

# АС3Filter

## Громкость и динамический диапазон

Александр Виговский

[АС3Filter home](#) :: [Donate](#) :: [Download](#) :: [Support forum](#)

Качество звука - это очень сложное понятие, которое зависит от огромного количества факторов и их взаимосвязи: типа акустической системы, условий прослушивания, качества исходной записи, самого слушателя и многих прочих. Невозможно дать однозначные рекомендации как добиться качества, нет единого рецепта, как нет и идеального звука: запись, подготовленная для одних условий прослушивания, в других условиях может звучать совершенно неприемлемо. Однако можно попытаться изложить некоторые главные принципы, которые помогут понять, что необходимо делать в каждом конкретном случае.

## **Содержание**

<b>1</b>	<b>Что такое громкость</b>	<b>3</b>
<b>2</b>	<b>Гистограмма громкостей</b>	<b>9</b>
<b>3</b>	<b>Проблема громкости DVD</b>	<b>12</b>
<b>4</b>	<b>Изменение уровня. Переполнение, обрезание и ограничение.</b>	<b>13</b>
<b>5</b>	<b>Нормализация</b>	<b>18</b>
<b>6</b>	<b>Однопроходная нормализация</b>	<b>20</b>
<b>7</b>	<b>Сжатие динамического диапазона</b>	<b>22</b>
<b>8</b>	<b>Микширование и громкость</b>	<b>25</b>
<b>9</b>	<b>Рекомендации</b>	<b>25</b>

# 1 Что такое громкость

Что такое громкость? Этот, казалось бы, очевидный вопрос достаточно трудно формализовать, поскольку в разных случаях подразумеваются совершенно разные вещи.

Наиболее очевиден смысл громкости, когда мы говорим о **звуковом давлении**, поскольку это то, что непосредственно воспринимается ухом.

**Звуковое давление** - давление, дополнительно возникающее при прохождении звуковой волны в жидкой и газообразной среде. Распространяясь в среде, звуковая волна образует сгущения и разрежения, которые создают добавочные изменения давления по отношению к среднему значению давления в среде. Т.о., звуковое давление представляет собой переменную часть давления, т.е. колебания давления относительно среднего значения, частота которых соответствует частоте звуковой волны. (Большая Советская Энциклопедия)

Таким образом, мы можем оценить любой звук - громкие звуки создают большое давление, тихие - малое. Давление измеряется в Паскалях, однако в акустике звуковое давление обычно измеряется в **децибелах** (дБ) относительно **порога слышимости**. По определению, величина порога принята равной  $p_t = 0.00002 \text{ Па} = 20 \text{ мкПа}$ . Порог слышимости принимается за 0дБ, а **громкость** вычисляется как  $l = 20 * \log(p / p_t)$ , где  $l$  [дБ] - громкость (в смысле звукового давления),  $p$  [Па]- звуковое давление,  $p_t$  [Па]- порог слышимости. При этом: все слышимые звуки имеют положительную величину громкости; неслышимые (ниже порога громкости) - отрицательную; изменение громкости на 6дБ соответствует двукратному изменению давления; изменение на 20дБ - изменению давления в 10 раз. Громкость в смысле звукового давления далее мы будем называть **абсолютной громкостью**.

Несколько типичных значений громкости:

Звук	Громкость	Давление
Порог слышимости	0dB	20 мкПа
Шелест листвы и слабого ветра Тиканье наручных часов Дыхание	10-20dB	60 - 200 мПа
Тихий шепот Тиканье настенных часов	20-30dB	200 - 600 мПа
Шум в помещении	30-40dB	0.6 - 2 мПа
Тихий разговор	40-50dB	2 - 6 мПа
Разговор средней громкости	50-60dB	6 - 20 мПа
Громкий разговор	60-70dB	20 - 60 мПа
Шумная улица	70-80dB	60 - 200 мПа
Двигатель грузового автомобиля	80dB	200 мПа
Шум в метро во время движения Отбойный молоток	90dB	600 мПа
Громкая дискотека	100-120dB	2 - 20 Па
Самолет на взлете	120dB	20 Па
Болевой порог	130dB	60 Па

Обратите внимание на диапазон воспринимаемых давлений: давление у порога слышимости и создаваемое самолетом отличаются в миллион раз! Поэтому логарифмическая шкала гораздо лучше согласуется с физиологией слуха - линейное изменение звукового давления не соответствует ощущению линейного изменения громкости. Например, изменение звукового давления на 50мПа при разговоре будет очень заметно, однако совершенно незаметно при взлете самолета. Изменение же звукового давления на 6дБ (в два раза) будет воспринято как примерно равное изменение громкости в обоих случаях, хотя в первом случае это будет соответствовать изменению давления на 25мПа, а во втором - 10Па.

Другая громкость - это **громкость записи (громкость сигнала)**. Эта громкость не является звуковым давлением (это может быть напряжение, намагниченность и т.д.), однако, звуковое давление создается в соответствии с громкостью сигнала записи при помощи системы воспроизведения. Каждой определенной громкости сигнала соответствует определенное звуковое давление. Громкость сигнала также может измеряться в децибелах. Однако если звуковое давление обычно измеряется относительно порога слышимости (минимального слышимого звукового давления), то громкость цифрового сигнала обычно измеряется относительно максимального цифрового уровня, принятого за 0дБ. Таким образом, громкость цифрового сигнала выражается отрицательными величинами (-3дБ, -20дБ) поскольку громкость записи всегда должна быть ниже максимальной. Чем меньше значение громкости, тем тише сигнал (-20дБ тише, чем -3дБ). Если громкость цифрового сигнала положительна, то это означает переполнение и, как следствие, возникновение цифровых искажений. (Суть этих искажений будет рассмотрена далее).

Регуляторы громкости на усилителе, в настройках системы, в плеере не создают звукового давления вообще. При отсутствии сигнала даже при максимальных установках громкости мы не услышим ничего (при условии, что сама система воспроизведения не создает шума). Таким образом, они влияют на громкость лишь косвенно и имеют смысл **усиления** сигнала (усиление может означать и ослабление сигнала). В дальнейшем термин громкость не будет использоваться для обозначения уровней усиления, за исключением **системной громкости**, поскольку это уже устоявшийся термин. Под системной громкостью будет пониматься как уровень усиления установленный в настройках операционной системы, так и в плеере, усилителе/ресивере и т.д.

Уровень усиления также может измеряться в децибелах. Это удобно, поскольку громкость сигнала и уровень усиления в таком случае просто складываются. Например, сигнал с громкостью 70дБ, усиленный на 10дБ будет звучать с громкостью 80дБ. Однако, несмотря на одинаковые единицы измерения необходимо всегда отличать громкость от усиления.

При воспроизведении сигнал записи преобразуется в звуковое давление при помощи акустической системы. Пусть максимальное давление создаваемое акустической системой составляет 100дБ. Тогда запись громкостью 0дБ будет создавать давление 100дБ, запись громкостью -30дБ будет создавать давление 70дБ и т.д. При изменении уровня усиления абсолютная громкость тоже будет изменяться. Таким образом, изменяя уровень усиления всегда возможно установить соответствие между громкостью записи и необходимой абсолютной громкостью. Например, если уровень диалогов в фильме составляет -30дБ, и мы хотим слышать их с естественной громкостью, то необходимо, чтобы громкость записи равная -30дБ соответствовала давлению 50дБ. Т.к. при максимальном уровне усиления (0дБ) уровень записи -30дБ создает давление 70дБ (что многовато), то при уровне усиления -20дБ, та же громкость записи создаст необходимые 50дБ давления. Отрегулированная таким образом акустическая система называется **калиброванной**, т.е. откалиброванная акустическая система - это система, где звуки воспроизводятся с правильной абсолютной громкостью. (На самом деле процесс калибровки может быть гораздо сложнее и включать намного больше параметров, однако здесь и далее мы будем говорить только о калибровке громкости). На большинстве бытовых приборов регулятор уровня усиления никак не маркируется (или наносятся абстрактные проценты или другие ничего не обозначающие цифры), поэтому для пользователя произвести точную калибровку акустической системы достаточно сложно.

