

ВОЛОГДИН Э.И.

**ДИНАМИЧЕСКИЙ ДИАПАЗОН
ЦИФРОВЫХ АУДИО ТРАКТОВ**

Конспект лекций

Санкт Петербург
2012

Содержание

1. Динамический диапазон звуков и музыки	3
2. Динамический диапазон фонограмм.....	5
3. Динамический диапазон цифрового аудио тракта	7
4. Расширение динамического диапазона с использованием технологии Dithering	10
5. Расширение динамического диапазона с использованием технологии Noise Shaping	15
.....	15
Список литературы	17

1. Динамический диапазон звуков и музыки

Человек слышит звук в чрезвычайно широком диапазоне звуковых давлений. Этот диапазон простирается от абсолютного порога слышимости до болевого порога 140 дБ SPL относительно нулевого уровня, за который принято давление 0,00002 Па (рис.1). Зона риска на этом рисунке обозначает область звуковых давлений, которые при

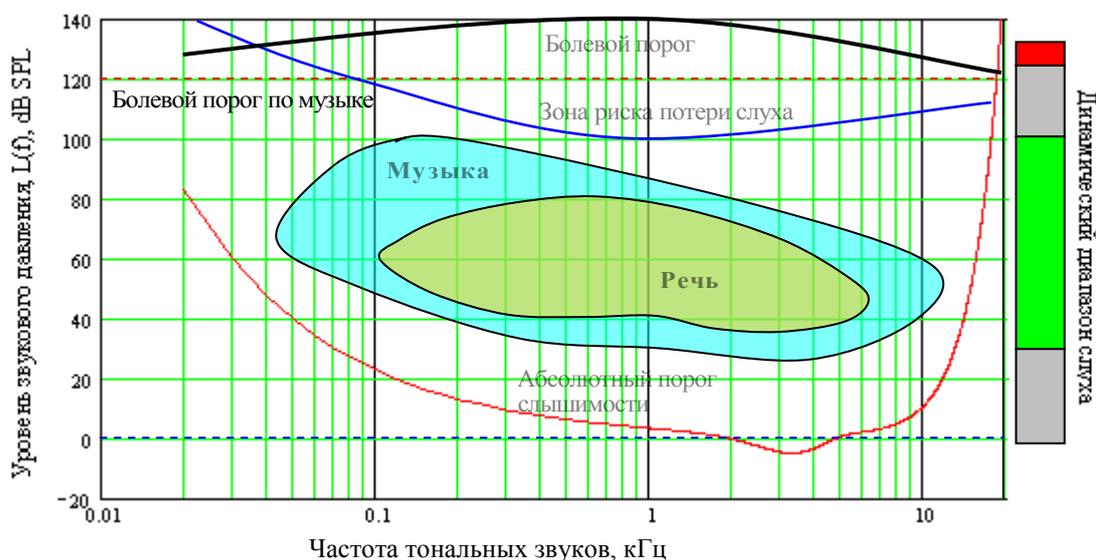


Рис. 1. Области слышимости слуха

длительном воздействии могут привести к полной потере слуха. Болевой порог для тональных звуков зависит от частоты, для звуков с произвольным спектром за болевой порог принят уровень давления 120 дБ SPL. График абсолютного порога слышимости достаточно точно описывается эмпирическим равенством

$$L(f) := 3.64 \cdot \left(\frac{f}{1000}\right)^{-0.8} - 6.5 \cdot e^{-0.6 \cdot \left(\frac{f}{1000} - 3.3\right)^2} + 10^{-3} \cdot \left(\frac{f}{1000}\right)^4$$

В тишине чувствительность слуха человека повышается, а в атмосфере громких звуков — понижается, слух адаптируется к окружающей звуковой среде, поэтому *динамический диапазон слуха* не такой большой — около 70..80 дБ. Сверху он ограничен давлением 100 дБ SPL, а снизу шумом с уровнем -30...35 дБ SPL. Этот динамический диапазон может сдвигаться вверх и вниз до 20 дБ. Для комфортного восприятия музыки рекомендуется, чтобы звуковое давление не превышало 104 дБ SPL в домашних условиях и 112 дБ SPL, в специально оборудованных помещениях.

Динамический диапазон музыки определяется отношением в децибелах самого громкого звука (фортиссимо) и самого тихого звука (пианиссимо). *Динамический диапазон симфонической музыки* составляет 65...75 дБ, а на концертах рок-музыки он возрастает

Табл.1. Динамические диапазоны музыки и голоса	
Источник звука	Динамический диапазон, дБ
Большой симфонический оркестр	60...80
Малый симфонический оркестр	40...50
Духовой оркестр	30...40
Эстрадный оркестр	35...40
Хор	30...45
Голос (основной тон)	25...35

до 105 дБ, при этом пики звуковых давлений могут достигать 122...130 дБ SPL. Динамический диапазон вокальных исполнителей - не превышает 35...45 дБ (табл.1).

Динамический диапазон музыки существенно зависит от выбора максимального звукового давления P_{max} , так как он ограничивается снизу абсолютным порогом слышимости. Эта зависимость наиболее сильно выражена на краях звукового диапазона. На рис. 2 приведены примеры изменения динамического диапазона тональных звуков. В

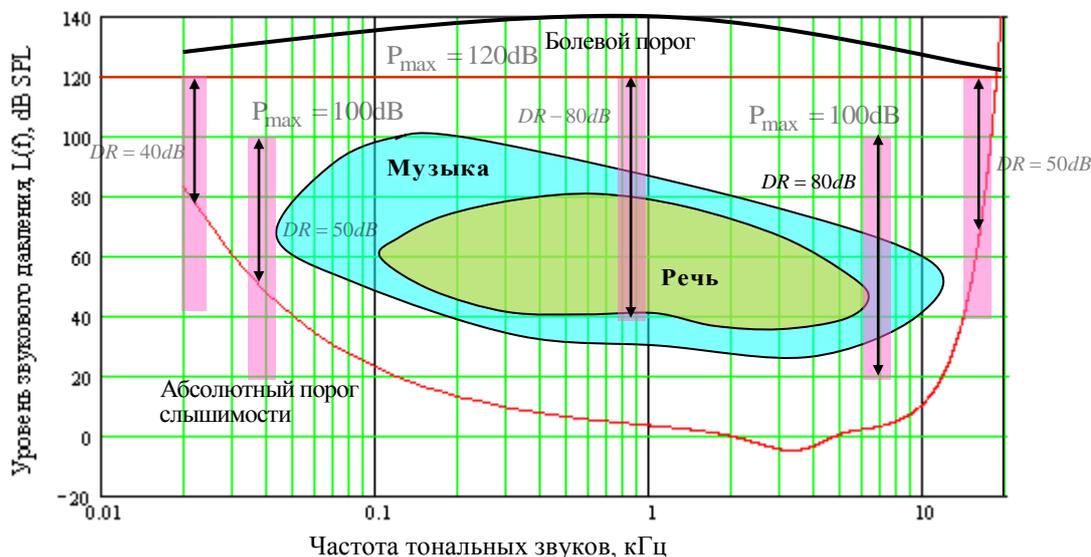


Рис. 2. Динамический диапазон музыки и пороги слышимости слуха

зависимости от выбора P_{max} и частоты тональных звуков динамический диапазон звука 80 дБ уменьшается на краях звукового диапазона до 40 – 50 дБ. Именно поэтому принято измерять динамический диапазон звуков на частоте 1 кГц, на которой он может достигать 117 дБ.

