

About MQA (Master Quality Authenticated)
新デジタルコーディング方式 - MQA (Master Quality Authenticated)
Bob Stuart, Keith Howard 訳：鈴木 弘明 (株式会社 ソナ)

当寄稿は JAS ジャーナル 2015 年 9 月号に記載した Bob Stuart 氏と Keith Howard 氏による *About MQA (for JAS)* の和訳です。従って英文オリジナルとの差異が生じた場合は英文オリジナルに順じます。また、参考文献は英文オリジナルを参照ください。

序論

1982/3 年に CD が登場して以来、「デジタルオーディオは本来の性能を発揮していないのでは？」という不満の声が絶えない。AD/DA の性能、過度のジッター、ディザの不当な印加、などの初期の諸問題が解決されてきたにも拘らず、16bit/44.1kHz、16bit/48kHz のシステムによるデジタルの音に対する不満の声は続いてきた。

今では、ハイレゾのデジタルオーディオはサンプリング周波数、量子化数を上げていくことで、音質を向上させる、ということが広く知られているが、コーディングの効率という面ではコストがアップすることになる。24bit/88.2kHz のレコーディングは、16bit/44.1kHz に比べるとデータレートは 3 倍となるし、更にサンプリングレートが 176.4kHz、352.8kHz と増えると、2 の倍数で増えていく。音質が向上するのは歓迎すべきことだが、データレート、データ容量の面では、負荷が飛躍的に増大することになる。

人間の聴覚

聴覚の研究においては、サンプリングレートを上げることで音質が向上するという事実に対して二つの明確な理由が指摘されている。

第一に、良く知られている人間の可聴上限周波数が 20kHz であることは良く知られているが、人間は超音波領域の周波数を感知でき、20kHz 以上の周波数の音が適正なレベルで入っていると、心地よく感じる、という研究発表もある。楽器が発生する超音波成分は、聴感上重要である。

第二は、最近の聴覚の研究では、アンチエイリアスと再構成フィルターの時間領域特性が音質の劣化をもたらしているという事が裏付けられている。デジタルオーディオで用いられるローパスフィルターが音色を持つ、という報告もある。

こうした知見は、人間の聴覚はある音が本人にとって“危険な音”か、“安全な音”かを、瞬間的に識別しなければならないという本能と密接に関連している。一方、視覚の場合は、人間は周囲全体 (360 度) を見渡せるわけではないから、音に対するほど人間は敏感というわけではない。

聴覚は、危険を予知する上で、最も重要な感覚であり、音の方向と距離を瞬時に知ることが肝要なのである。また直接音と短時間遅れた音、時間的に近接した反射音を識別する能力も重要である。そのため、時間間隔=分解度は周波数、帯域とは独立していると言えるのである。

人間の聴覚は周波数に対するよりも時間に対して鋭い。この時間に対する感度の鋭さは、音楽の受聴とも密接に関連している。時間・周波数の関係を最も良く感知出来るのは、音楽家であるという実験結果もある。時間に対する感覚の鋭さは、音楽家になる過程で強まり、磨かれるのである。従来、オーディオ・システムの性能を周波数領域で評価してきたが、これは基本的に音楽に対する感覚とは相容れないのである。これからは、オーディオのエンコード、装置の設計においては、時間領域の性能をより考慮した新しいアプローチを試みなければならない。

新たな取り組み

MQA (Master Quality Authenticated (マスターの品質「音質」の保証：本文中で明らかになる) は、前述の要素に対する新たなアプローチである。単なる新技術というだけでなく、ハイレゾリューションの音の再生の新たな“哲学”とでも言うべきものである。人間の聴覚に基づいた最適性能基準となるものであり、スタジオのマイクロホンから DA コンバーターまでの全録音/再生行程を包括した技術=哲学である。

アナログ及びデジタルのマスターの固有の音質を前例のない高忠実度で伝送するのを目的として、MQA は特に時間領域の特性を重視している。

従来、デジタルオーディオの分野で使われてきた急峻なローパス、アンチエイリアス、再構成フィルターは“時間の滲み” “時間のボケ”を増加させる。MQA は、これに対しこれまでにない新たな試みを行い、既存の録音に対しては、時間の分解能は $10\mu\text{s}$ をターゲットにしている。将来は人間の聴覚の時間分解能とされている $5\sim 8\mu\text{s}$ よりも更に高い分解能 $3\mu\text{s}$ が録音音楽に適用されることが望まれる。

96kHz の録音ではサンプリング周期が $10.4\mu\text{s}$ であり、前述のターゲット $10\mu\text{s}$ を満足するように見える。しかし、それはアンチエイリアス、再構成フィルターがシステムのインパルスレスポンスが 1 サンプル時間間隔を遥かに超えて拡張するという悪影響を無視していることになる。

MQA の時間領域の特性を実現するためには録音・再生において介在するフィルターについて理解していなければならない。スタジオのマイクロホンから、家庭での DA 変換までの全行程が問題となるのである。アナログ録音の素材をエンコードする場合はアナログテープ録音機の性能も、考慮されなければならない。

MQA はマスタリング・スタジオでの再生プロセスにおいて適用されるので、最終的に得られた音は、そのままの音質で家庭に届くことになる。再生の際に MQA 対応の機器では LED が点灯し、MQA が正しく再生されていることが表示される。Master Quality が Authenticated (保証) されているのである。

