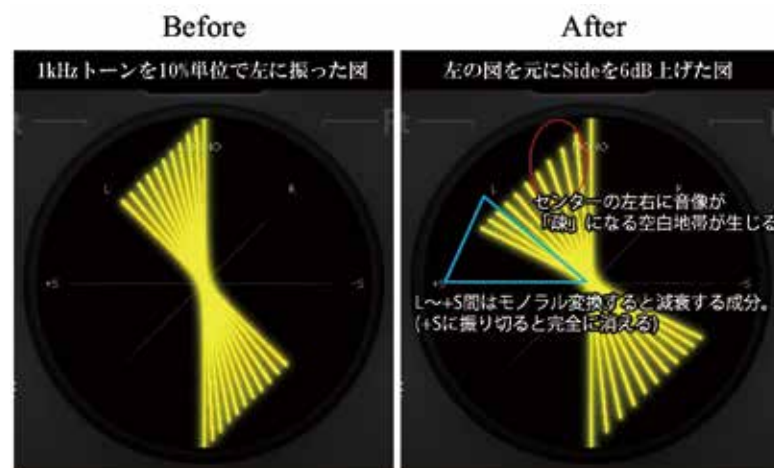
ステレオ音像拡張の副作用

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。

図中のステレオ・メータ(リサージュ、リサージュともいう)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。

さて、左ページの図「Before」は、1kHz のサイン波をステレオ・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。



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、**ステレオ・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（さもなくばモノラルに聴こえます）、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。

8章 YouTube ラウドネス規定 導入後のミックス手法

2015年3月、YouTubeでラウドネス・ノーマライゼーションを開始したよという情報が飛び交いました^{*1}。どうやら前年の2014年末から翌年初頭にかけて、いつのまにかそのようにシステムがアップデートされていたようなのです。世界中のユーザがいろいろな実験を重ねた結果、掲載動画の音声レベルが-13LUFS/LKFS^{*2}付近に揃えられているそうです。これは簡単にいうと、アップロードされた動画の音量が一定の聴感レベルに揃えられるようにYouTube側で調整されるようになったということです。再生する動画ごとの極端な音量差がなくなるのは、ユーザにとっては利便性が増します。これまで聴感レベルを上げるため「だけ」にマキシマイズを行った音楽作品を投稿していた場合、再生時にYouTubeのシステムに音量を下げられることになります。音質が犠牲になることに目をつぶってまで収録レベルを上げていた場合に至っては、デメリットしかありません。

また、2016年10月には、TIDALが配信レベルを-14LUFSに統一することを公表しました。他にも同様の処理がiTunes Radio、Spotifyでも配信時に施されていることは、4章「収録レベルの話」で述べたとおりです。

音楽配信サービスが次々とラウドネス・ノーマライゼーションを開始する中で、投稿者の意思に反してレベルを下げられることなく、また音質を犠牲にしないようにするために、ユーザはどのようにミックスやマスタリングを行えばよいのでしょうか？

本章はBob Katzが推奨するK-Systemでのミックス手順をベースにアレンジを加え、YouTubeに最適化した作品をミックスする際の具体的手順を紹介します。