Шум помещения маскирует звук и этим уменьшает его динамический диапазон музыки снизу. На рис.3. показано как при уменьшении звукового давления от 120 до 80 дБ SPL динамический диапазон музыки из-за шума помещения уменьшается с 90 до 50 дБ.

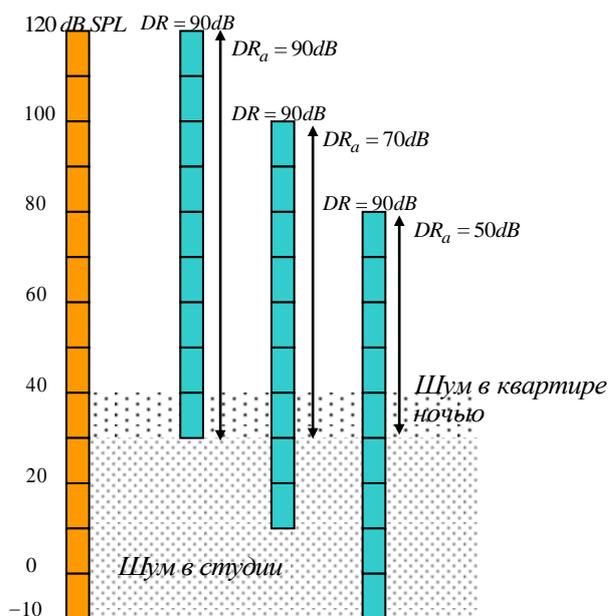


Рис. 3. Динамический диапазон музыки в студии

Влиянием шума можно полностью пренебречь только при условии, что его уровень на 10...20 дБ SPL ниже минимального уровня музыкальных звуков. В студиях звукозаписи уровень шума не превышает 20...30 дБ SPL, ночью в квартирах «тихих» домов этот уровень равен 40 дБ SPL, любой разговор увеличивает уровень шума до 60 дБ SPL. Именно поэтому тихая музыка часто тонет в шумах помещения прослушивания и невольно возникает желание увеличить громкость.

Шум квантования, являющийся белым шумом, заметен на слух при его интенсивности всего 4 дБ SPL, даже когда общий шум аудио аппаратуры в помещении достигает 50 дБ SPL. Эти цифры необходимо сопоставить с тем, что полной шкале FS цифрового измерителя уровня соответствуют уровень между 105 и 112 дБ SPL. Поэтому для бытовых помещений динамический диапазон музыки не должен превышать 101 - 108 дБ.

Динамический диапазон микрофонов определяется так же, как это обычно делается в электрических трактах. Верхняя граница ограничивается допустимой величиной нелинейных искажений, а нижняя - уровнем собственных шумов. Современные студийные микрофоны допускают максимальное звуковое давление 125...145 дБ SPL, при этом нелинейные искажения не превышают 0,5% ... 3%. Уровень собственных шумов микрофонов составляет 15...20 дБА, динамический диапазон – от 90 до 112 дБА, а отношение сигнал/шум- от 70 до 80 дБА. Эти микрофоны с запасом перекрывают весь диапазон слуха человека от 120 дБ SPL до уровня шума студии 20 дБ SPL. В современных студиях запись производится с использованием 22 или 24 разрядных АЦП, иногда используется квантование с плавающей запятой, поэтому проблем с динамическим диапазоном не возникает. Стоит такая аппаратура крайне дорого.

2. Динамический диапазон фонограмм

Музыкальный и речевой сигналы представляют собой последовательность быстро нарастающих и более медленно затухающих звуковых импульсов (рис.4). Такой сигнал характеризуется *среднеквадратическим и пиковым значениями уровней*, разность этих уровней называется *пик-фактором*. Прямоугольная волна (меандр) имеет единичный пик-фактор 0 дБ, пик-фактор синусоиды равен 3 дБ. Фонограммы музыкальных и речевых сигналов имеют пик-фактор до 20 дБ и более. *Время определения пик-фактора* связано временем интегрирования при вычислении среднеквадратического значения сигнала, и обычно, оно равно 50 мс.

Динамический диапазон и пик-фактор музыкальной фонограммы определяют путем статистической обработки *мгновенных* значений сигналов. Наиболее подробно статистические характеристики рассчитываются в звуковом редакторе Audition 3 (рис.4).

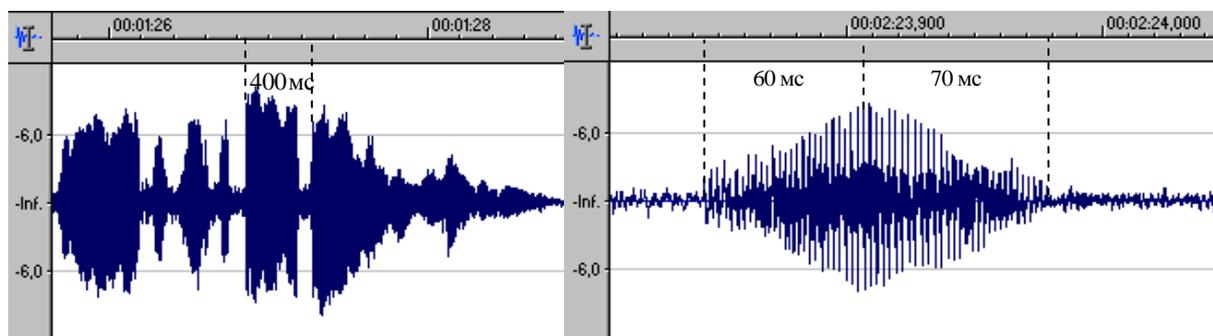


Рис.4. Фрагменты фонограмм музыкальных отрывков различной длительности

Из них основными являются следующие: Peak Amplitude (L_{pic}), Maximum RMS Power (L_{max}), Minimum RMS Power (L_{min}) и Average RMS Power (L_{avr}) (уровни максимального, минимального и среднего среднеквадратического (эффективного) значения мощности сигнала).

Динамический диапазон фонограммы по данным этой таблицы определяется как

$$DR_m = L_{pic} - L_{min}$$

пик-фактор рассчитывается по формуле

$$PF_m = L_{pic} - L_{avr}$$

Динамический диапазон может быть также определен по гистограмме распределения уровней фонограммы, приведенной на рис.5. Такие операции удобно быстро делать до и после динамической обработки фонограммы.

Время интегрирования при измерении уровней гистограммы сильно влияет на ее вид. Но это время никак не регламентируется и выбирается самим исполнителем в

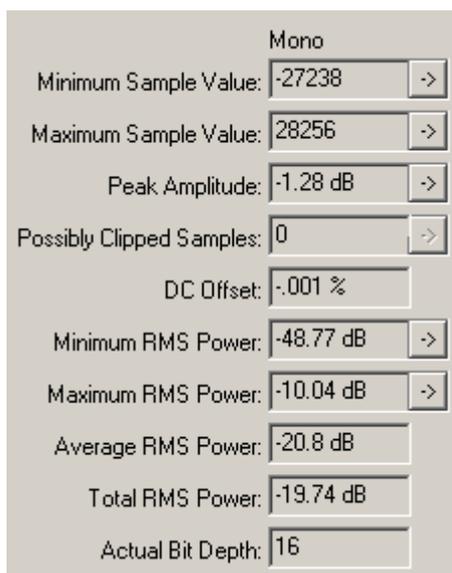


Рис.4. Статистические характеристики фонограммы музыки Бетховена «Элизе»