MQA の時間分解能は従来の 24bit/192kHz の録音・再生に比べて、一桁は高い。立ち上がり縁(前縁)は $250\mu\text{s}$ から $3\mu\text{s}$ まで減少し、トータルのインパルスレスポンスの持続時間は $500\mu\text{s}$ から $25\mu\text{s}$ まで減少する。聴覚上の音の“滲み”は少なくとも $100\mu\text{s}$ から $10\mu\text{s}$ まで減少する。MQA のトータルの周波数特性、インパルスレスポンスを Fig.1 に示す。Fig.2 は MQA のレスポンスが、音が空気中を数メートル伝達する際のレスポンスと大差がない、ということを示している。

Fig.1 のインパルスレスポンスを 24bit/192kHz の AD/DA の行程のレスポンスと比較して表したのが Fig.3 である。

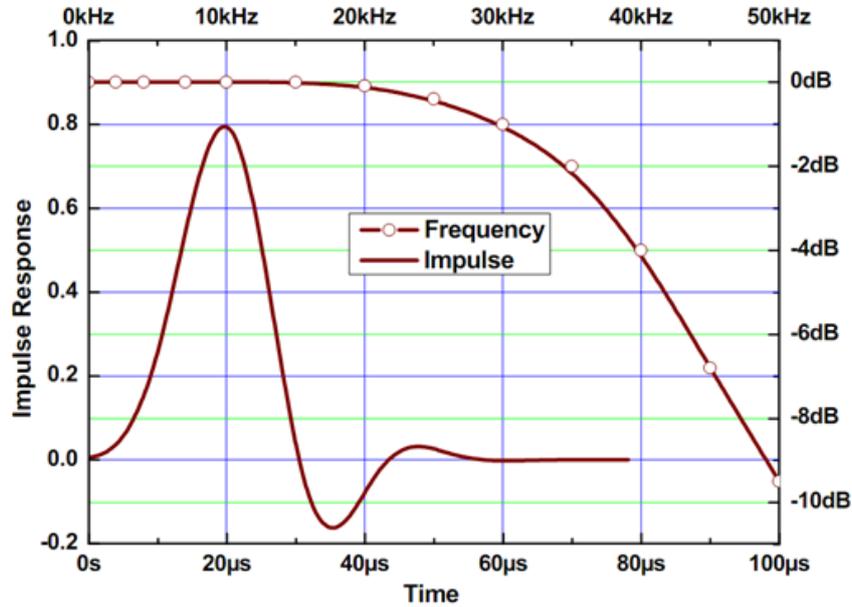


Figure 1. 48kHz 伝送の MQA の周波数特性とインパルスレスポンス

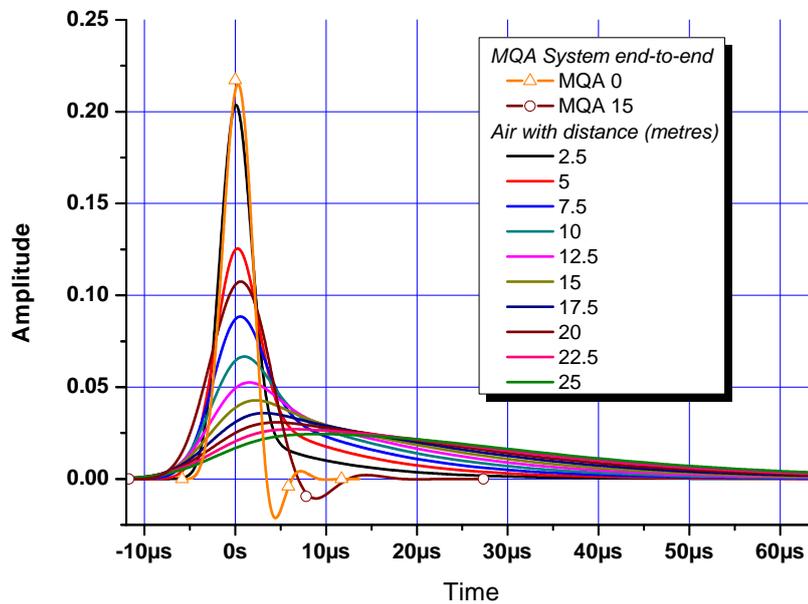


Figure 2. 空气中 (湿度 30%)を数メートル伝播する音と MQA のインパルスレスポンスの比較

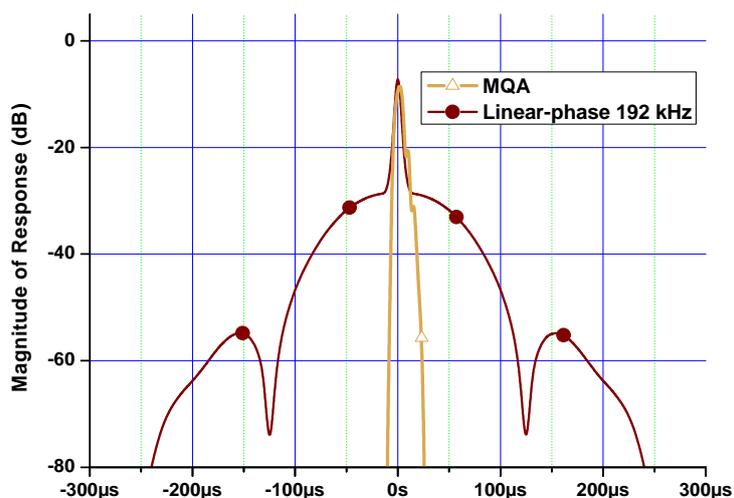


Figure 3. MQA と 192kHz システムのインパルスレスポンスの比較.