На восприятие громкости также влияет свойство **адаптации** слуха. Слух приспосабливается к громкости окружающего шума, и соответственно корректирует восприятие громкости. Многие наверняка замечали, что человек в наушниках, в которых играет громкая музыка, начинает говорить громко (согласно восприятию окружающих), однако с его точки зрения это нормальная громкость речи (относительно громкости звука в наушниках). И наоборот - в полной тишине люди начинают говорить шепотом, и при этом шепот все равно кажется громким. В условиях обычного городского шума тиканье механических часов услышать очень сложно - этот звук

воспринимается как очень тихий. То же самое тиканье часов в полной тишине ночью может восприниматься как четко слышимое. Таким образом, возникает третья громкость - **субъективная громкость**. Мы не будем говорить о каком-либо измерении субъективной громкости, а только о качественном сравнении звуков (громче/тише).

Таким образом, звук одной и той же абсолютной громкости может восприниматься как громкий или тихий в зависимости от окружающих условий. Допустим, мы отрегулировали уровень усиления так, чтобы громкость диалогов в фильме приблизительно соответствовал реальности. Тогда, если в фильме есть звук тикающих в тишине часов, то при просмотре фильма в условиях городского шума мы не услышим часов вовсе, поскольку звук часов значительно тише окружающего шума, а слух адаптирован отсеивать окружающий шум. В хороших условиях прослушивания, когда окружающего шума нет, те же самые часы будут ясно слышны.

Во многих случаях громкость записи удобно измерять относительно некоторого опорного уровня. Например, если громкость записи составляет -20дБ - это громко или тихо? А если нам известно, что громкость диалогов в той же записи составляет -30дБ, то сразу можно сказать, что это достаточно громко. Если же громкость диалогов равна -10дБ, то можно сказать, что это достаточно тихо. **Уровень диалогов** (средняя громкость разговора) - это очень удобный опорный уровень, относительно которого можно легко ориентироваться. Если громкость звука на 10дБ выше уровня диалогов, то это громко, а если на 10дБ ниже - то это тихо. При этом сам уровень диалогов может быть произвольным и зависеть от записи - в одной записи уровень диалогов может составлять -10дБ, а в другой и все -30дБ. В любом случае звуки громкостью ниже уровня диалогов будут восприниматься как тихие, а выше уровня диалогов - как громкие, даже в случае если акустическая система будет неправильно откалибрована, и диалоги будут звучать с абсолютной громкостью в 40дБ или 60дБ. Благодаря свойству адаптации слух приспособливается к текущей средней громкости и вносит соответствующую 'поправку'. Однако правильное восприятие громкости будет нарушено, когда громкость окружающего шума приближается к громкости диалогов (или даже превышает его) - в этом случае все воспроизводимые звуки будут казаться тихими.

Таким образом, можно ввести еще одну шкалу громкости (в дополнение к шкале громкости цифрового сигнала и шкале абсолютной громкости) - громкость относительно уровня диалогов. При сопоставлении разных громкостей получим:

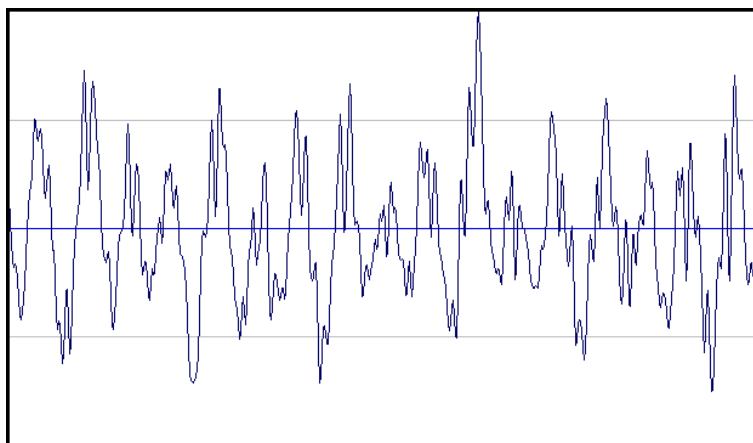
	Громкость реальных звуков	Громкость относительно уровня диалогов	Громкость записи-1	Громкость записи-2	Громкость при воспр. записи 1	Громкость при воспр. записи 2
<b>Тихие звуки</b>						
Тихий шепот	10dB	-40dB	-70dB	-50dB	10dB	10dB
Норма шума в жилых помещениях	20..30dB	-20..-30dB	-50..-60dB	-30..-40dB	20..30dB	20..30dB
Тихий разговор	40dB	-10dB	-40dB	-20dB	40dB	40dB
<b>Уровень диалогов</b>						
Разговор средней громкости	50dB	0dB	-30dB	-10dB	50dB	50dB
<b>Громкие звуки</b>						
Шум пишущей машинки	70dB	+20dB	-10dB	0dB	70dB	60dB
Двигатель грузового автомобиля	80dB	+30dB	0dB	0dB	80dB	60dB
Громкий автомобильный сигнал на расстоянии 5–7м	100dB	+50dB	0dB	0dB	80dB	60dB
Шум работающего трактора на расстоянии 1м	120dB	+70dB	0dB	0dB	80dB	60dB
Болевой порог	130dB	+80dB	0dB	0dB	80dB	60dB

В таблице также приведены две гипотетические записи, записанные в разных условиях: уровень диалогов в записи-1 составляет -30дБ, а в записи-2 -10дБ. Видно, что при воспроизведении на откалиброванной системе первая запись значительно лучше передает громкие звуки - возможно воспроизводить звуки громкостью до 80дБ, в то время как вторая запись воспроизводит звуки громкостью всего лишь до 60дБ.

Заметим также, что для двух приведенных записей требуется разная калибровка акустической системы. Так, для акустики, способной создавать давление в 100дБ уровень усиления, необходимый для первой записи составляет -20дБ, а для второй - -40дБ. Т.о. первая запись требует значительно бОльшего усиления, и при воспроизведении обеих записей при одинаковых настройках уровня усиления первая запись будет звучать значительно тише. Поэтому вторая запись удобна для некалиброванных систем, поскольку допускает сильные отклонения в уровне усиления.

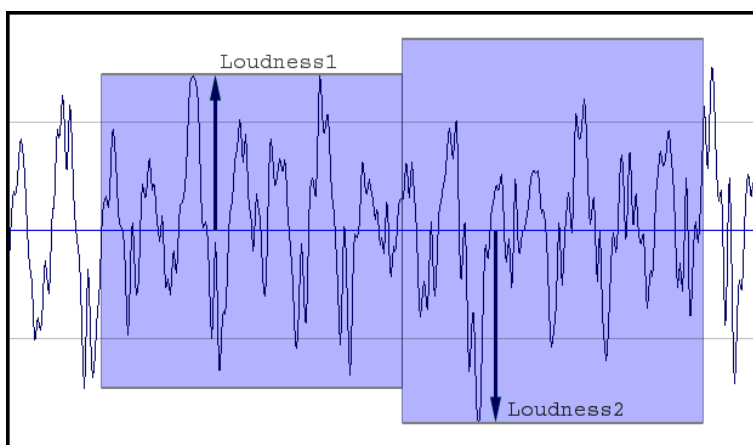
Таким образом, запись-1 хорошо передает громкие звуки, но требует большего уровня усиления; при недостаточном усилении и наличии внешних шумов будет восприниматься как чрезмерно тихая. Запись-2 не требует большого усиления, хорошо слышна даже на малых уровнях усиления и при наличии шума, но не может хорошо передавать громкие звуки.

А теперь вспомним, что звуковой сигнал - переменный:



Что брать для расчета громкости? Очевидно, что изменение давления за один период звуковой волны не имеет смысла изменения громкости звука, поскольку мы не слышим отдельные колебания. Поэтому громкость определяется не для одной точки, а на определенном промежутке времени. Есть много разных способов для определения громкости. Простейшие - это определение максимума и вычисление энергии сигнала. Более сложные методики учитывают неравномерность слуха к звукам разной частоты и интенсивности.