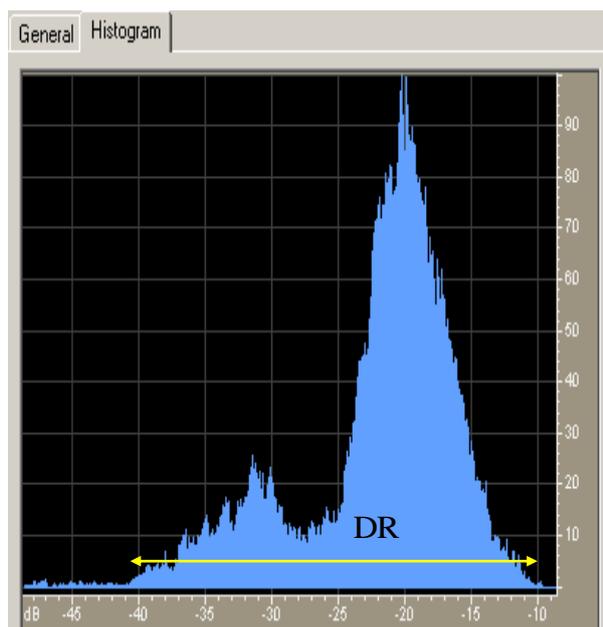


Рис.5. Гистограмма распределения музыки Бетховена «Элизе»

зависимости от задачи исследования. Если, например, важным является динамический диапазон мгновенных значений уровней фонограммы, то время интеграции должно быть 1-5 мс. Если измеряется динамический диапазон музыки с учетом слухового восприятия, то время интеграции выбирается равным 60 мс, это постоянная времени слуха.

Обычно по вертикальной оси нормализованной гистограммы фонограммы откладывается плотность вероятности распределения уровней фонограммы, что позволяет определять динамический диапазон и пик-фактор с заданной вероятностью при выбранном времени интеграции. В звуковом редакторе Adobe Audition 3 используется нормализация гистограммы, при которой максимальной вероятности событий всегда соответствует значение 100. Такая гистограмма описывает распределение вероятностей уровней сигналов фонограммы относительно максимального значения. При ее построении автоматически подбирается масштаб по оси X, поэтому затруднительно сравнивать гистограммы различных фонограмм.

Практическое применение. Кому и зачем нужна статистическая информация и гистограмма фонограммы. Эти данные прежде всего оказывают неоценимую помощь при динамической обработке фонограммы, так как они позволяют обоснованно выбрать характеристики компрессора и экспандера. Статистические результаты обработки фонограмм с музыкой различных жанров позволяют определить необходимый динамический диапазон электроакустического тракта, сформировать требования по пиковой и средней мощности головок акустических система. Они играют существенную роль при разработке алгоритмов компрессии звуковых сигналов.

Эмоциональную музыку с широким динамическим диапазоном и большим пик-фактором можно слушать только на высококачественной дорогой аппаратуре с хорошими акустическими агрегатами. На дешевых плеерах с наушниками и в автомобилях из-за шумов динамический диапазон сокращается и она звучит просто отвратительно. Поэтому широким спросом такие записи не пользуются и, неизбежно, с каждым годом динамический диапазон и пик-фактор

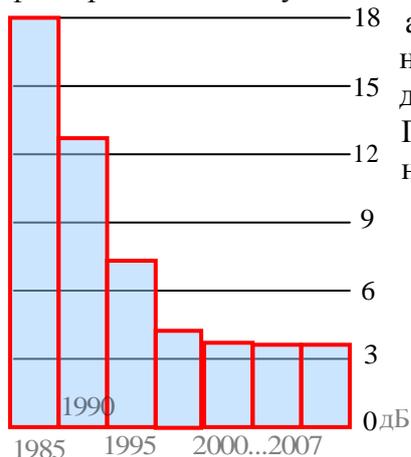


Рис. 6. Пик-фактор CD дисков

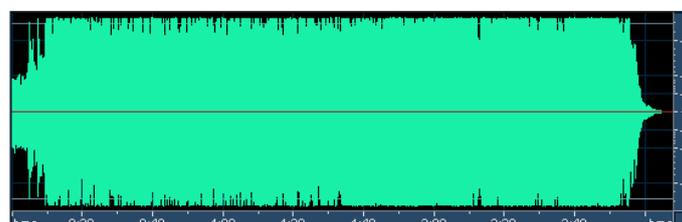


Рис.7. Фонограмма песни "I'll Be There For You"

фонограмм преднамеренно изготовителями уменьшается (рис.6.). На современных CD дисках в большинстве случаев динамический диапазон не превышает 20 дБ, а пик-фактор чуть больше 3 дБ, что вполне достаточно для танцевальной музыки. На рис.7. приведена картинка современной фонограммы с компакт диска.

3. Динамический диапазон цифрового аудио тракта

Обычный цифровой тракт включает в себя АЦП и ЦАП. Первый осуществляет квантование аналоговых сигналов, и преобразование их в цифровой поток. Второй производит обратное преобразование цифрового потока в аналоговый сигнал.

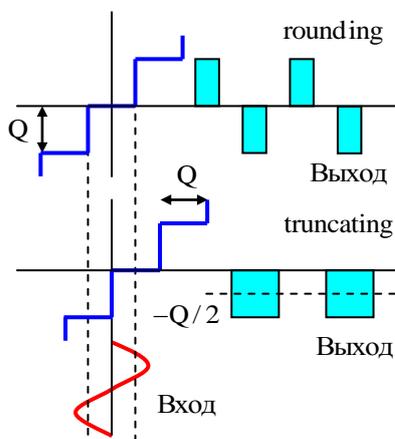


Рис. 8. Передаточные функции квантователей Mid-Tread и Mid-Riser

Квантование – это округление при АЦП последовательности выборок ЗС до целого двоичного значения. При импульсно-кодовой модуляции (ИКМ) такая операция осуществляется с помощью линейного квантователя, называемого в технической литературе *Mid-Tread*. У него передаточная функция имеет форму «лестницы» с одинаковыми ступеньками и, обязательно, **нечетное число уровней квантования**. Округление цифровых данных в этом квантователе производится до ближайшего двоичного значения (рис.8). Этот алгоритм принято называть *rounding*.

В этом алгоритме выходной сигнал квантователя симметричен относительно оси времени, и квантование осуществляется **с порогом, равным 0,5 шага квантования Q** . Пока входной сигнал меньше этого порога выходной сигнал квантователя равен нулю, это значит, что квантование осуществляется с центральной отсечкой. При входном сигнале несколько выше порога квантования выходной сигнал имеет вид последовательности импульсов со скажностью, зависящей от уровня ЗС. При дальнейшем увеличении уровня ЗС формируется выходной сигнал ступенчатой формы.

Округление цифровых данных в квантователе *Mid-Riser* производится до ближайшей меньшей величины (рис.8), поэтому данный алгоритм принято называть - *truncating*. Квантователь *Mid-Riser* отличается тем, что у него отсутствует порог квантования, поэтому он передает звуковые сигналы очень маленьких уровней, даже ниже уровня

шума. Однако, при отсутствии ЗС любой ничтожный шум порождает на выходе последовательность случайных импульсов с амплитудой 1 квант, это значит, что такой квантователь усиливает шумы.