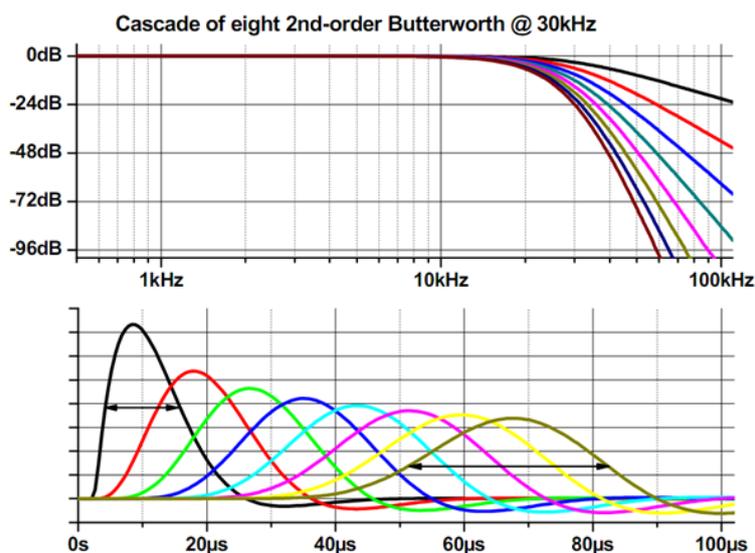


Figure 4. 二次のバターース・ローパス・フィルタ（コーナー周波数が 30kHz）を 8 台カスケード接続した場合の周波数特性、およびインパルスレスポンス

付属装置・設備

MQA により時間特性が格段に向上するので、オーディオの行程においてのアナログ機器が音質に対し、制限的な要素になり得る。従来行われてきた、超音波領域のノイズ、歪みを除去するための高周波数帯域の狭帯域化は必要なくなる。DA コンバーターのローパスフィルタは制御不能で、システムのインパルスレスポンスを劣化させるからである。

高分解能 (High Resolution) の再生

JAS は、High Resolution の規格において、アンプ、スピーカー、ヘッドホンの再生帯域幅を最低でも 40kHz と規定している。これは、DA コンバーター以降の機器によって、音質が制限を

受けることがなくなるということの意味する。

変換器（トランスジューサー）の場合、広帯域だけでは不十分である。しばしばメタルドームツイーターに見られるが、高域では、明らかな共鳴があってはならない。

操作性と音質の関係

Fig.5 は各種オーディオキャリアの操作性と音質との関係の年代による変化を示す。音質は音楽が広く楽しめるようになるにつれ、操作性・簡便性の犠牲となってきている。MQA は 24bit/96kHz, 24bit/192kHz よりも高音質の音を極低レートで提供出来るので、操作性・簡便性と音質のトレードオフの問題は無くなる。MQA はストリーミング配信の場合、転送レートは平均で 1 Mbps と、CD のレート 1.411 Mbps よりも更に低い。

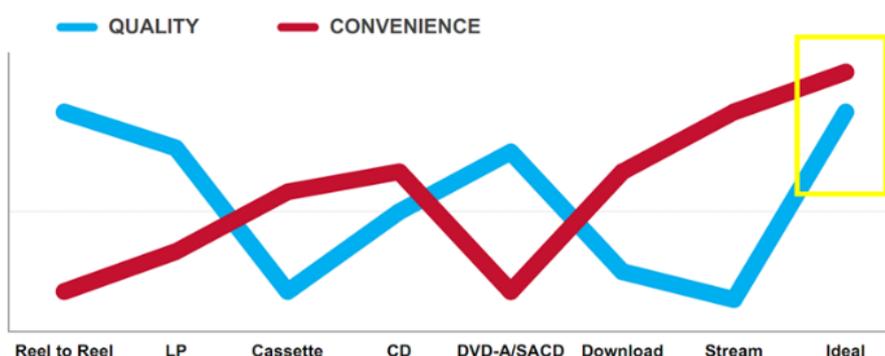


Figure 5. 音楽格納メディア、配信技術の変化と機器の操作性・簡便性との関係を示す。MQA を使えば、伝送、保存、再生の簡便性と音質の間のギャップは極小にすることが出来る

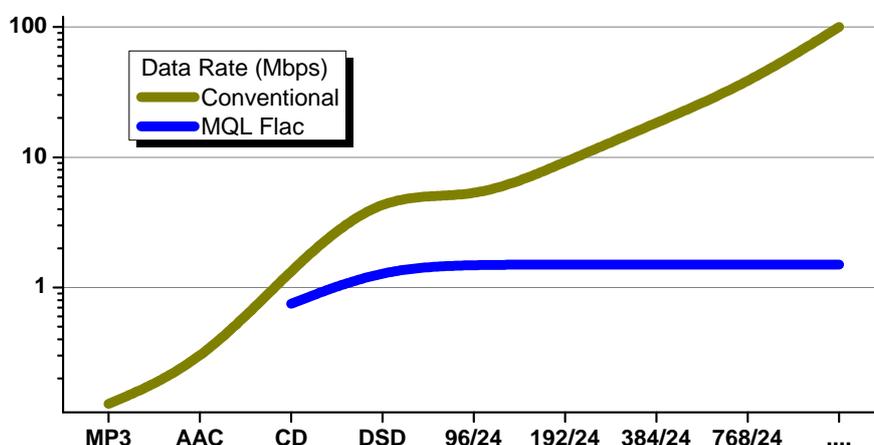


Figure 6. 既存のコーディング方式とMQAのデータレートの比較

MQAのダウンロードファイルはデータレートが平均で1.2Mbpsで、完全なバックワードコンパチ性を保有している。デコーダーが無くても、ファイルの再生は可能である。但し、その場合、MQAの音質はフルに再生されない。16bitの”パイプ”しか用意されていない場合でも、MQAのデコーダーは従来の16bit/44.1kHzもしくは16bit/48kHz以上の音質を提供出来る。