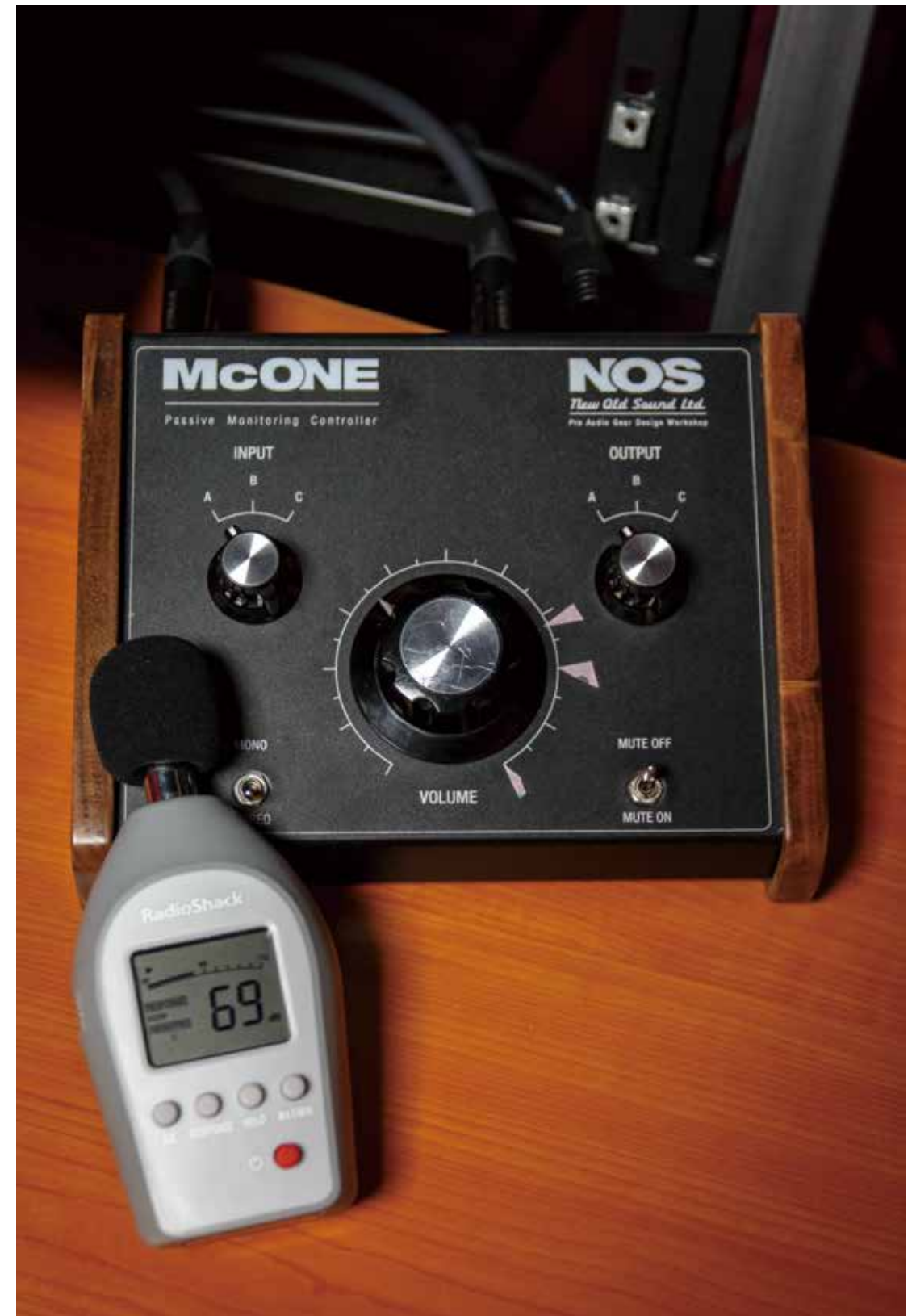
1. 準備物

まずはこの2つを準備します。



*1 <http://productionadvice.co.uk/youtube-loudness-normalisation-details/>

*2 LUFS/LKFSについては、4章「収録レベルの話」をご参照ください。



9章 抑揚をアレンジに活かす

この章まで関心を持って読んでくださった方は、次に曲作りに取り組む際にはさっそくダイナミック・レンジをフルに活用した、グルーヴ溢れるミックス / マスタリングを実践してみたいとお考えではないかと想像します。ここまで、K-System の運用手順などを通じ「VU メータは通常 0 VU 付近、強いパッセージで +4 ～ 6dB 付近」を 1 つの開始地点として提案してきました。しかし、具体的には曲のどのセクションで、メータがどのぐらい振れるのが理想的なのでしょうか？ ここから先は、正解はありません。極端にダイナミック・レンジが広いと、場合によってはハイファイなリスニング環境でないと逆に聴きづらくなることがあるのも事実です。しかしどう転んでも、完全にピークが抑えられ、曲中の音圧変化がビートに追従しない平面的なミックスより表現豊かになることは間違いないでしょう。

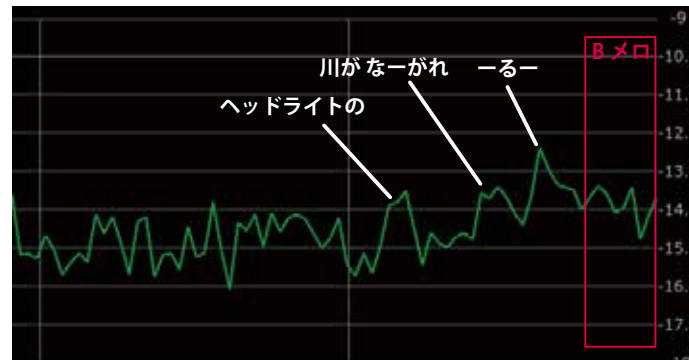
とはいえ、少ないヒントだけをお渡しして荒野に置き去りにするのも申し訳ないので、この章ではわずかながら、実例より先人方の足跡に学んでみることにします。