Динамический диапазон АЦП с квантователем *Mid-Tread* определяется через логарифм отношения максимального и минимального значений сигнала синусоидальной формы на входе квантователя

$$DR_A = 20 \log \frac{A_{\max}}{A_{\min}},$$

где $A_{\max} = Q \cdot 2^{(q-1)}$, $A_{\min} = \frac{Q}{2}$, Q – шаг квантования, q – число разрядов. Поэтому

$$DR_A = 20 \log \left(\frac{Q \cdot 2^{(q-1)}}{Q/2} \right) = 6.02q \quad (1)$$

При $q = 8$ этот динамический диапазон равен 48 дБ, а при $q = 16$ он увеличивается до 96 дБ. Значение DR_A определяет нижнюю границу динамического диапазона по уровню входных сигналов квантователя типа *Mid-Tread*.

Динамический диапазон ЦАПа измеряется в соответствии с рекомендациями стандарта *EIAJ* через отношение максимального среднеквадратического значения сигнала синусоидальной формы \bar{A}_{\max} на его выходе к среднеквадратическому значению шума квантования σ , измеренному в полосе от 0 до частоты Найквиста F_N

$$DR_R = 20 \log \frac{\bar{A}_{\max}}{\sigma}, \quad (2)$$

где, $\bar{A}_{\max} = \frac{Q \cdot 2^{(q-1)}}{\sqrt{2}}$, $\sigma = \frac{Q}{\sqrt{12}} = 0,289Q < Q/2$.

Поэтому $DR_R = 6,02 \cdot q + 1,76$; $q > 1$. (3)

При $q = 16$ $DR_R = 98$ дБ, что на 2 децибела больше динамического диапазона квантователя, определяемого формулой (1). Измеренный таким образом динамический диапазон ЦАПа отождествляется со значением его *SNR*.

Если верхний диапазон частот ограничивается значением $F_{\max} < F_N$, то расчетная формула для *SNR* и DR_R принимает вид

$$SNR_R = DR_R = 6.02q + 1,76 + 10 \cdot \log \left[\frac{f_s}{2 \cdot F_{\max}} \right], \quad (4)$$

где f_s – частота дискретизации, F_{\max} – максимальная частота звукового диапазона. При $f_s = 44,1$ кГц и $F_{\max} = 20$ кГц и $SNR_R = DR_R = 98,5$ дБ. Как видно, отношение сигнал/шум лишь на 2 децибела больше динамического диапазона. Надо обратить внимание на то, что величина *SNR* зависит от частот f_s и F_{\max} , тогда как *DR* от этих параметров не зависит. Тем не менее, в большинстве технических публикаций динамический диапазон отождествляется с отношением сигнал/шум. Это подтверждается и стандартами AES 17 и IEC 61606.

В стандарте IEC 61606 рекомендуется измерять *SNR* и *DR* при подаче на вход АЦП синусоидального сигнала с частотой 997 Гц и уровнем минус 60 дБ FS с обязательным использованием технологии *TPDF Dithering*. При этом расчетное соотношение для *SNR* из-за вносимого дополнительного шума предлагается в виде

$$SNR_T = DR_T = 6.02q - 3,01 + 10 \cdot \log \left[\frac{f_s}{2 \cdot F_{\max}} \right].$$

При прежних условиях $DR = SNR = 93,7$ дБ, а не 96 дБ, как это часто встречается в технической литературе. Следовательно, уменьшается и расчетный динамический диапазон. Вместо SNR часто используется его обратная величина, определяющая интегральный уровень шума квантования

$$\bar{L}_{nT} = -SNR_T.$$

В соответствии со стандартом [IEC 61606](#) измерение динамического диапазона DR_R производится в соответствии со схемой, приведенной на [рис.9](#). В этой схеме тестовый

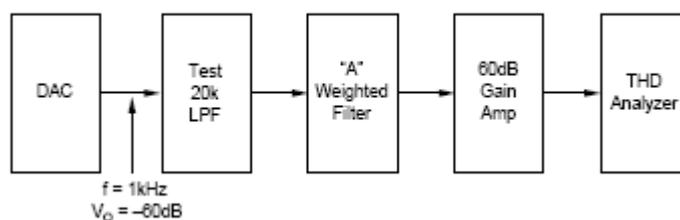


Рис..9. Схема измерения динамического диапазона ЦАП

цифровой сигнал с частотой 1 кГц, и уровнем минус 60 дБ, сформированный с использованием технологии *TPDF Dithering*, подается на вход ЦАПа. Аналоговый сигнал с ЦАПа поступает на вход ФНЧ с частотой среза 20 кГц, ограничивающий спектр шума квантования. Далее производится фильтрация с помощью взвешивающего фильтра типа *A*, учитывающего особенности слухового восприятия шума квантования, что увеличивает динамический диапазон на 2-3 дБ. Тестовый сигнал и шум усиливаются на 60 дБ и подаются на измеритель уровня *THD+N*. В этом измерителе тональный сигнал подавляется режекторным фильтром и вольтметром эффективных значений измеряется уровень шума в децибелах. Это измеренное значение уровня шума отождествляется, с обратным знаком, с динамическим диапазоном ЦАПа.

При квантовании сигналов минимального уровня возникают громадные искажения, достигающие 100% ([рис.10](#)). В связи с этим на практике приходится руководствоваться *реальным динамическим диапазоном* АЦП. При определении этого диапазона необходимо учитывать: пик-фактор музыкальных сигналов, достигающий 12...20 дБ, необходимость поднимать уровень нижней границы динамического диапазона над уровнем шума квантования хотя бы на 20 дБ и иметь дополнительный запас в верхней части динамического диапазона около 10...12 дБ для предотвращения случайной перегрузки.

В результате *реальный динамический диапазон записи* 16-ти разрядного АЦП ИКМ

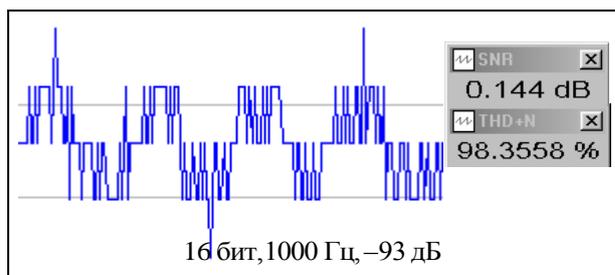


Рис.10. Последовательность выборок искаженной формы синусоидального сигнала

тракта не превышает 48...54 дБ. Этого даже близко не хватает для хорошей студийной звукозаписи. При автоматической регулировке уровней, что имеет место при записи компакт-дисков, этот диапазон может быть расширен до 74 дБ, только с заметным ухудшением качества звука сигналов низкого уровня.

Запас динамического диапазона сверху предохраняет от возможности перегрузки, когда пиковые уровни превышают ожидаемое значение. При записи танцевальной музыки вполне достаточно запаса в 6 дБ.

При записи симфонической музыки иногда приходится иметь запас до 20...30 дБ. Запас динамического диапазона снизу предотвращает возможность тихих пассажей оказаться ниже уровня шума и, тем более, ниже порога слышимости.

В цифровых трактах *верхняя граница динамического диапазона* ограничивается уровнем сигнала 0 дБ FS. Без использования технологии *Dithering* *нижняя граница динамического диапазона* ограничивается уровнем

$$L_A = -1/DR_A.$$

При $q = 8$ бит он равен минус 48 дБ, а при $q = 16$ бит – минус 96 дБ. Неизбежный шум тракта повышает этот уровень.