MQA は、時間分解能が高いオーディオ信号をバックワードコンパチの 44.1kHz もしくは 48kHz のファイルに格納するので、従来の PCM オーディオファイル再生用のすべての機器で再生が可能である。

デジタルオーディオのコンバーターとチャンネル

原音を最も良く再生できるのは、マイクロホンを直接スピーカーに接続した場合であるが、それでも完全とは言えない。マイク、スピーカーそのものがロッキーであり、不確定な面を持っているからである。原音を格納、もしくは伝送するための一つの方法が Fig.7 に示すようにデジタル変換であるが、AD および DA のゲートウェイにおいて、極めて重要な音質上の制限が発生するのである。

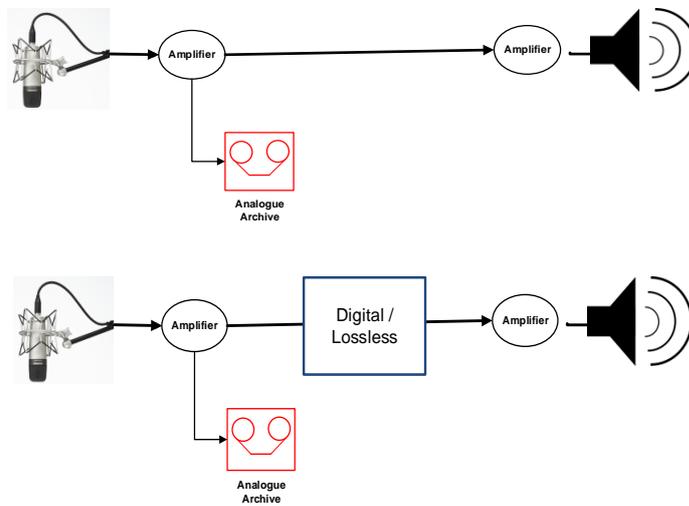


Figure 7. 録音の概念モデル

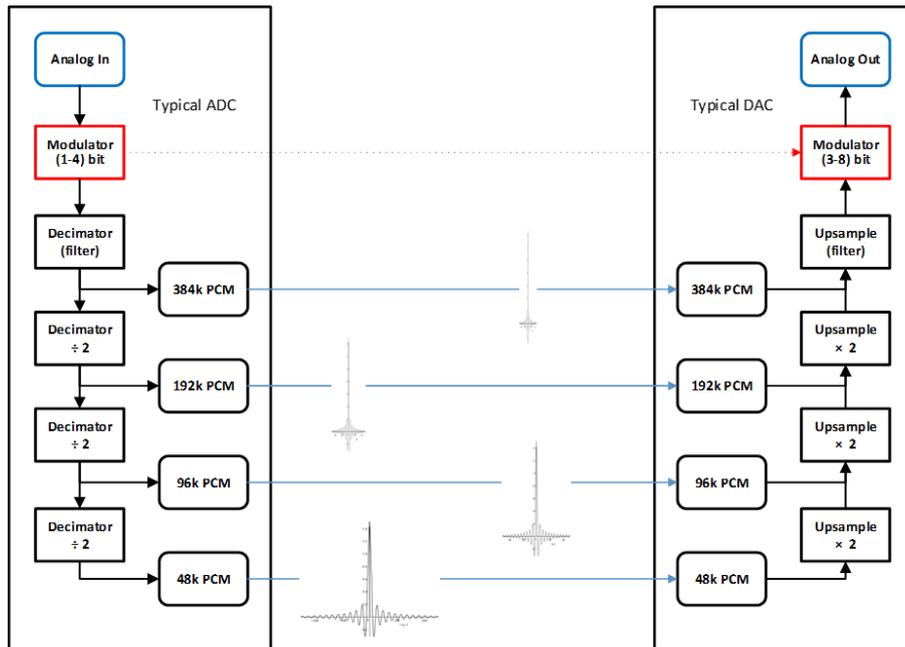


Figure 8. PCM のデルタシグマ AD/DA コンバーターの内部ブロック図

Fig.8 は、最近の AD/DA コンバーターのブロック図である。高速の短語長のモジュレーターが問題なのだが、PCM 信号が逐次伝送される毎にサンプリングレートはモジュレーターのレートの分数で任意に決められる。

信号が AD コンバーターを通ると、各ダウンサンプリングの段階で、量子化ノイズが加わり、時間的な“ブレ・ボケ”が増加し、DA コンバーターの各アップサンプリングの段階で更にノイズが加わり、時間の“ブレ・ボケ”も増加することになる。デシメーション及び、アップサンプリングフィルタの特性と、各段階での処理の品質により、トータルな音質は大きな影響を受けることになる。AD のモジュレーターから DA のモジュレーターまでの信号の流れは決してロスレスではなく、デバイスによるエリアシングや熱変調ノイズ、音の“ブレ・ボケ”などが誘引されるため、リニアでさえないことになる。AD、DA のゲートウェイはアナログのマスター・オーディオ信号を伝送するに当たっての主要な障害となっている。