以下、ラウドネス・メータを使用し、アレンジに沿ってメータの表示がどのように動くか、具体的な市販曲を例に見てみます。よく考え抜かれたミックスは、メータが曲の各セクションで示す最高地点も、きちんと曲の抑揚に追従していることが目でも確認できます。

作例 1 PSY・S「Angel Night ～天使のいる場所」アルバム「Non Fiction」収録バージョン (1988)

タイトルに覚えのない方でも、アニメ「City Hunter2」のオープニング曲と聞けば、アラフォーのみなさんはピンと来るかもしれません。ボーカルの圧が耳で感じ

られるのと同様にラウドネスも全体的にボーカルの出入りに追従しているのがグラフからもわかります。右図は A メロの終わり、次のセクションへの期待を煽るべく音節ごとにレベルが増していっているのがわかります。これがフェーダ操作によるものなのか演奏者の技量によるものなのかはわかりませんが、いずれにせよ現在風の高音圧ミックスがされていれば、すべて平らに均されたフレーズだったのではないかと想像します。同じような傾向は右図、B メロの終わりでも見られます。サビに向けて一音一音が徐々に大きくなり、盛り上がりきったところでサビ頭のフレーズ「Angel Voice」が最高値となります。



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、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が、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あ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つち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。とかちつくちて。

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「ああ編集しているなァ」と思わせます。



付録 A-2 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断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つたらどうしよう？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、クスッと笑ったり、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あ？」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たせてもらえないうちに次期・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 枚のなんと摩訶不思議な盤に直面します。



付録 A-3 波形とは？ 海苔波形とは？ きざみ海苔波形とは？

市販タイトルの収録レベルについてインターネット上での議論に目をやると、よく波形だの海苔波形だのという語句を目にします。本書でも、これまでに音声信号波形の画像をたびたび掲載してきました。

しかし、そもそも波形とはなんのことでしょう？

どのような波形が海苔波形であり、何が問題とされているのでしょうか？

初歩的な話ではありますが、信号波形の見方や解釈についてたびたび誤解を目にすることがありますので、このセクションでは主に波形ウォッチング(?)を始めて間もないリスナーの方々を対象に、これらの疑問に答えてみようと思います。ついでに、最近…海苔波形の次のステージでしょうか…私が耳にするようになった**きざみ海苔**サウンド（ついさっき勝手に命名しました）についてもユルく考察してみます。

これまでの章に記載した内容をだいたい理解されている方は、直接きざみ海苔の項 (p.88) まで進んでいただいてもよろしいかと思います。

そもそも波形とはなにか

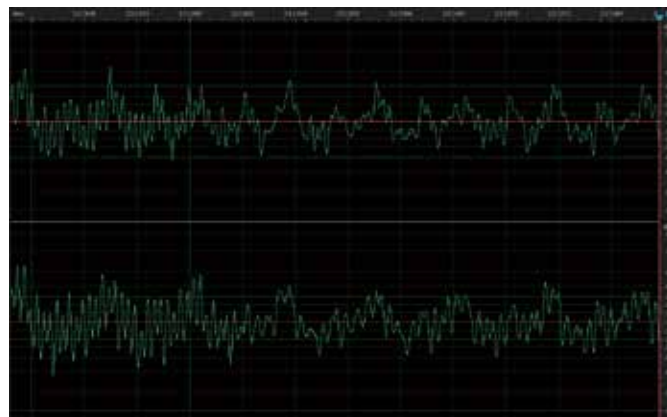
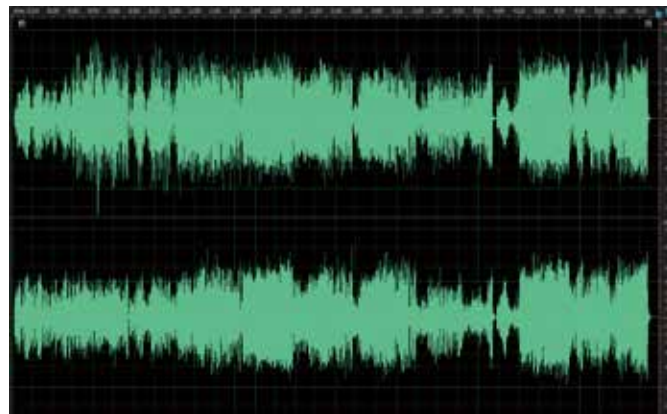
広義の波形とは「時間にともなう数値の変化をグラフに描いたもの」です。レコーディング界限で波形と呼ばれるものは、音声ケーブルを流れる信号レベル（電圧）の高低を表します。