Интегральный уровень шума минус 93,7 дБ – это много или мало. Важно насколько этот уровень превышает порог слышимости. С использованием технологии *Dithering* шум

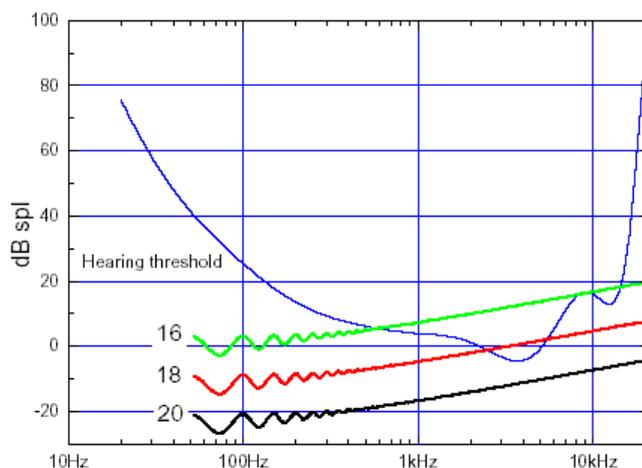


Рис.11. Пороги слышимости шума квантования в зависимости от числа разрядов

квантования становится белым шумом, порог слышимости которого равен 4 дБ SPL . Это значит, что вблизи 3 кГц шум квантования при $q = 16$ бит будет превышать порог слышимости на 22,3 дБ (рис.11). Как видно из этого рисунка, для того чтобы шум квантования был не слышен требуется использовать 20-разрядное квантование.

4. Расширение динамического диапазона с использованием технологии *Dithering*

Для расширения динамического диапазона ИКМ тракта с квантователем типа *Mead-Tread* без увеличения числа разрядов и частоты дискретизации разработано множество специальных алгоритмов. Например, запись в студии осуществляется с использованием 20 или 24- разрядов, а затем цифровые данные реквантуются до 16 разрядов с применением технологий *Dithering* или *Noise Shaping*.

Технология Dithering реализуется двумя путями. Первое, при квантовании аналоговых сигналов к ЗС добавляется небольшой аналоговый шум. Более часто эта технология используется при *реквантовании цифровых ЗС*, когда производится уменьшение числа разрядов кодовых слов. В этом случае в технологии *Dithering* применяется *цифровой шум*. В настоящее время исходная запись в студии осуществляется с 24 разрядами, а затем производится реквантование, обычно до 16 разрядов, как это принято в стандарте CD. При этом качество такого CD по шумам соответствует 20- разрядной записи.

В процессе реквантования чаще применяют операцию *truncating*, при которой просто отбрасываются младшие разряды кодовых слов. В этом случае в выходном сигнале

реквантователя возникает отрицательная постоянная составляющая, равная половине кванта (рис.8.). При таком алгоритме округления пока входной синусоидальный сигнал меньше одного кванта выходной сигнал квантователя имеет вид однополярных прямоугольных импульсов с постоянной скважностью 0,5, а шум паузы любого уровня преобразуется в случайную однополярную последовательность импульсов.

В настоящее время стандартом AES17 принята технология декорреляции ошибок квантования, при которой используется случайный или псевдослучайный белый шум с треугольной формой огибающей плотности вероятности (*TPDF – Triangle Probability Density Function*) мгновенных значений шума с размахом от пика до пика 2 LSB (*Least Significant Bit* – биты младших разрядов) и с равномерной спектральной плотностью мощности от 0 до частоты Найквиста.

Эта технология сокращенно называется «*TPDF dither*». Она легко реализуется практически и гарантирует: нулевое среднее значение ошибки квантования, отсутствие коррелированных искажений, отсутствие больших вариаций амплитуды шума и шума модуляции, зависимого от сигнала. При использовании такой технологии цифровая система ведет себя также как идеальная аналоговая систем, у которой беспредельное разрешение ниже LSB , отсутствуют искажения и шум модуляции. Надо подчеркнуть, что при использовании технологии «*TPDF dither*» шум становится более приятным на слух без использования эффектов маскировки искажений.

При использовании технологии дизеринг, устраняется порог квантования, линейаризуется передаточная функция при малых уровнях сигналов, погрешности квантования преобразуются *в белый шум с равномерной спектральной плотностью*. В обычном q – разрядном квантователе типа *Mid-Tread* с динамическим диапазоном $DR = 6q$ сигналы в шкале FS с уровнем меньше минус $-DR_A$ на выход не передаются. Благодаря шуму, используемому в технологии *Dithering*, передаточная функция квантователя линейаризуется и его порог размывается. Благодаря этому через квантователь могут передаваться сигналы с уровнем ниже минус $-DR_A$.

Физической основой такого феномена является способность человека слышать тональный звук, *маскируемый белым шумом*, когда уровень интенсивности тона меньше уровня мощности шума. Это связано с тем, что *маскировка тона происходит в критической полосе слуха*, в которой находится маскируемый тон. Этот феномен поясняется рис. 12. При такой маскировке уровень порога слышимости маскируемого тона является нижней границей динамического диапазона звукового сигнала.

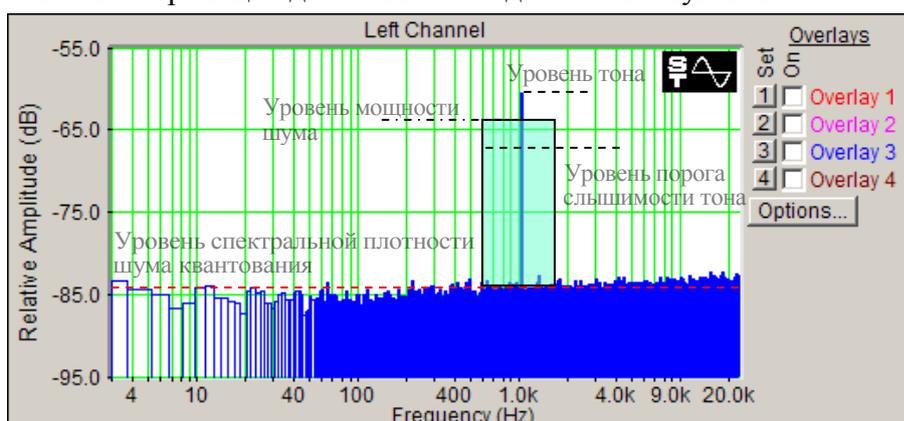


Рис. 12. Маскировка тона белым шумом в критической полосе слуха

Уровень мощности белого шума в критических полосах определяется равенством

$$L_{mn}(f) = L(f)_{sdn} + 10 \log(W_{cbw}),$$

где $L(f)_{sdn}$ - уровень спектральной плотности шума реквантования, W_{cbw} - ширина критической полосы. Она рассчитывается по эмпирической формуле

$$W_{cbw} = 25 + 75[1 + 1,4(f/1000)^2]^{0,69}, \text{ Hz}.$$

Ширина критической полосы меняется от 100 Гц на НЧ до 6 кГц – на высоких. Во столько

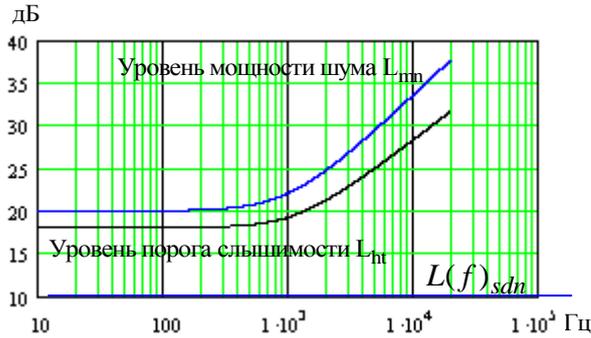


Рис.13. Уровень порога слышимости тона, маскируемого белым шумом

же раз возрастает уровень мощности шума реквантования в критических полосах (рис.13.). До частоты 500 Гц ширина критических полос одинакова и равна 100 Гц, поэтому в этом диапазоне уровень порога слышимости тона, маскируемого белым шумом, на 18 дБ выше уровня спектральной плотности мощности шума квантования. Выше 500 Гц уровень порога слышимости тона поднимается с крутизной немного меньше 10 дБ на декаду.