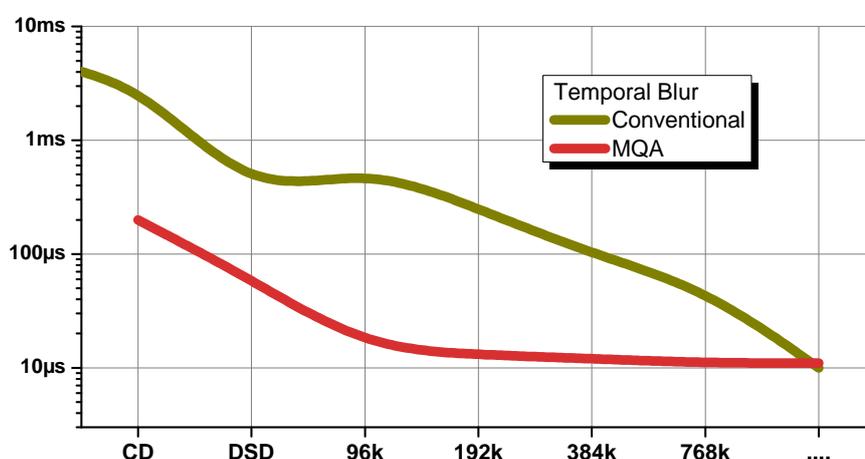


Figure 9. 従来のシンク・カーネルのサンプリングと MQA での再構成サンプリングでの音の時間分解能の比較

高サンプリング周波数で伝送すれば、処理段階数は減り、フィルターによる音の“滲み”は少なくなるが、これは転送レートが高くなる場合の広帯域化のメリットとは無関係である。

理想的には、これらのアップサンプリング、ダウンサンプリングを避けるためにはマルチビットコンバーターとモジュレーターを直接接続すればよい。しかし、ハードウェアから最も正確なアナログ品位の忠実再生を引き出すことを目的に、MQA は、ディザをうまく組み合わせ、注意深くエンコード/デコードのフィルター・カーネルを選ぶことで、この直接接続を仮想的に作り出すことが出来る。Fig.10 は CD、24bit/96kHz、32bit/96kHz、などの伝送フォーマットのコーディング空間を比較した図で、DSD のノイズフロアも表示している。

32bit LPCM のコーディングのためには、縦軸のスケールは膨大なものとなり、上限の 120dB は聴感上、痛みを感じる閾値であり、0dB（人間が検知出来る最も静かなレベル）が縦軸の中央に位置している。32bit のチャンネルのノイズフロアは静寂レベル（人間が検知出来る最も静かなレベル）よりも 120dB 低くなっている。ビット数を上げていけば、縦軸（音圧）方向の精度は上がることになるが、チャンネル自体の情報量は、膨大なものになるのである。

Fig.10 中の四本の曲線が示しているのは：
 赤；ノイズに対する人間の聴覚の閾値（この曲線以下のレベルの音は直接には聞こえない）
 緑；静かな屋外の環境騒音レベル
 青；最も静かな録音スタジオの騒音レベル
 茶；マイクロホンの熱雑音の限界値（この曲線以下のデータはブラウン運動、ということになる）