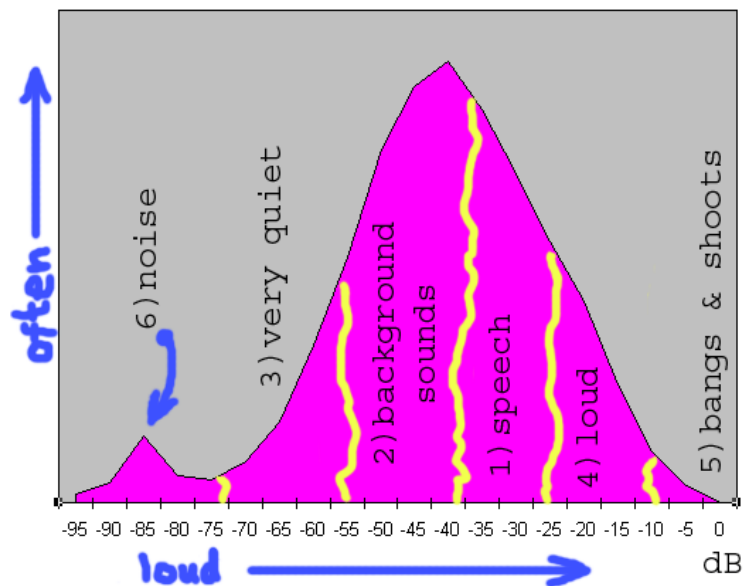
При определении громкости как максимума сигнала мы просматриваем диапазон и ищем максимум сигнала:



Найденный уровень, выраженный в дБ, и будет характеризовать громкость. Чтобы отличить эту громкость от полученных другим способом, эта громкость также называется **пик-уровнем**. В дальнейшем мы будем пользоваться только таким определением громкости. В действительности это определение не очень хорошо отражает реальное восприятие громкости, однако оно удобно для последующих рассуждений и поэтому мы не будем более подробно останавливаться на других, более точных методах определения громкости.



## 2 Гистограмма громкостей



Этот рисунок иллюстрирует распределение громкостей типичного DVD фильма (DVD FightClub, здесь и далее берутся только оригинальные аудио-треки, без перевода и без микширования, дабы не исказить звуковую картину). По оси X откладывается громкость в децибелах, по оси Y - насколько часто звук с этой громкостью появляется. Проще говоря, на протяжении фильма постоянно контролировалась текущая громкость, и чем чаще встречается какая-то громкость, тем выше в этом месте график. Т.е. можно сказать, что громких взрывов было мало, зато очень много разных фоновых звуков. График поделен на несколько условных областей:

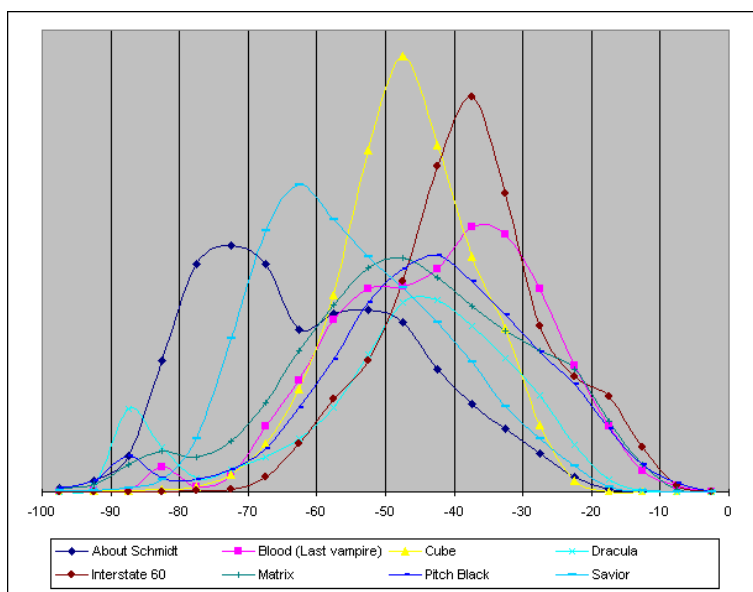
1. Голос. Уровень диалогов для фильмов - это один из наиболее важных параметров. Он является отправной точкой для всех остальных звуков: все, что ниже - считается тихими звуками, все, что выше - громкими. Хорошая слышимость диалогов это один из главных критериев настройки (о которой мы будем говорить дальше).
2. Фоновые звуки - проезжающие автомобили, звуки шагов, фоновая музыка
3. Очень тихие звуки - легкий ветер, шуршание травы и т.д.
4. Громкие звуки - звонок телефона, удары и прочее
5. Взрывы и прочие глобальные катастрофы
6. Шумы записи

Границы между областями весьма условны и могут изменяться для разных записей. В данном случае известно, что уровень диалогов равен -27дБ.

Рассматривая график можно отметить несколько интересных моментов. Во-первых, чаще всего в фильме встречались звуки с громкостью -40дБ. Пока просто отметим этот факт. Во-вторых, главный горб расположен от 0дБ (максимальный уровень) до -70дБ (минимальный уровень). Разумно предположить, что все важные звуки расположены внутри этого горба, а все, что тише - просто шумы. Разность максимального и минимального уровней составляет 70дБ. Эта величина называется **динамическим диапазоном** и будет далее иметь огромное значение.

Для фильмов на DVD приведенная картина достаточно типична. Диалоги занимают значительное время, немного больше - фоновый звук и музыка, периодические всплески громкости и периодический же постепенный уход в тишину. В целом, звуковая картина этого фильма достаточно сбалансирована и использует практически весь доступный динамический диапазон.

Теперь посмотрим и на другие фильмы:



На этих примерах очень хорошо видно разделение на две большие группы. Для первой группы фильмов характерно практически полное отсутствие громких звуков (Cube, Dracula, About Schmidt, Savior). Для них диапазон громкостей от 0дБ до -15дБ (очень громкие звуки) практически не используется. Для другой группы фильмов, наоборот громкие звуки чрезвычайно важны (Interstate60, Blood (Last Vampire), Pitch Black, Matrix). Казалось бы, что этих звуков относительно мало, поэтому особой роли они не играют. Однако, например, продолжительность звука выстрела (его громкой части) составляет доли секунды, но именно большая громкость выделяет этот звук. Если бы этот звук был тише, то создалось бы впечатление 'плоскости' и невыразительности звука. В фильмах первого типа ('тихие' фильмы) громкий звук, даже если встретится один раз на фильм, он все равно не несет большой смысловой

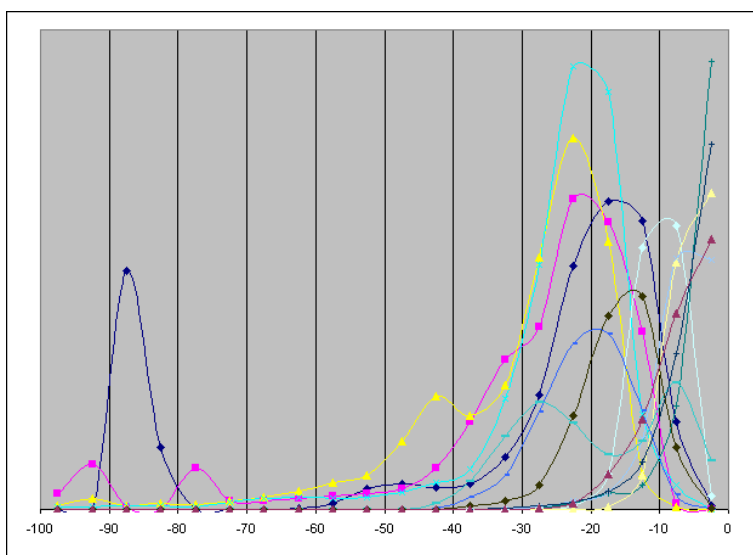
нагрузки, поэтому его громкость не так критична. В фильмах второго типа эти звуки (даже при относительно малой продолжительности) формируют настроение фильма.

Другая бросающаяся в глаза особенность - это высокие 'горбы' в фильмах Cube и Interstate60. В фильме Cube присутствует постоянный тихий фоновый гул и именно поэтому график имеет горб на громкости около -50дБ. То же самое касается фильма Interstate 60 - это путешествие на автомобиле, и как следствие, постоянный звук движения. Формы обоих гистограмм очень похожи, однако максимум гистограммы для фильма Cube находится на 10дБ ниже. Так как уровень голоса в обоих фильмах одинаковый, то сразу можно заключить, что фильм Cube значительно тише. Это далеко неочевидный вывод. Вспомним про свойство адаптации слуха - если сдвинуть графики так, чтобы их максимумы находились на одном и том же уровне, фильм Cube все равно будет восприниматься как 'тихий'!