Спектральная плотность мощности шума реквантования $L(f)_{sdn}$ связана с

интегральным уровнем шума \bar{L}_n равенством

$$L(f)_{sdn} = \bar{L}_n - 10 \cdot \log(f_s / 2).$$

Для 16-ти разрядного АЦП с TPDF Dithering $\bar{L}_n = -93,7$, $L(f)_{sdn} = -137,2$ дБ.

Уровень мощности шума квантования в критической полосе рассчитывается по формуле

$$L_{ncw}(f) = L(f)_{sdn} + 10 \log(W_{cw}),$$

где $L(f)_{sdn}$ – уровень спектральной плотности мощности шума квантования, W_{cw} – ширина критической полосы.

Коэффициент маскировки в критической полосе это разность между уровнем мощности шума квантования в критической полосе и уровнем порога слышимости тона, он может быть представлен в виде равенства

$$K_m(f) = -\log\left[1 + 1,4 \left[\frac{f}{502}\right]^{2,5}\right] - 2, \text{ (дБ)}$$

Уровень порога слышимости тона $L(f)_{ht}$, маскируемого белым шумом, до частоты 500 Гц на 18 дБ выше уровня спектральной плотности мощности шума реквантования (рис.13). На более высоких частотах $L(f)_{ht}$ возрастает, с крутизной немного меньше 10 дБ на октаву. Аналитически эта зависимость может быть представлена эмпирическим равенством

$$L_{ht}(f) = L_{mn}(f) + K_m(f) = -DR_T.$$

При использовании технологии Dithering нижняя граница динамического диапазона квантователя Mid-Tread L_q зависит от числа используемых разрядов и абсолютного порога слышимости (АПС). Поэтому для определения величины расширения динамического диапазона необходимо на одном графике совместить шкалы FS и SPL. Для этого достаточно, чтобы в шкале FS 0 дБ соответствовал уровень максимального звукового давления (L_{max}) в шкале SPL, а 0 дБ SPL в шкале FS соответствовал уровню максимального звукового давления в шкале SPL со знаком минус ($-L_{max}$).

На рис.14 приведены графики, которые позволяют определить величину расширения динамического диапазона при реквантовании $16 \Rightarrow 8$ бит с использованием технологии

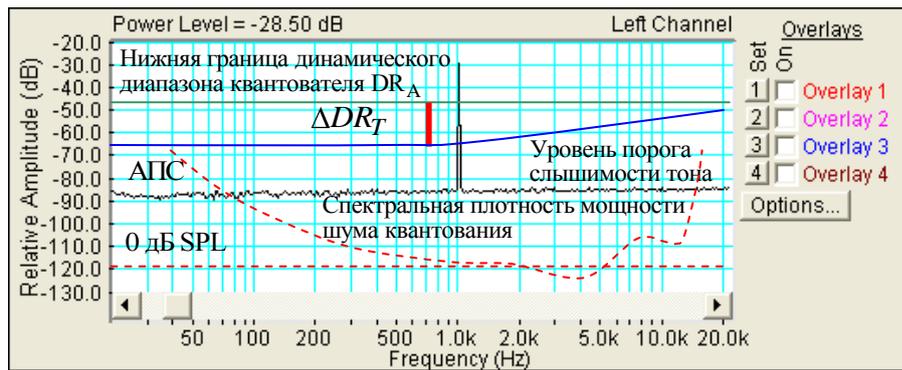


Рис.14. К определению динамического диапазона реквантователя Mid-Tread 16-8 бит с использованием технологии TPDF Dither

TPDF Dithering при уровне звукового давления $L_{\max} = 120$ дБ SPL. В этом случае нижняя граница динамического диапазона, определяемая квантователем, рассчитывается по формуле

$$L_q = -DR_A = -48 \text{ дБ}$$

По графику уровень спектральной плотности мощности шума квантования равен -86 дБ.

Как видно из рисунка, при использовании технологии *Dithering* нижняя граница динамического диапазона $L_{ht}(f)$ от 30 до 500 Гц определяется порогом слышимости тона, маскируемого белым шумом, и равна -67 дБ. Выигрыш в динамическом диапазоне может быть рассчитан по формуле

$$\Delta DR_T = -L(f)_{sdn} + L(f)_{ht} + L_q,$$

или определен по графику (красная линия). В данном примере этот выигрыш равен 19 дБ и, следовательно, динамический диапазон благодаря технологии *Dithering* расширяется с 48 до 68 дБ. На частотах выше 1 кГц выигрыш в динамическом диапазоне с повышением частоты постепенно уменьшается до 5 дБ.

В данном примере значение L_{\max} , равное 120 дБ SPL, практически никакого влияния на динамический диапазон не оказывает. Из графика видно, что с уменьшением L_{\max} будет постепенно уменьшаться ΔDR_T на низких частотах из-за повышения порога слышимости, который будет определять нижнюю границу динамического диапазона на этих частотах.

На рис.15 приведены графики, которые позволяют определить величину расширения динамического диапазона при реквантовании $24 \Rightarrow 16$ бит с использованием технологии

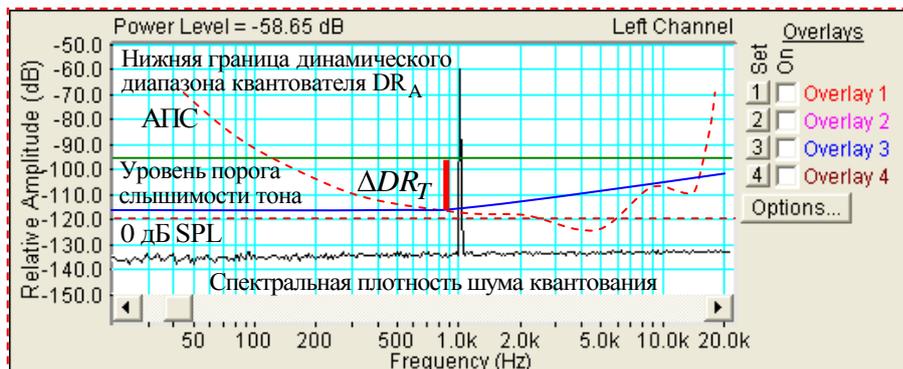


Рис. 15. К определению динамического диапазона реквантователя Mid-Tread 24-16 бит с использованием технологии TPDF Dither

TPDF Dithering при уровне звукового давления $L_{\max} = 120$ дБ SPL. В этом случае нижняя граница динамического диапазона, определяемая квантователем, рассчитывается по формуле

$$L_q = -DR_A = -96 \text{ дБ}$$

По графику уровень спектральной плотности мощности шума квантования равен – 135 дБ.

Как видно из рисунка при использовании технологии *Dithering* нижняя граница динамического диапазона L_d от 120 до 1000 Гц определяется абсолютным порогом слышимости. Выигрыш в динамическом диапазоне может быть рассчитан по формуле

$$\Delta DR_T = -L_{AIC} + L_q,$$

или определен по графику (красная линия). В данном примере этот выигрыш на частоте 1 кГц равен 21 дБ и, следовательно, динамический диапазон благодаря технологии *Dithering* расширяется с 96 до 117 дБ (теоретический предел). На частотах выше 1 кГц нижняя граница динамического диапазона определяется порогом слышимости тона, маскируемого белым шумом. При этом выигрыш в динамическом диапазоне с повышением частоты постепенно уменьшается до 5 дБ.