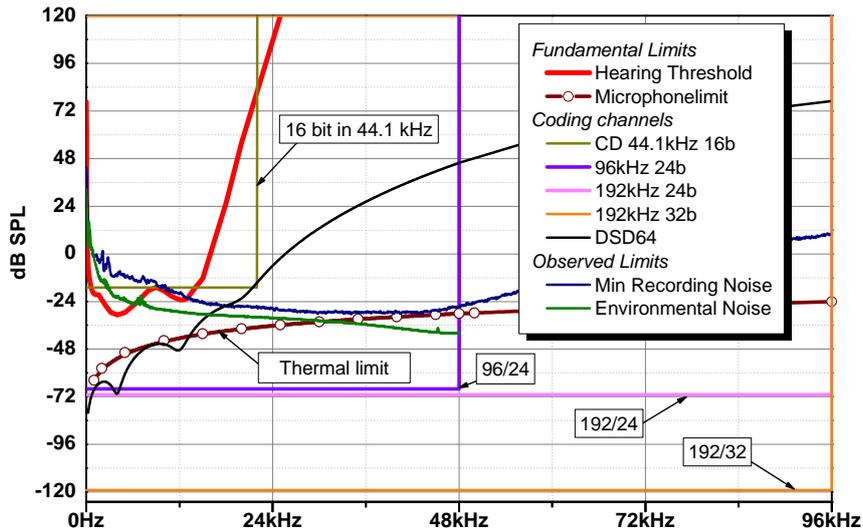


Figure 10. 各種デジタル・チャンネルのデータ量を示すシャノン・ダイアグラム
 （面積がデータレートになる）

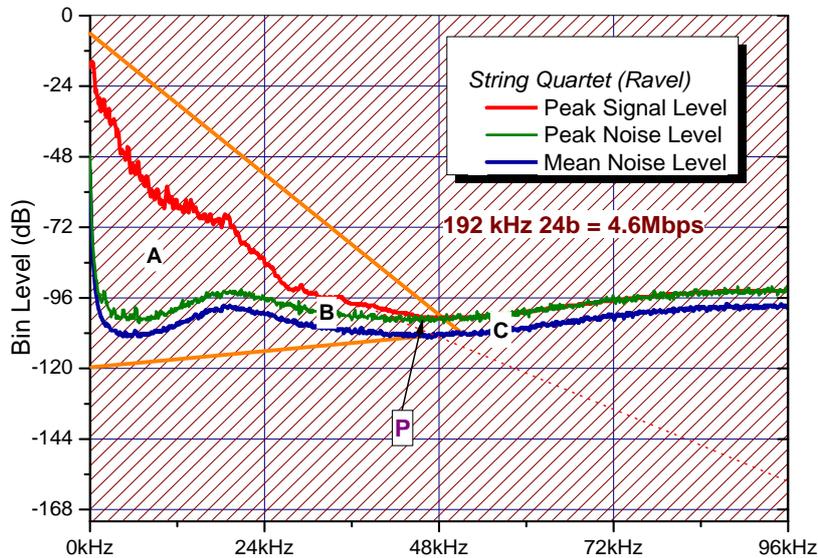


Figure 11. 24bit/192kHz の録音サンプルの信号レベル、ノイズレベルを示すシャノン・ダイアグラム

データ伝送

Fig.11 のシャノン・ダイアグラムは、24bit/192kHz のソース信号（この場合はラベルの弦楽四重奏）に対しての全コーディングスペース（データ量）を示したもので、赤の曲線が周波数スペクトラムのピーク値を表している。緑の曲線、青の曲線はそれぞれ録音環境の暗騒音のピーク値、平均値を示している。暗騒音は、音楽再生の前・中・後に聴くことになるもので、録音空間の環境騒音、及びアナログのノイズである。

ラベルの弦楽四重奏を選んだのは、自然な音響環境で、実際の楽器の音を捉えた録音であるからである。弦楽器のピチカート音は再生が難しく、スペクトラムは高調波成分が 20kHz よりも上まで存在している。

マイクロホンがピックアップするピークレベルは周波数が高くなるにつれて小さくなるが、これは自然音の特徴であり、うまく利用すれば時間的なボケ、転送レートを低下させることが出来る。

音楽信号の曲線とノイズ曲線は P 点で交わっているが、この交差周波数より上の周波数の信号成分はノイズに埋もれることになると考えられる。

“A”領域は従来の（CD の）オーディオ帯域で、20kHz を上限として、音を聴いていることになる。“B”領域にも音楽成分が存在しているが、この帯域を切り出して聴いても何も聞こえない。しかし“B”領域の成分は時間的な分解能、音のエンベロープに大いに影響を与えている。経験的にも、この成分を取り除くと音質が低下するということが知られている。

“C”領域は、A、B とは全く異なる領域であり、目立った音楽信号は皆無で、帯域はほとんどすべてのマイクロホン、スピーカーの周波数帯域よりも上の帯域で、レベルは人間のノイズに対する聴感の閾値を上下している。

この定常的なノイズは全く聞こえないが、サンプリング周波数を上げていけば、“A”領域の分解能は向上し、音の”ボケ”も小さくなる。

各種録音では、P 点は 30kHz から 60kHz の間、40kHz 付近に位置することが多い。従って、サンプリング周波数を 96kHz 以上、例えば 192kHz、352.8kHz にすれば、“C”領域はより広い帯域にまたがることになる。音楽に関連した成分はすべて、ノイズレベルよりも十分小さいレベルであるが、転送レートが高くなっていることにより、フィルターによる音の“ボケ”が小さくなり、分解能も向上することになる。

Fig.11 のオレンジ色の三角形は、音楽に関するすべての信号成分（残る成分はノイズと静寂信号である）を包含していて、全コーディング領域（データ量）空間の約六分の一の面積を占めている。つまり、データレートの六分の五は浪費されていると言えることになる。

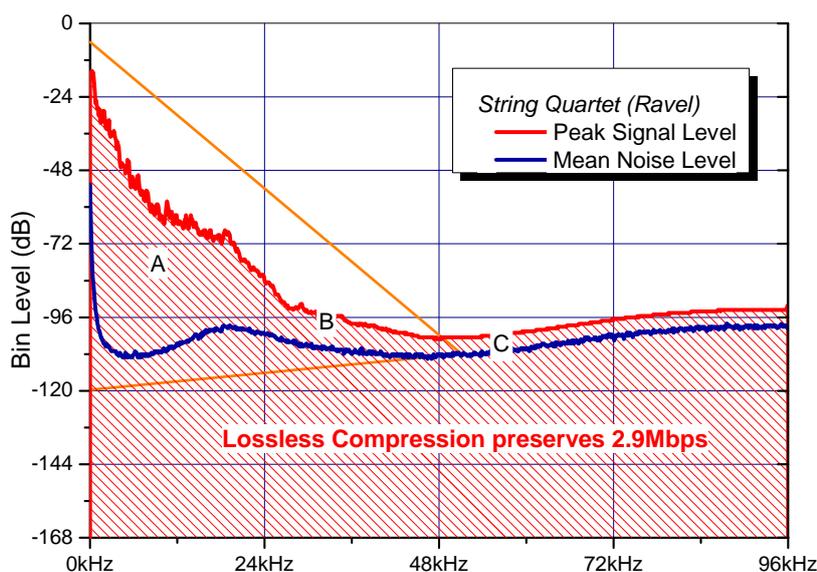


Figure 12. MLP ロスレス圧縮の使用によりデータレートが 4.6 から 2.9Mbps に低減

高性能のロスレス圧縮を使えば、データレートを 2.9Mbps まで低減させ、データ量を 37%減らすことができるが、それでもオンラインのストリーミングに対するデータ量としては過大であるといえる。