Для прочих фильмов нет настолько хорошо выраженных фоновых звуков, поэтому графики более 'размазаны' по динамическому диапазону.

Фильм About Schmidt имеет большой горб на уровне -70дБ..-80дБ. Это шумы записи - в этом фильме они относительно громкие по сравнению с другими фильмами, однако достаточно тихие, чтобы они не мешали просмотру.

До сих пор мы рассматривали исключительно фильмы. Посмотрим теперь на музыку:



Сразу бросается в глаза совершенно другое распределение громкостей. (Здесь собраны гистограммы музыки разных направлений и разного качества, для того чтобы показать разнообразие. Они не показывают реального соотношения разных типов композиций.) Все звуки сильно сдвинуты в 'громкую' часть динамического диапазона. Разница в средней громкости фильмов и музыки достигает 40дБ. Кроме того, для фильмов характерен плавный спад гистограммы в громкой области. Для музыки же максимум гистограммы зачастую находится на максимальном уровне (0дБ). Таких

графиков приведено немного, однако такая ситуация является очень распространенной.

Понятие уровня диалогов к музыке зачастую неприменимо, поэтому разделить громкие и тихие звуки очень затруднительно. Поэтому также затруднено определение субъективной громкости композиции. Так, субъективно громкая композиция может быть объективно тише субъективно громкой. Большое значение здесь имеют положительные или отрицательные перепады громкости, которые на данных гистограммах не видны.

Отсутствие единого ориентира приводит к хаосу - звук из разных источников имеет разную громкость - разница в максимумах гистограмм композиций достигает 20дБ. Да и вообще понятие 'громкости композиции' очень сложно как-то формализовать. Многим наверняка знакома ситуация, когда есть много разной музыки из разных источников. При проигрывании ее подряд (например, если записать AudioCD) громкость постоянно меняется при переходе от одной композиции к другой, что неприятно (а упомянутый эффект несоответствия реальной и субъективной громкости может еще более запутать восприятие). Приведенные графики очень хорошо иллюстрируют такую ситуацию.

Динамический диапазон музыкальных композиций (разница между самым громким и самым тихим звуком) составляет 20-40дБ, что значительно меньше динамического диапазона фильмов (70дБ).

### **3 Проблема громкости DVD**

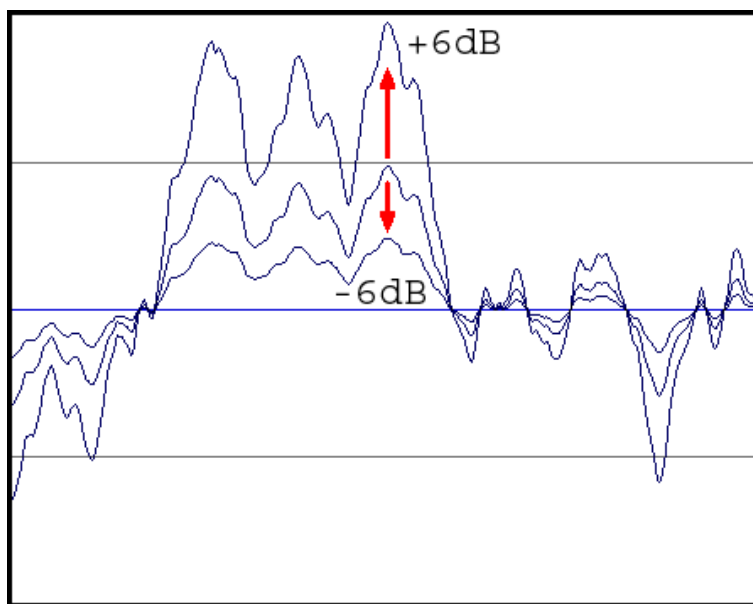
В среднем уровень громкости на музыкальных записях (а также системные звуки Windows) оказывается значительно выше, чем на DVD-фильмах. Поэтому при одинаковых настройках системной громкости абсолютная громкость фильмов будет значительно ниже и при наличии внешних шумов может оказаться недостаточной, слышно будет плохо (sic!) и будет казаться, что звук некачественный. Увеличение усиления в настройках системы и на усилителе даже до максимальной может не решить проблемы: разница в средней громкости достигает 40дБ, что очень много. Однако даже если мощности усилителя окажется достаточно и проигрывание DVD происходит с достаточной абсолютной громкостью это не всегда удобно, поскольку звуки операционной системы, которые звучат с нормальной абсолютной громкостью при нормальных настройках усиления, при большом усилении будут просто громкоподобными.

Эта проблема характерна в основном для компьютерного проигрывания, поскольку в аппаратных проигрывателях управление уровнем усиления осуществляется самим декодером. Некоторые программные DVD-проигрыватели умеют управлять системной громкостью, однако это тоже не всегда удобно, поскольку изменяется громкость всех звуков системы (можно случайно оглушить соседей), да и управлять громкостью на внешнем усилителе программа все равно не может. Поэтому это только частичное решение проблемы.

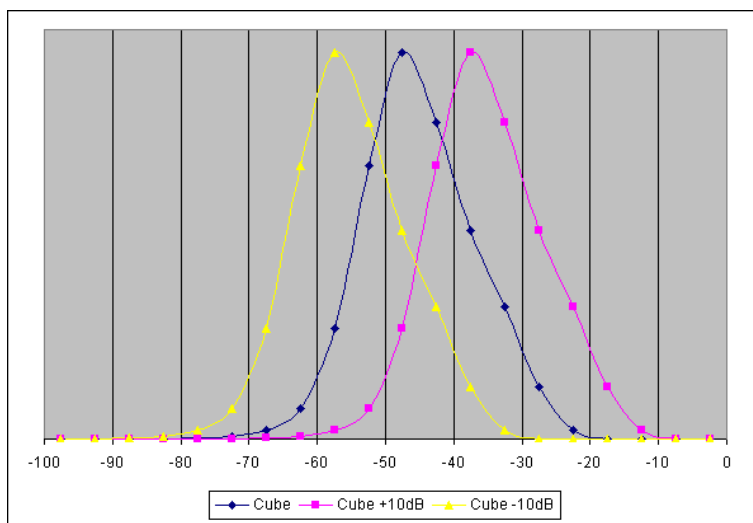
Компромиссное решение состоит в обработке звука непосредственно перед проигрыванием. Подобная обработка может сильно повысить качество проигрывания конкретной записи и в конкретных условиях. Многие могут возразить, что при этом теряется 'качество', однако, как уже упоминалось - нет абсолютного качества, Мы не ставим себе целью монтировать звук, наша цель - добиться того, чтобы в наших условиях слушать его было приятно. Если акустическая система не обладает достаточной мощностью или у нас нервные соседи смотреть фильм с пониженной громкостью, когда не разобрать половины слов и неслышно половины тихих звуков будет просто неприятно. Даже самовнушение о максимально достигнутом 'качестве' не сможет сгладить этого впечатления. Еще раз повторю основную мысль: качество - это то, что мы слышим. Звук проходит много стадий обработки до того, как попасть к слушателю, а разнообразие акустики и ее свойств настолько велико, что последняя стадия обработки непосредственно перед воспроизведением является практически необходимостью.

## 4 Изменение уровня. Переполнение, обрезание и ограничение.

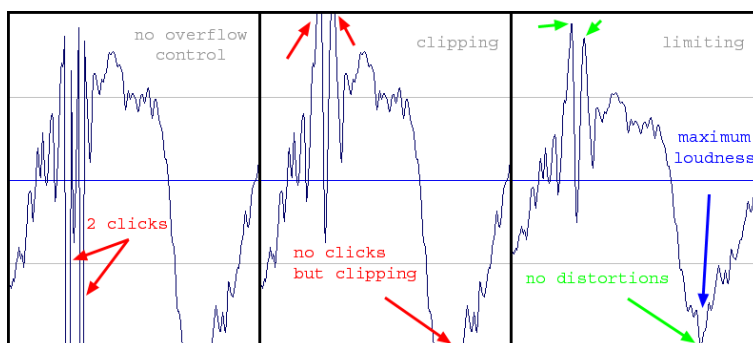
Изменение уровня - это просто умножение амплитуды сигнала на определенное значение, в результате чего изменяется громкость всего сигнала (увеличивается или уменьшается).