В данном примере значение L_{\max} , равное 120 дБ SPL, фактически определяет динамический диапазон. Из графика видно, что с уменьшением L_{\max} будет постепенно уменьшаться ΔDR_A на низких частотах из-за повышения абсолютного порога слышимости, и при $L_{\max} = 100$ SPL ΔDR_A будет равно нулю.

Таким образом, эффект расширения динамического диапазона при использовании технологии *TPDF Dither* реализуется лучше при мало разрядном реквантовании. При этом также допускаются значительно более широкие возможности изменения максимального звукового давления, при которых этот эффект работает. При 16-ти разрядном реквантовании применение этой технологии имеет смысл только при звуковом давлении не менее 110-120 дБ SPL.

В качестве практического примера применения технологии *Dithering* на [рис.16](#) приведены спектр и временная диаграмма ЗС с частотой 1 кГц и уровнем минус 68 дБ при

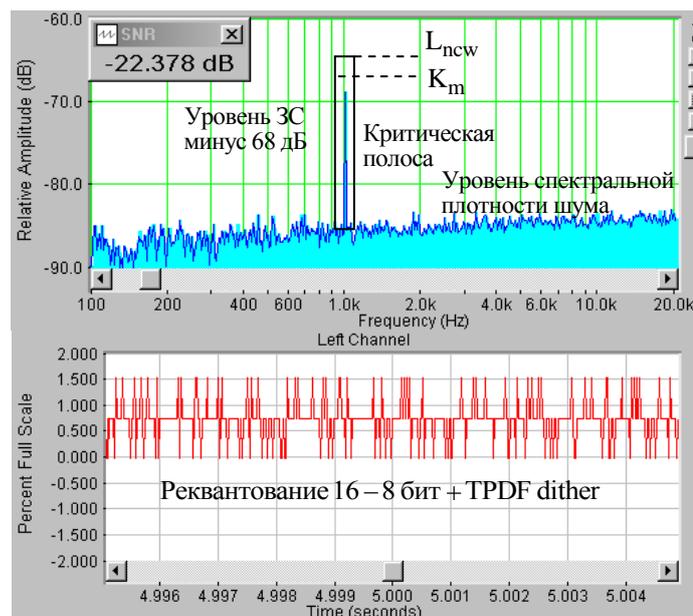


Рис.16. Спектр и временная диаграмма реквантованного ЗС

реквантовании 16 – 8 бит. При обычном 8-разрядном квантовании с округлением

динамический диапазон не превышает 48 дБ и ЗС с уровнем ниже минус 48 дБ через квантователь не передается.

Как видно, в случае применения технологии *Dithering* в спектре реквантованного сигнала отчетливо выделяется тональная составляющая, хотя сигнал на 22 дБ ниже эффективного значения шума квантования ($SNR = -22,3$ дБ) и на временной диаграмме он не просматривается.

По приведенным выше формулам рассчитаем динамический диапазон реквантованного ЗС. На частоте 1 кГц $W_{cw} = 160$ Гц, $L(f)_{sdn} = -85$ дБ (из графика). При этом уровень мощности шума квантования в критической полосе рассчитывается по формуле

$$L_{ncw}(f) = L(f)_{sdn} + 10\log(W_{cw}) = -85 + 22 = -63\text{dB},$$

Коэффициент маскировки в критической полосе определяется равенством

$$K_m(f) = -\log\left[1 + 1,4 \left[\frac{f}{502}\right]^{2,5}\right] - 2 = 2,8\text{dB}.$$

Уровень порога слышимости

$$L_{ht}(f) = L_{mn}(f) - \log\left[1 + 1,4 \left[\frac{f}{502}\right]^{2,5}\right] - 2 = -DR_T = 65,8\text{dB}$$

Таким образом, при 8 разрядном квантовании благодаря технологии *Dithering* динамический диапазон расширяется с 48 до 65,8 дБ, то есть на 17,8 дБ.

5. Расширение динамического диапазона с использованием технологии Noise Shaping

Под технологией *Noise Shaping (NS)* понимается преобразование огибающей спектральной плотности мощности шума квантования. Она является логическим развитием технологии *Dithering* и с ее помощью шум квантования из белого преобразуется в окрашенный шум с передаточной функцией такой же, как у фильтра высоких частот - с завалом низких частот и подъемом в области высоких частот.

В современной аудиотехнике эта технология наиболее широко используется в аналого-цифровых и цифро-аналоговых преобразователях, работающих на основе сигма-дельта модуляции. В этих структурах технология *NS* реализуется при частоте дискретизации значительно выше, чем это требуется по теореме В.П. Котельникова и основная мощность шума квантования переносится в область выше частоты Найквиста, поэтому достигается значительное увеличение *SNR* и расширение динамического диапазона.

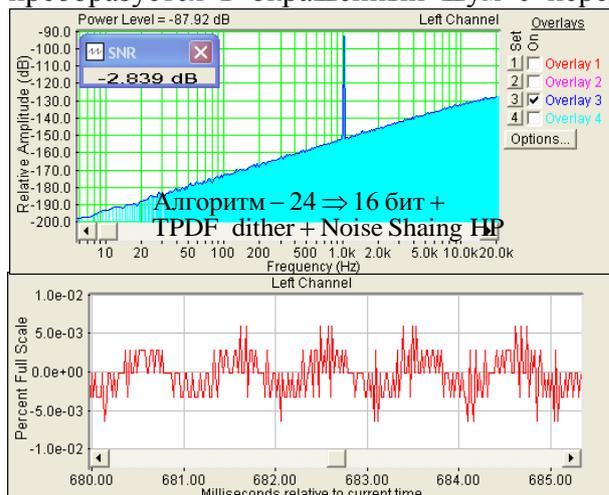


Рис.17. Спектр и временная диаграмма сигнала с уровнем минус 93 дБ FS

рассматривается применение такой технологии только при импульсно-кодовой модуляции в процессе реквантования, проводимой перед записью музыкальных программ, например, с помощью звуковых карт. В этом случае передискретизация не

используется и происходит изменение огибающей спектра шума квантования только в звуковом диапазоне.

При этом обеспечивается понижение уровня спектральной плотности шума в области максимальной чувствительности слуха и повышение в области высоких частот, близких к частоте Найквиста. Поэтому слуховая заметность шума квантования существенно уменьшается, но значение SNR резко падает и может стать отрицательным, так как сигнал оказывается значительно ниже уровня шума квантования (рис.17.). В этом случае, естественно, используемый в ИКМ трактах метод измерения динамического диапазона применять невозможно в принципе, и динамический диапазон можно определить лишь с помощью FFT анализа.