右の図をご覧ください。

PC ベースの音楽編集ソフトは、編集対象となる音声を視認できるよう、このように波形を表示します。しかし、これは PC モニタの制約から、時間軸方向に極端に縮小されたものです。

先の波形を拡大表示すると、実際には下図のように、上下にウネウネと揺れ動く一本の線であることがわかります。

ちなみに右の波形は、時間にして 60 ミリ秒 (0.06 秒) の間に起こる変化にすぎません。一般的な音声信号は 1 秒間に何千、何万回と上下に揺れるわけですから、まして数分間の曲ともなれば、波形のすべてを単一の線として PC モニタの解像度で描くことは到底不可能です。このため、曲全体の波形が表示できる程度に縮小すると、右上図のようにブロックとして表示されます。



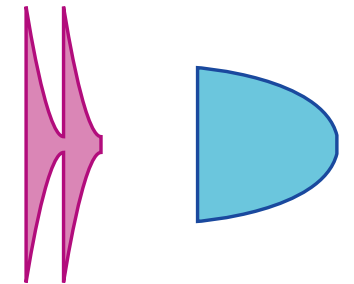
ちなみにいずれの波形も上下 2 段で表示されているのは、左と右の 2 チャンネル分の信号を並べているためです。

信号レベルはゼロを中心に上下に変動し、ほとんどの場合、上下対称になります（シンセサイザーなど、電氣的に音を生成する機材では上下いずれかに偏ることもあります）。

さて、波形を拡大表示したときに見られた 1 本の揺れ動く線。先ほどは信号レベルの高低と書きましたが、これは同時に、スピーカ・コーンがどのように前後に動くかを図示したもののでもあります。一般に中心点のゼロより上が前方、下が後方を表します。

アナログ変換された音声信号は、指示書たる波形の動きのとおりスピーカ・コーンを前後に振動させます。その振動が空気に伝わり、空気振動がリスナーの鼓膜に伝わり、ヒトははじめて音を知覚できます（あるいは低音であれば、空気振動は直接身体に伝わり、感じられます）。

聴感上の音の大きさは、スピーカが前後に動いた幅ではなく、移動量の時間平均で決まります。つまり、右図、左の波形のようにスピーカが瞬間的に大きく前後運動しても、さほど大きな音には感じられませんが、右の青い波形のようにそこそこの範囲を長時間継続して運動していれば（≒波形の面積が大きければ）より大きな音に感じられます。

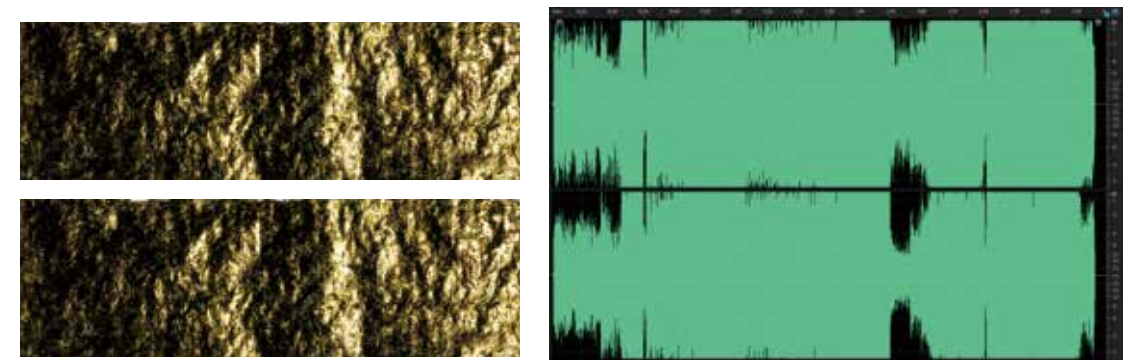


海苔波形と呼ばれるもの

音声における波形の意味がだいたいわかったところで、海苔波形と呼ばれるものの特性を考察してみましょう。

聴感上の音量が波形の面積である程度決まることは、先ほど述べたとおりです。一般に海苔波形と呼ばれるものは、再生時の音が大きく聴こえるよう、信号の加工により波形の面積を広げたものです。曲のはじめから終わりまで、信号レベルはゼロを中心に上限値、下限値の間を絶えず往復します。このため、下図のように波形を縮小表示したときに海苔のようなベタ塗りの矩形になることから、このように形容されます。