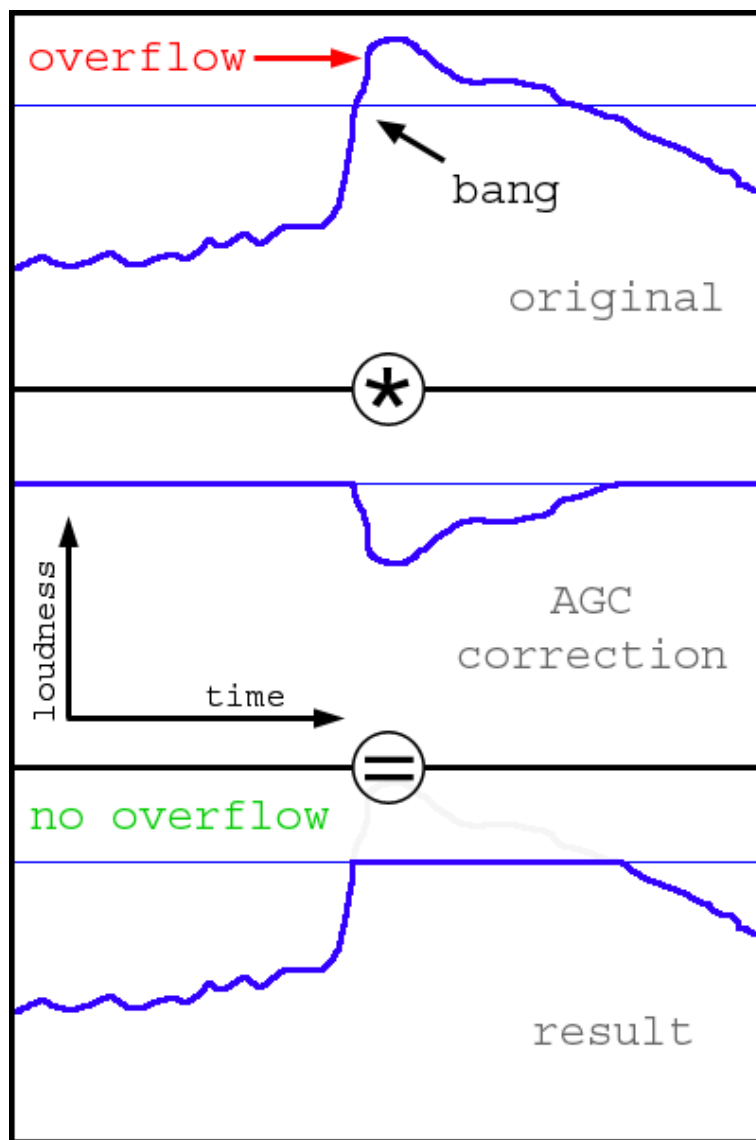
В логарифмическом виде умножение на число - это просто прибавление константы. Поэтому если ко всем уровням прибавить одно и то же число, то гистограмма уровней просто сдвинется:



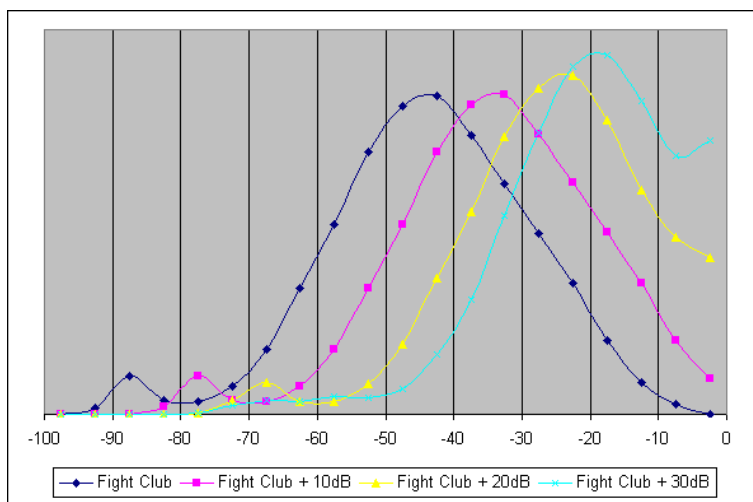
Если мы умножим сигнал на слишком большое число, то возможно переполнение. Если переполнение не отслеживать, то значение амплитуды принимает практически случайные значения (см. рисунок) и это приводит к очень заметным на слух частым щелчкам. Самый простой способ борьбы с этим дефектом - **обрезание сигнала** (клиппинг, clipping), т.е. там, где амплитуда сигнала выходит за пределы мы 'обрезаем' его (см. рисунок). При небольшом переполнении клиппинг практически незаметен на слух, однако при повышении уровня он проявляется как 'песок' в звуке.



Более сложный, но и более эффективный способ - это ограничение сигнала (лимитинг, limiting). Суть его состоит в том, чтобы автоматически понижать уровень сигнала так, чтобы переполнения просто не происходило. Для этого необходима система автоматической регулировки усиления (АРУ, Automatic gain control, AGC), которая будет вносить поправку в текущий уровень сигнала. Пока переполнения нет, АРУ не изменяет сигнал. Как только сигнал превышает максимальный уровень, вносится поправка, которая изменяет сигнал так, чтобы его уровень не превышал максимальный. Как видно на рисунке, сигнал полностью сохраняет свою форму! Однако громкость при этом остается максимальной:

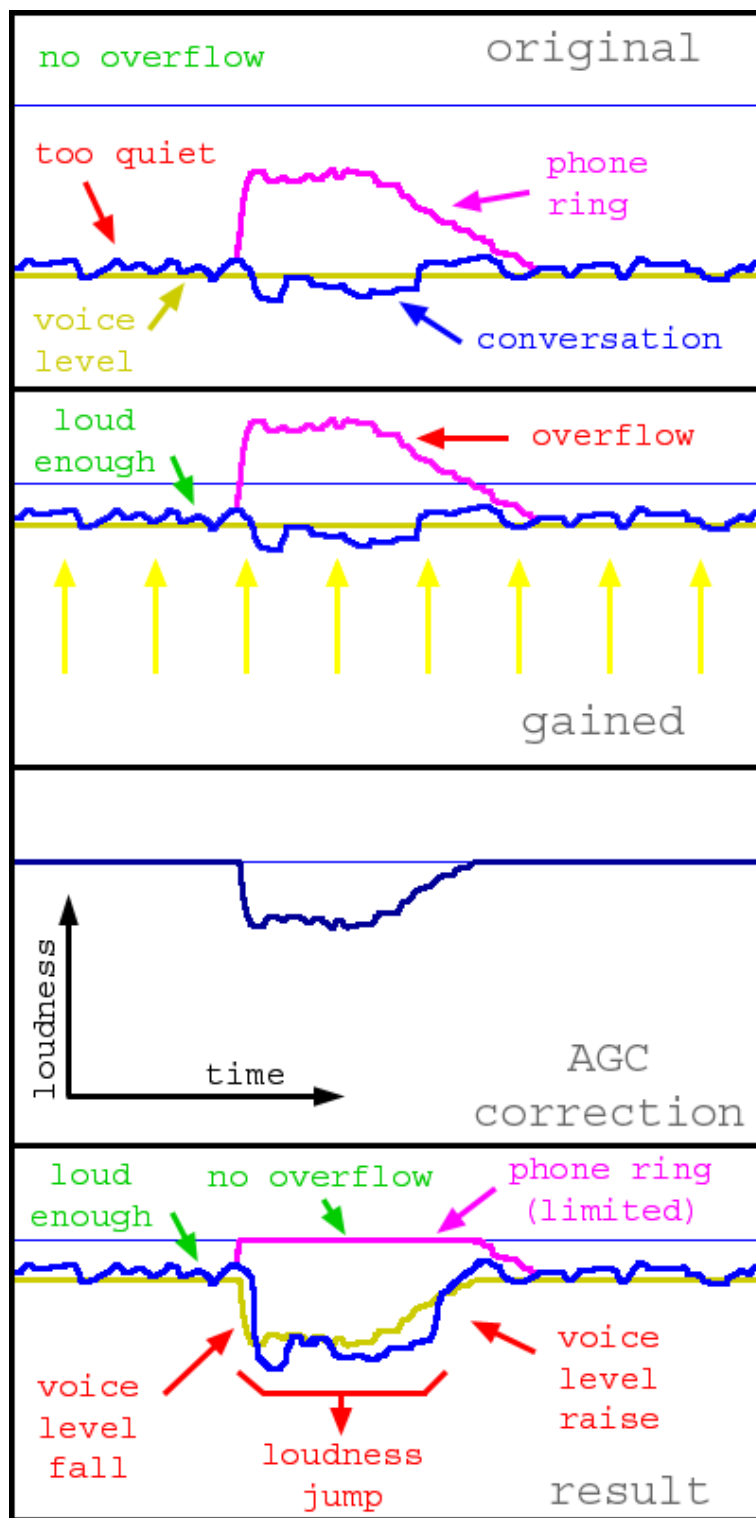


Необходимо обратить внимание на то, что на предыдущем рисунке приведена амплитуда, а на этом - громкость (определенная по *большой группе амплитуд*). Поэтому, несмотря на то, что график громкости 'обрезан' форма самого сигнала искажается минимально. Этот способ также не лишен недостатков. Изначально звук имел разную громкость, а после ограничения все звуки выходящие за пределы имеют одну и ту же громкость и звук может получиться 'плоский' и невыразительный. Посмотрим на гистограммы (Fight Club):



Здесь приведен пример неудачного усиления. При усилении +10дБ искажения еще мало заметны (мест в фильме, где возникает переполнение еще относительно мало), однако с ростом усиления ограничение уровня появляется все чаще и чаще и при уровне +30дБ начинает ограничиваться голос, что крайне заметно. Особенно заметны недостатки ограничения при значительных изменениях громкости за малое время - субъективно громкость начинает прыгать вверх и вниз. Если на фоне разговора (который при усилении +30дБ уже воспроизводится с максимальной громкостью) происходит громкий звук (который громче голоса в оригинале, например, звонок телефона), то уровень усиления уменьшается так, чтобы громкий звук не искажался, но при этом громкость разговора также резко падает. Когда громкий звук заканчивается разговор так же резко возвращается к предыдущей громкости:





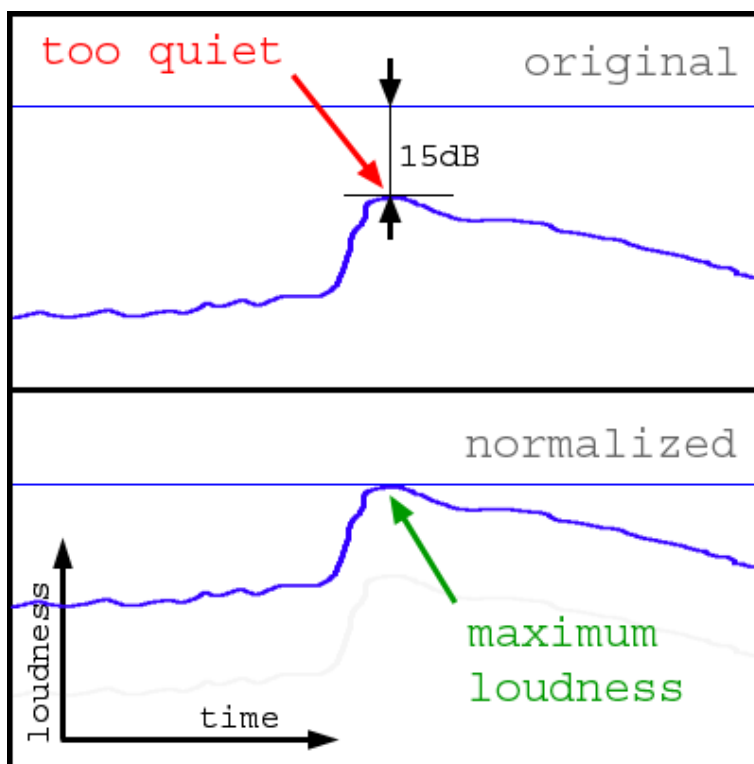
Т.о. относительная громкость разных звуков воспроизводимых одновременно сохраняется, однако уровень усиления постоянно изменяется, что хорошо заметно и очень неприятно. Таким образом, при больших уровнях усиления (20-30дБ и больше)

ограничение тоже дает плохой результат.

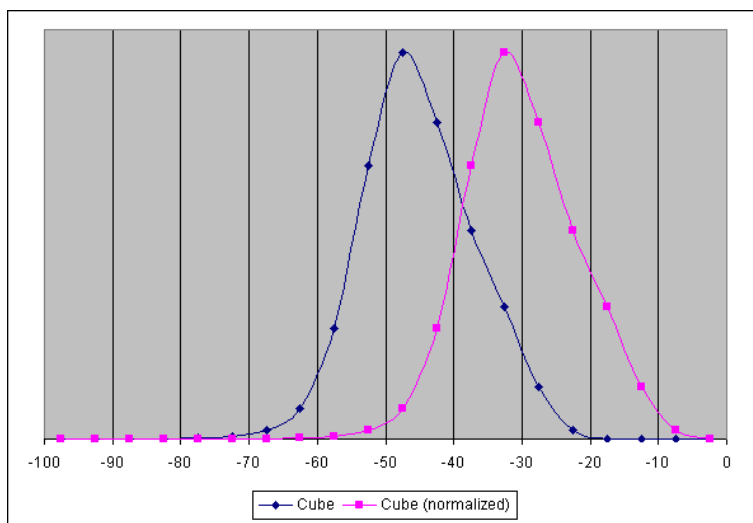
При малых уровнях усиления дефекты ограничения сигнала практически незаметны, тогда как обрезание сигнала хорошо слышно практически всегда. Т.е. при прочих равных условиях ограничение дает более качественный результат. Фильтр всегда производит либо ограничение (опция **“AGC”** включена) либо обрезание (опция **“AGC”** отключена). В дальнейшем если не оговорено явно будет подразумеваться, что включено ограничение. Поэтому рекомендуется всегда держать опцию **“AGC”** включенной.

## 5 Нормализация

Зададимся целью поднять громкость без потери качества вообще. Такое возможно? Да, если есть 'запас' динамического диапазона. Посмотрим снова на гистограммы с фильмами, фильм Cube (желтый график). На гистограмме видно, что громкость не поднимается выше -15дБ (возможно, что за весь фильм и есть один-два таких момента, однако, для целей данного рассмотрения будем считать, что громких звуков нет вообще). А так как громких звуков нет, то можно повысить громкость на 15дБ без потери качества!



Кроме увеличения громкости никаких изменений в сигнале не происходит. При этом гистограмма, не изменяя формы, подвинется вплотную к правому краю:



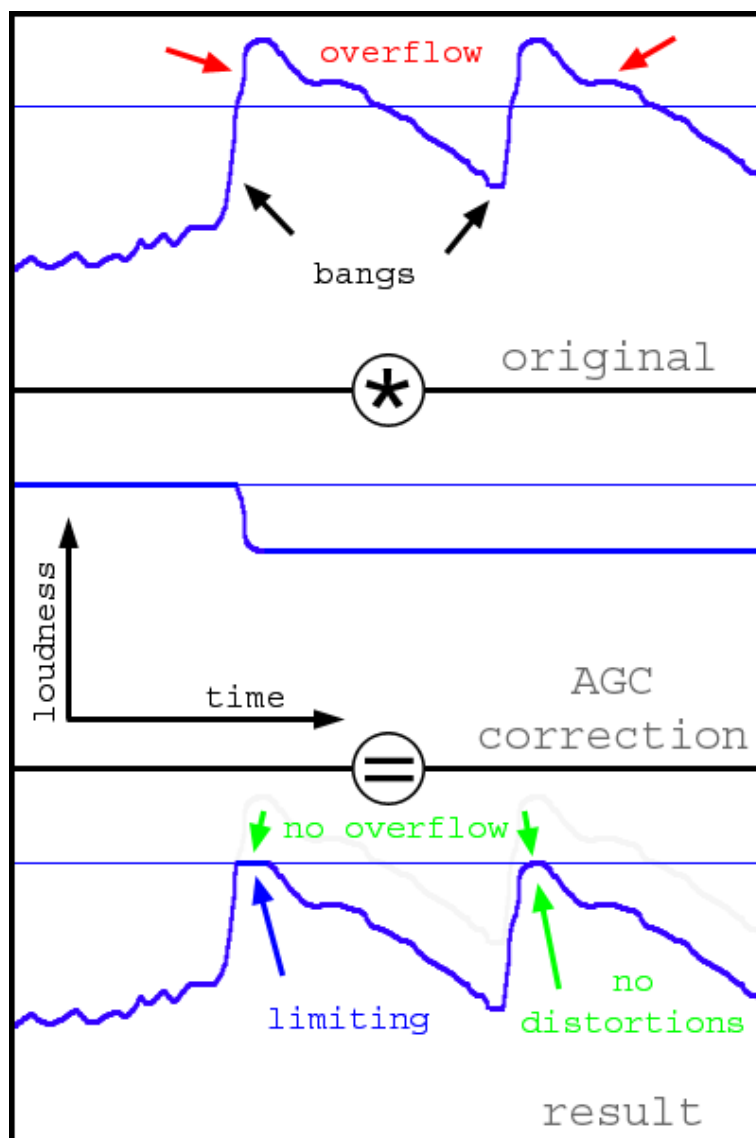
Таким образом, окажется, что в фильме теперь присутствует звук с максимальной громкостью. Но если мы будем увеличивать громкость дальше, возникнут искажения (описанные выше). Такая форма сигнала, где присутствует звук максимальной громкости без внесенных при этом искажений, называется **нормализованной**, а процесс - **нормализацией**. Нормализация часто производится на стадии подготовки музыкальных записей и к нам звук попадает уже нормализованным (например, AudioCD). Однако для фильмов нормализации не производится. Почему?

Вспомним гистограммы для музыки и хаос с громкостью. Для музыки важно воспроизведение с максимальной громкостью, поскольку она рассчитана на максимальную целевую аудиторию - CD-плееры, шум на улице, метро, дешевые наушники, хрипящие радиостанции, через которые тихие звуки ниже -40дБ просто не услышать. (Еще раз посмотрим на графики для музыки - минимальная громкость примерно равна -40дБ...)

Для DVD это не так. Они в первую очередь предназначены для качественных домашних кинотеатров. Воспроизводящая аппаратура должна быть откалибрована так, чтобы всегда воспроизводить диалоги с одной и той же абсолютной громкостью для любого фильма, будь то боевик с его потрясающими стену взрывами или мелодрама с тихим шуршанием травы. При этом воспроизводящая система всегда должна быть готова воспроизвести эти звуки без необходимости изменять уровень усиления вручную. Таким образом, на первый план выступает не необходимость сделать запись максимально громкой, а необходимость жесткого опорного уровня, чтобы декодер имел возможность подстроить уровень усиления автоматически. Этот опорный уровень может быть любым (на самом деле это не важно, т.к. если есть опорный уровень, то соответствующую коррекцию можно сделать всегда). Стандартом де-факто для DVD является уровень диалогов равный -27дБ. Поэтому даже если фильм громких звуков не содержит, то на этапе создания диска проводить нормализацию не следует и динамический диапазон остается неизрасходованным.

## 6 Однопроходная нормализация

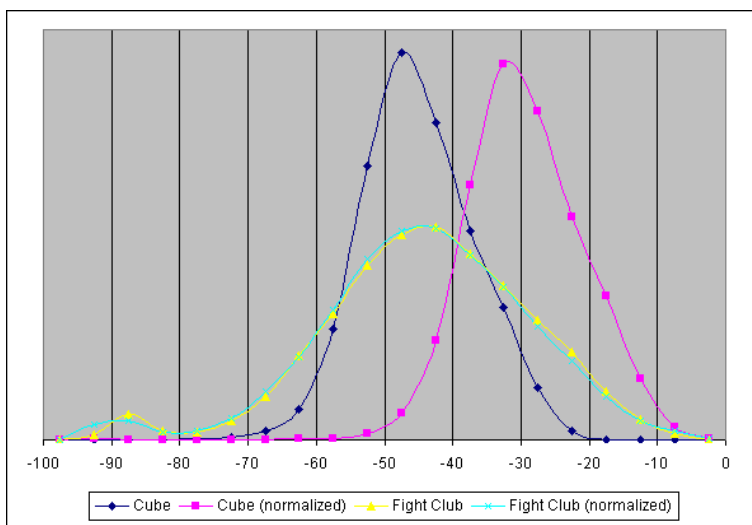
Итак, для целей повышения громкости нормализация это хорошо. Но для того, чтобы провести нормализацию, необходимо знать максимальный уровень записи, а чтобы его найти, нужно просмотреть всю запись заранее. Однако это не всегда возможно, да и просто неудобно. Есть метод нормализации, для которого не нужно предварительно просматривать всю запись - однопроходная нормализация. Суть метода состоит в том, что при просмотре фильма мы ищем максимум громкости постоянно. В начальный момент времени усиление максимально. В момент, когда возникает переполнение, мы уменьшаем усиление:



Как видно на рисунке первый пик громкости срезается почти так же, как и при обычном переполнении, однако продолжительность 'среза' значительно меньше (срав-

ните с простым ограничением) и впоследствии переполнений не происходит. Таким образом, усиление будет корректироваться при нахождении каждого нового максимума и в результате гистограмма автоматически сдвигается в сторону достижения максимальной громкости.

Как и все прочие, метод не лишен недостатков. Во-первых - метод не лишен искажений (хотя и малозаметных - нужно точно знать, что слушать, чтобы заметить работу однопроходной нормализации). Во-вторых, громкость постоянно снижается во время просмотра фильма (хотя обычно основной процесс снижения громкости заканчивается на первых 10-15 минутах фильма, поэтому практически незаметен). В-третьих, основная цель - повышение громкости, может быть не достигнута - все зависит от самого фильма (это касается и обычной нормализации). Вот пример гистограмм полученных в случае однопроходной нормализации для фильмов Cube (где нормализация возможна) и Fight Club (где нормализация не даст желаемого эффекта):

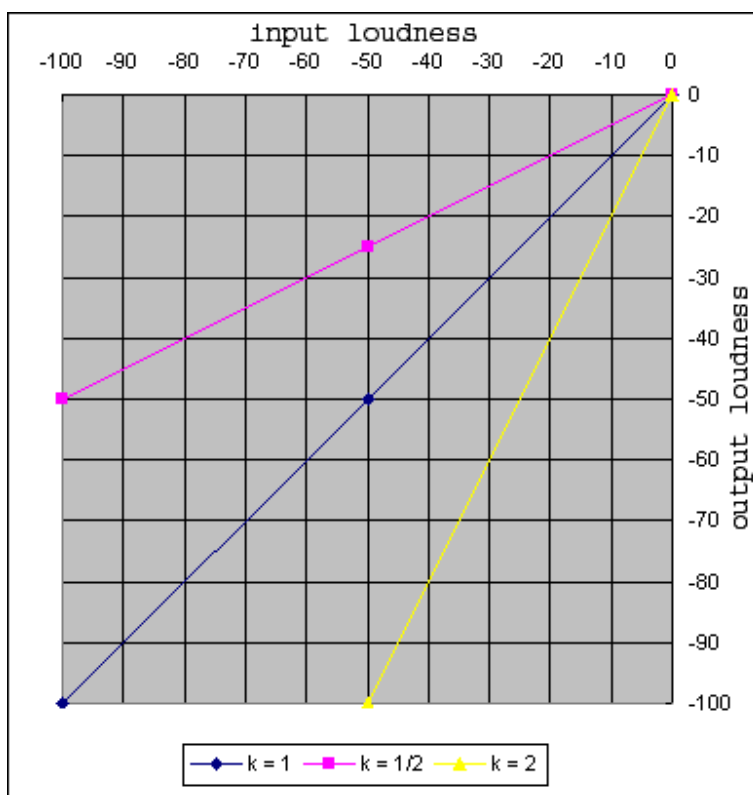


Как видно на этих гистограммах результат отличается разительно. Включение однопроходной нормализации дает существенный положительный эффект для фильма Cube - стало заметно громче (и подтвердилось наше предположение, что звуки громче -15дБ в фильме отсутствуют - сравните с графиком для нормализации). Но для фильма Fight Club увеличения громкости вообще не произошло - таким образом нужный эффект не был достигнут, но были внесены дополнительные искажения.