MQA は、Fig.11 のオレンジ色の三角形の中身のデータのすべての伝送をコンバーターの時間の”ブレ”やノイズ変調なしに正確に行おうとするものである。

16bit/44.1kHz のデジタルファイルはロスレスで伝送出来るが、音質も問題なく伝送されるということが保証されるわけではなく、むしろ劣化する、と言わなければならない。

従来のハイレゾデジタルについても、16bit/44.1kHz よりも音質面では上であると考えられるが、時間領域での人間の聴覚の鋭さに対応しているとは言えない。

従って、MQA の目標は、狭義でのロスレス処理（コアのデジタル・パスはロスレスで、ビット単位でデコーダーにおいて照合される）にあるのではなく、人間が聴くことのできるすべてを時間の”ボケ”なしに、またコンバーターの量子化の不確定性とは無関係に伝送し、マスタリングスタジオでのアナログ音をなんら損なうことなく、最終ユーザーに届けることにある。これこそ、より深い意味での”ロスレス”伝送である。

音楽の折り紙

MQA が高サンプリングレートの信号を 44.1kHz もしくは 48kHz のサンプリングまで低減させる処理を日本の折り紙細工にちなんで”音楽の折り紙”と呼ぶことにする。Fig.14~16 には、192kHz の音楽ソースに対するこの”音楽の折り紙”の作用を示す。

“折り紙”プロセスは、二次元的に表示しているが、実際には三次元的なプロセスである。最初の“折り” (Fig.13) で、伝送レートが 192kHz から 96kHz に低減し、二番目の“折り” (Fig.14) では、96kHz から 48kHz に低減することになる。この“折りたたみ”プロセスはフィルタリングではなく、固有のサンプリングレートとビット長はそのままである。”折たたむ”毎に得られる

情報（データ）は増えることになる。MQA では、最初の“折り”プロセスを”エンキャプシュレーション”と呼ぶ。

この“エンキャプシュレーション”プロセスは、階層的にスケラブルであり、ソースが 352.8kHz のファイルであった場合は、Fig.14 に示すように、三回”折り”を行い、最終の伝送レートは 44.1kHz となる。同様にソースが 96kHz であった場合は、Fig.14 のロスレスのプロセスから始めることになる。MQA も階層的にスケラブルであるため、各“折り”をターゲット高く行い、転送のスピードを倍に上げることも可能である。

Fig.14 - 15 に示す二回目の“折り”はロスレスで行うことが可能である。オリジナルのアナログ信号のノイズ・フロア（青色の曲線）が示すように、24bit のエンコーディングで得られるダイナミックレンジの大半がノイズで占められるからである。二回目の“折り”プロセスでは、データ埋め込みの手法を用い、24kHz 以上の信号情報を 24kHz 以下のノイズフロア中に埋め込んでいることになる。

録音環境のノイズ、マイクロホンの自己ノイズ（併せて、アナログ・マスターのテープ・ノイズ）などがあるため、音楽の録音では、16bit 以上のダイナミックレンジを得ることすら難しいのが現実である。これに加えて、人間が聴くことの出来るノイズ中の信号はノイズレベル以下 10dB まで（Fig.13 のオリーブ色の曲線）であり、19bit から 24bit では重要な情報はなんら伝えられないということになる。

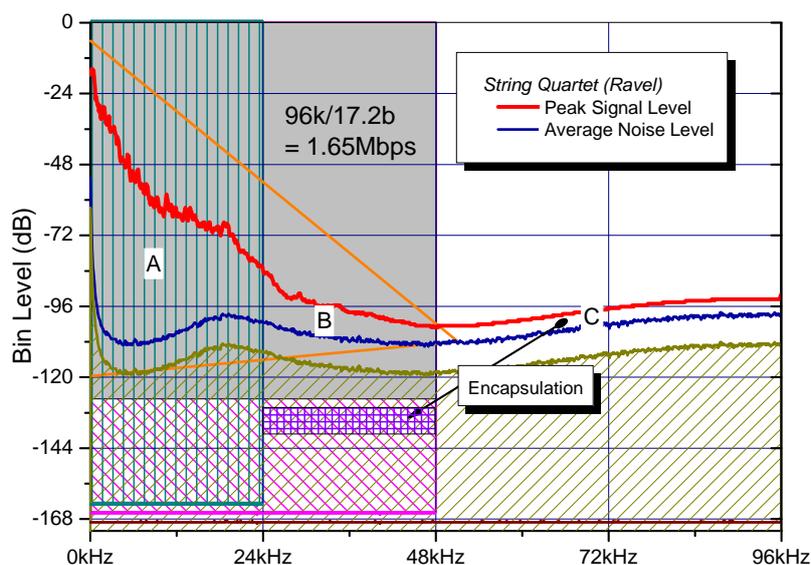


Figure 13. 最初の”折り”プロセスは E (Encapsulation)。転送レートが半分となり、C 領域の 信号成分は B 領域のノイズレベル以下に格納される。