こういった波形をもつ楽曲は、信号レベルの時間平均を測ることを目的とする「VU メータ」が比較的動かないという特徴があります。これに対して、90 年代ぐらいまでの楽曲であれば（聴感上の音量の違いはさておき）VU メータがビートに合わせて上下することが観察できます。



さて、このような海苔波形を生み出す加工に対する批判の中身は、おおよそ次のようなものです。

本書第2部でも述べたとおり、海苔波形の音源は聴感上の音量が大きいことから、えてして再生時には相対的にボリュームが絞られ、p.37の図のように波形がタテに縮小することになります。前述のように、波形はスピーカの動きを示す指示書でもあるわけですから、振れ幅が一定であるということは、必然的にスピーカの可動範囲も制限されることになります。

海苔波形のできるまで

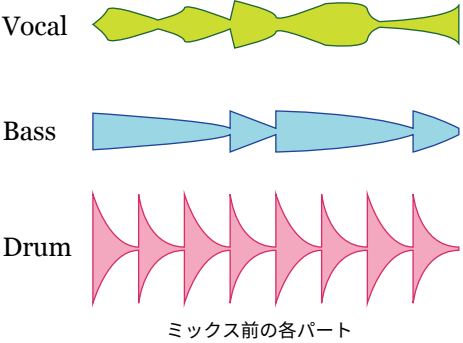
すべての市販タイトルの制作過程を見たわけではないので想像の域を出ませんが、海苔波形への加工は多くの場合、マスタリング段階で2MIX（各パートのバランスを調整して、一旦ステレオ2chに整理したもの）に対して行うことで実現しているものと思われます。

平たくいうと順序としては、各パートがよいバランスになるよう整える作業である「ミックス」を終えたあと、続く「マスタリング」という工程で、信号レベルが曲を通して均一かつ最大になるよう、処理を行っているものと思われます。

元のパートごとの波形とマスタリング後の波形を比べると、だいたい下の図のようになります。

20世紀型マスタリングの信号波形

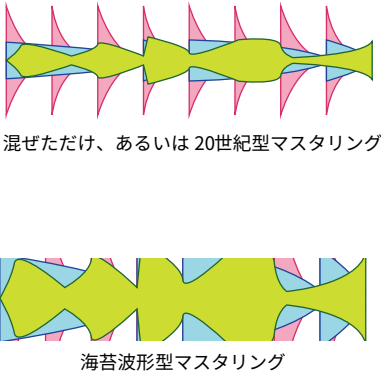
各パートが元々持っている抑揚は、基本的にマスタリング後も維持されます。また、全体のレベル変動は、鼓膜が感じる圧の変化に連動します。たとえばドラムが鳴るタイミングで鼓膜は一段強く振動しますし、ボーカルの声の張り具合によっても、鼓膜の振動量が変化します。これは例えではなく、実際80年代末までの録音のよい盤であれば、ある程度大きな音量で再生したときに、ボーカルが声を張り上げる箇所が一段強い圧として感じられ、その圧の変化に連動してVUメータが動くのが見られます。右図でも、ボーカルのレベルを示す緑色の部分が、波形全体の面積変化に寄与しているのがわかります。



音圧競争時代の波形

ミックス時点での各トラックは、先の例と変わりません（右図、「混ぜただけ〜」）。ミックスを終えてステレオ2chに整理した信号に対して、瞬間的にレベルが高くなる部分を下げたのち、全体のレベルを持ち上げることでレベルの時間平均を上げています。

スピーカの振動範囲、またVUメータの移動範囲は曲の大部分を通して一定になります。





付録 A-4 ボーカルへの「返し」

ボーカルを録音をする際のマイクの選定方法や立て方については、少し探せばネットでも書物でも様々なアドバイスが見つかります。しかし「返し」と呼ばれる録音時のモニタの音作りについては、あまり初心者向けの情報は出てこないように思います。筆者自身、講釈を垂れるほどの現場経験はありませんが、僭越ながら歌好きでもある私なりのこだわりポイントをご紹介します。特に初学者で、ご自身ではあまり歌われない方々にとってヒントになれば幸いです。



付録 A-6 ハイレゾの話

96kHz/24bit といった高ビット、高サンプル・レー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・レートまたはサンプル・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。変換作業の理不尽さを認識しながらも、所属レーベルより 96kHz 納品を求められたため、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、メーカー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ー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タイトルです。この手の商品が問題になるのは、たとえばその変換時のアルゴリズムが、変換を行う技術者の主観が入る余地が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、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「残っているアナログ・マスターはハイファイだが、44.1kHz/16bit や CD 登場初期の A/D コンバータでは取りこぼしがあった」ためにハイレゾで再販される商品です。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。

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。