На рис.18 приведены графики, которые позволяют определить величину расширения динамического диапазона при реквантовании $16 \Rightarrow 8$ бит с использованием технологии

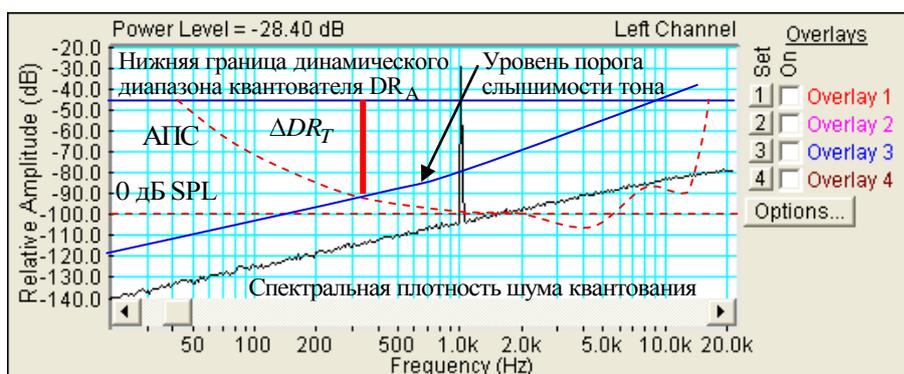


Рис.18. К определению динамического диапазона реквантователя Mid-Tread 16-8 бит с использованием технологии TPDF Dither+Noise Shaping

TPDF Dithering + Noise Shaping HP при уровне звукового давления $L_{max} = 100$ дБ SPL. В этом случае нижняя граница динамического диапазона, определяемая квантователем, рассчитывается по формуле

$$L_q = -DR_T = -48 \text{ дБ}$$

При использовании технологии Noise Shaping HP в цепи обратной связи квантователя включается фильтр высоких частот из-за которого график спектральной плотности мощности шума квантования имеет подъем высоких частот с крутизной 10 дБ на декаду. На частоте 1 кГц уровень спектральной плотности равен -115 дБ.

Как видно из рисунка при использовании технологии TPDF Dithering + Noise Shaping HP нижняя граница динамического диапазона L_d от 40 до 350 Гц определяется абсолютным порогом слышимости. Максимальный выигрыш в динамическом диапазоне достигается на частоте около 350 Гц (красная линия) и он равен 42 дБ. Благодаря этому динамический диапазон на этой частоте расширяется с 48 до 90 дБ.

На частотах ниже 350 Гц динамический плавно сокращается и на частоте 40 гц выигрыш снижается до нуля. На частотах выше 350 Гц нижняя граница динамического диапазона определяется порогом слышимости тона, маскируемого белым шумом. С повышением частоты динамический диапазон сокращается и на частоте 10 кГц выигрыш снижается до нуля.

В приведенном примере значение $L_{max} = 100$ дБ SPL, с увеличением L_{max} выигрыш по динамическому диапазону будет увеличиваться, а частота максимума снижаться. При этом ΔDR_T на частоте 1 кГц будет оставаться постоянной, равной 32 дБ, а динамический диапазон на этой частоте будет 80 дБ.

На рис.19 приведены графики, которые позволяют определить величину расширения динамического диапазона при реквантовании $24 \Rightarrow 16$ бит с использованием технологии TPDF Dithering + Noise Shaping HP при уровне звукового давления $L_{\max} = 120$ дБ SPL. В

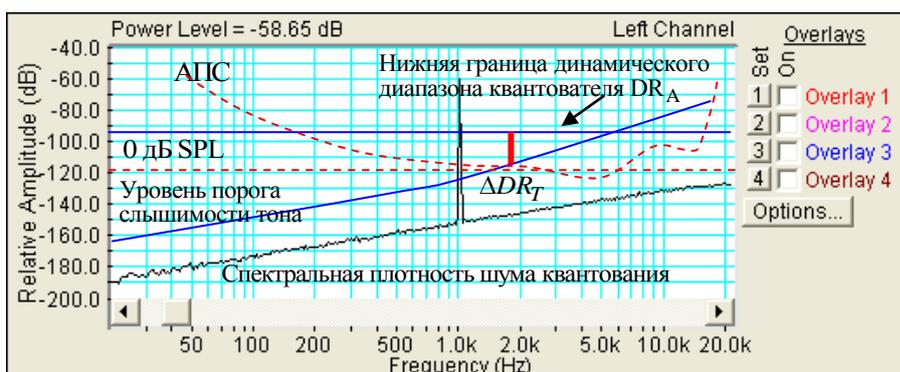


Рис. 19. К определению динамического диапазона реквантователя Mid-Tread 24-16 бит с использованием технологии TPDF Dither+Noise Shaping

этом случае нижняя граница динамического диапазона, определяемая квантователем, рассчитывается по формуле

$$L_q = -DR_T = -96 \text{ дБ}$$

При использовании технологии *Noise Shaping HP* в цепи обратной связи квантователя включается фильтр высоких частот из-за которого график спектральной плотности мощности шума квантования имеет подъем высоких частот с крутизной 10 дБ на декаду. На частоте 1 кГц уровень спектральной плотности равен -115 дБ, как и в предыдущем примере.

Как видно из рисунка, при использовании технологии TPDF Dithering + Noise Shaping HP нижняя граница динамического диапазона L_d от 200 до 2000 Гц определяется абсолютным порогом слышимости. Максимальный выигрыш в динамическом диапазоне достигается на частоте около 2000 Гц (красная линия) и он равен 21 дБ. Благодаря этому динамический диапазон на этой частоте расширяется с 96 до 117 дБ (теоретический максимум).

На частотах ниже 2000 Гц динамический плавно сокращается и на частоте 200 гц выигрыш снижается до нуля. На частотах выше 2000 Гц нижняя граница динамического диапазона определяется порогом слышимости тона, маскируемого белым шумом. С повышением частоты динамический диапазон сокращается и на частоте 6 кГц выигрыш снижается до нуля.

В приведенном примере значение $L_{\max} = 120$ дБ SPL, с уменьшением L_{\max} выигрыш по динамическому диапазону будет уменьшаться, а частота максимума повышаться. При $L_{\max} = 100$ дБ SPL выигрыш ΔDR_T будет равняться нулю.

Список литературы

1. Pohlman K.C. *Principles of Digital Audio*, 5rd Ed. McGraw-Hill, 2005.-860 с.
2. Watkinson J.R. *The Art of Digital Audio*. 2nd Ed.. Boston, MA: Focal Press, 1994.
3. Александр Радзишевский. Основы аналогового и цифрового звука. Из - во дом «Вильямс», М, СПб, Киев, 2006, -281.
4. Цифровая звукозапись. Под редакцией Дж. Мааса и М. Веркамена. Пер. с англ.- М.: Мир, 2004.-352 с., (Мультимедиа для профи).
5. Nika Aldrich. *Digital Audio Explained*. Second edition. Sweetwater, 2004, -373.
6. Julian Dunn. *High Dynamic Range Audio Applications for Digital Signal Processing*. Presented at AES 93rd Convention, October 1992.

7. *Stuart, J.R. and Wilson, R.J.*, 'Dynamic Range Enhancement Using Noise-shaped Dither Applied to Signals with and without Preemphasis', *AES 96th Convention*, Amsterdam, (Preprint #3871) (1994)
8. *Stuart, J.R. and Wilson, R.J.*, 'Dynamic Range Enhancement using Noise-Shaped Dither at 44.1, 48 and 96 kHz', *AES 100th Convention*, Copenhagen (May 1996).
9. *Stuart, J. Robert* - 'Predicting the Audibility, Detectability, and Loudness of Errors in Audio Systems', Preprint 3209, presented at the 91st AES Convention (October 1991)
10. *Larry Gaddy, Hajima Kawai*. Dynamic Performance Testing of Digital Audio D/A Converters. Burr-Brown Application Bulletin. 2002.
11. *Julian Dunn*. Measurement Techniques for Digital Audio. Audio Precision, 2004,- 203.
12. *Bruce E. Hover*. Testing High Performance Audio A-D and D-A Converters. Audio Precision. 2001.
13. AES17-1998. Revision of AES17-1991. AES Standard Method for Digital Audio Engineering Measurement of Digital Audio Equipment