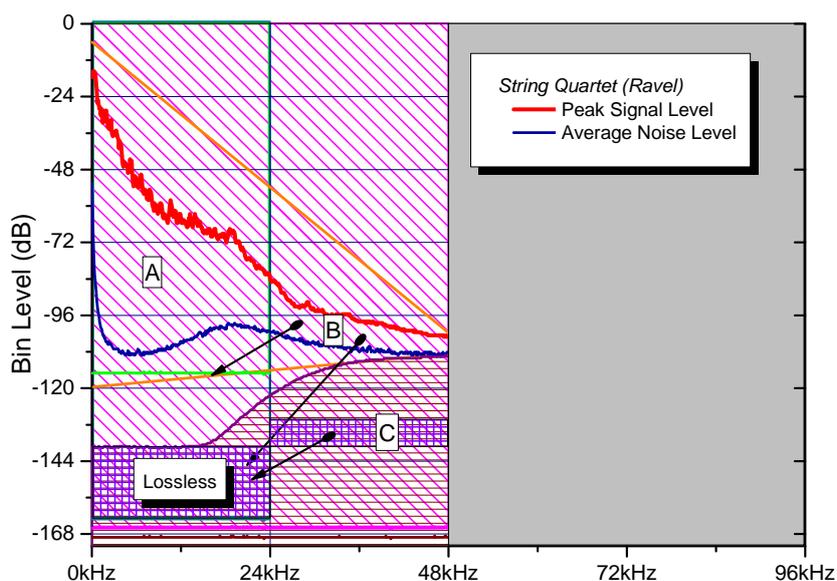


Figure 14. 二番目の”折り”プロセスはL (Lossless)。転送レートは更に半分に低減。B領域の信号はCの信号と共にベースバンドのスペクトルのノイズ以下にロスレスで埋め込まれる。

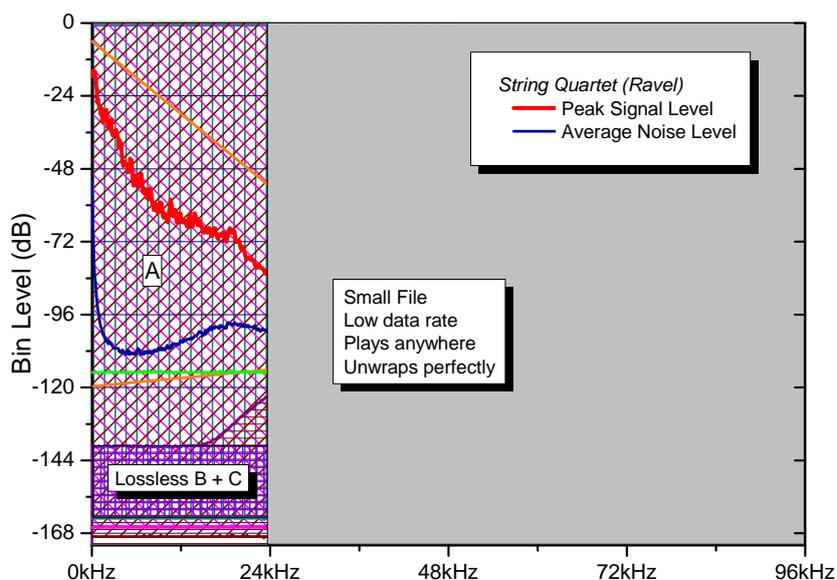


Figure 15. 最終的な MQA のダウンロードファイル

再生時にはデコーダーにおいて、二段階の“折り”プロセスが完全に元の状態に“開かれる”ことになる。デコーダーがない場合は互換性のあるベースバンド信号が再生される。

Fig.13~15 の図には表示していないが、MQA では、革新的なプロセスを用いて、“折り”とデコーダーによるロスレスの”開き”の処理を行っている。

例えば、環境と音楽信号に合った、新たなサンプリング・カーネルを使って人間の聴覚に適合した処理を高時間分解能で行っている。

MQA ではサブトラクティブディザがデコーダーで作用するため、他のフォーマットとは違い、ノイズ変調がなく、また、ロッキーなコーデックとは違い、MQA は決して音楽信号に影響を及

ぼすことはなく、音楽信号の特性を厳密に保ち続ける。

再生

MQA のデコーダーでは、前述の二回にわたる”折り紙”のプロセスが”開かれ”、24bit/192kHz の出力が Fig.1 に示す周波数特性、インパルスレスポンスで完全に再現される。デコーダーは更に、D/A コンバーターへの入力信号をビット対ビットの精度で再構築し、マスタリング・スタジオでの音と同じ音を再現する。また MQA は、再生音がマスタリング・スタジオでの音であることを認証し、その結果は再生デバイス上に表示されるようになっている。

デコーダーが利用出来ない場合は、ファイルはバックワードコンパチにより 44.1kHz もしくは 48kHz で再生される。

伝送路が 16bit にしか対応していない場合でも、MQA のデコーダーはベースバンド (Fig.10 中の領域 A) とロッキーな領域 B (及び C) を再構築するので、再生音はハイレゾのオリジナルに近く、CD よりも格段に良いということになる。

同様に、再生デバイスが高サンプリングレートに対応していない場合は、折り紙を“開く”プロセスを止め、デコーダーがインパルスレスポンス、周波数特性の最適化を行う。

まとめ

MQA は、従来相容れないものとされてきた、“音質”と“利便性・操作性”という二つの特性を兼ね備えた音楽再生を実現させようとするものである。現行のハイレゾのデジタルフォーマットを遥かに凌ぐ音質の音楽再生が実現出来るが、高転送レートを必要としない。しかも、デコーダーを持たないユーザーでも、音楽再生が楽しめる。MQA 音楽は従来の PCM と同様に配信が可能であり、ストリーミング、ファイル・ダウンロードにも対応し、CD などの光ディスクに搭載することも可能である。

MQA は、世界中の録音エンジニア、音楽プロデューサー、レコード会社の間で好評を博している。これまでにプレス、メディアへの発表・デモを行ってきたが、MQA はデジタルオーディオの未来を代表する技術であると、広く称賛されるようになってきている。

以上