Необходимо еще раз отметить, что для работы однопроходной нормализации необходимо предварительное (начальное) усиление - без повышения уровня однопроходная нормализация просто бессмысленна. Также как и обычное усиление, оно задается уровнем Master. Однако если оно будет слишком сильным - с начала работы алгоритма падение громкости будет очень заметно.

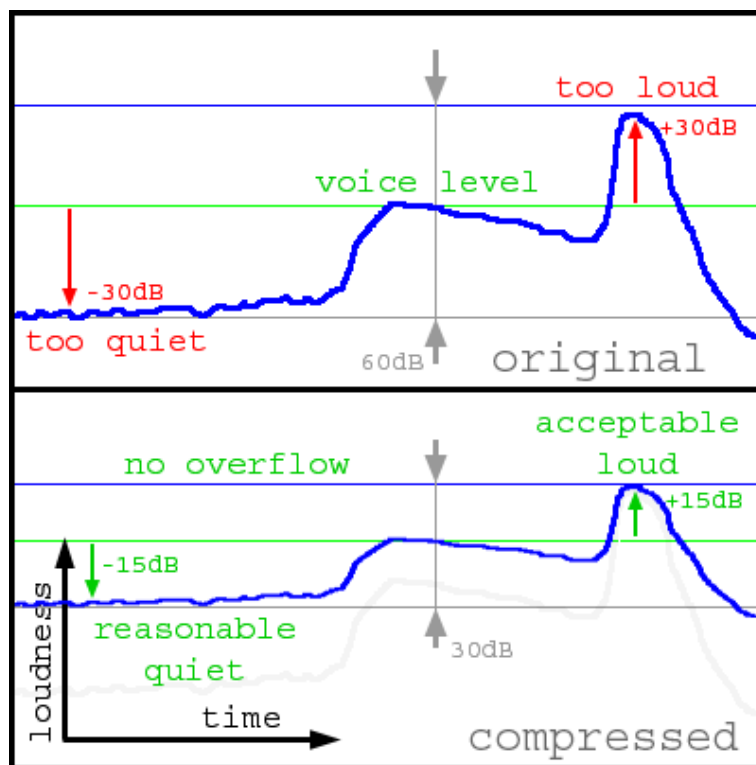
## 7 Сжатие динамического диапазона

Задумавшись над вопросом - а зачем нам поднимать громкость? Для того чтобы слышать тихие звуки, которые не слышны в наших условиях (например, если нельзя слушать громко, если есть посторонние шумы в комнате и т.д.). А можно ли усилить тихие звуки, а громкие не трогать? Оказывается можно. Эта техника называется сжатием динамического диапазона (компрессия, Dynamic Range Compression, DRC). Для этого необходимо изменять текущую громкость постоянно - тихие звуки усиливать, громкие - нет. Самый простой закон изменения громкости - линейный, т.е. громкость изменяется по закону  $\text{output\_loudness} = k * \text{input\_loudness}$ , где  $k$  - коэффициент сжатия динамического диапазона:



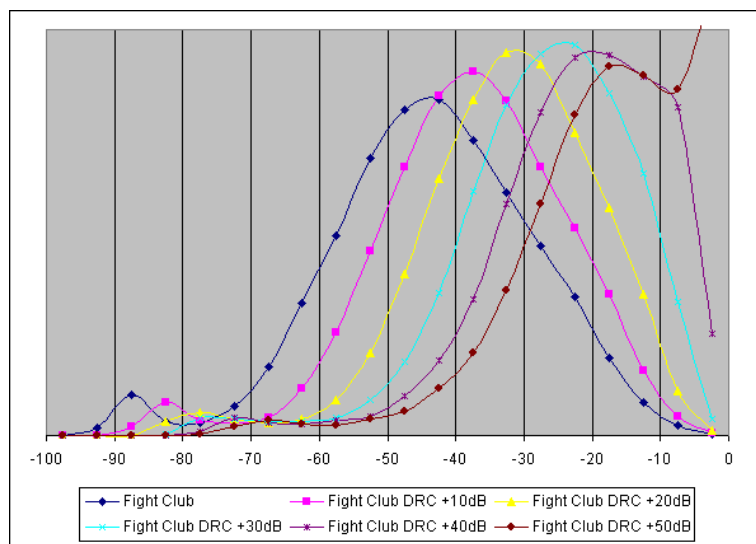
При  $k = 1$  никаких изменений не производится (выходная громкость равна входной). При  $k < 1$  громкость будет увеличиваться, а динамический диапазон - сужаться. Посмотрим на график ( $k=1/2$ ) - тихий звук, имевший громкость -50дБ станет громче на 25дБ, что значительно громче, но при этом громкость диалогов (-27дБ) повысится всего лишь на 13.5дБ, а громкость самых громких звуков (0дБ) вообще не изменится. При  $k > 1$  - громкость будет уменьшаться, а динамический диапазон - увеличиваться.

Посмотрим на графики громкости ( $k = 1/2$ : сжатие ДД в два раза):



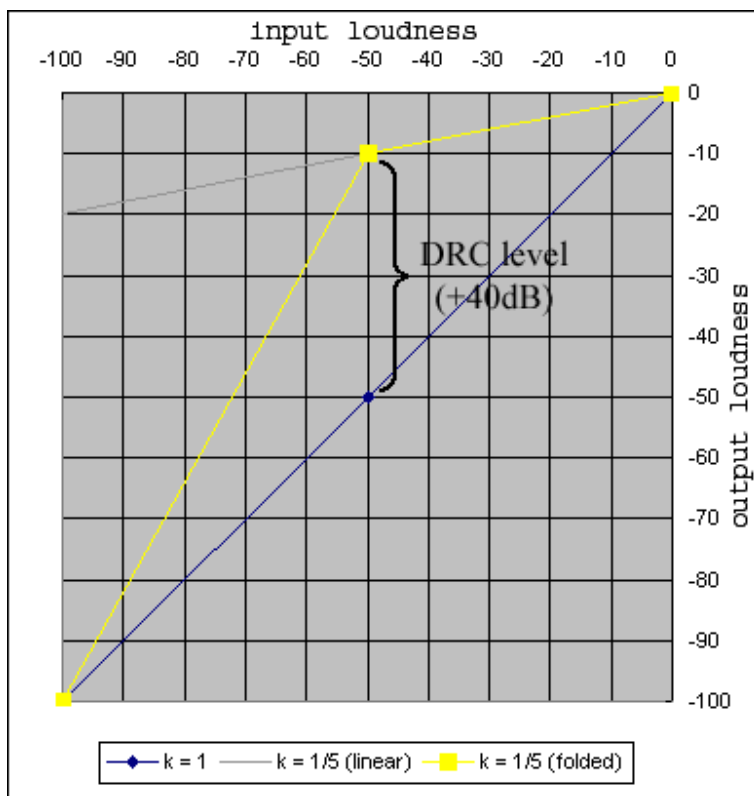
Как видно в оригинале присутствовали как очень тихие звуки, на 30дБ ниже уровня диалогов, так и очень громкие - на 30дБ выше уровня диалогов. Т.о. динамический диапазон составлял 60дБ. После компрессии громкие звуки всего лишь на 15дБ выше, а тихие - на 15дБ ниже уровня диалогов (динамический диапазон теперь составляет 30дБ). Таким образом, громкие звуки стали значительно тише, а тихие - значительно громче. При этом переполнения не происходит!

Теперь обратимся к гистограммам:



Как хорошо видно - при усилении до +30дБ форма гистограммы хорошо сохраняется, что означает, что громкие звуки остаются хорошо выраженными (не уходят в максимум и не обрезаются, как это происходит при простом усилении). При этом выделяются тихие звуки. Гистограмма это показывает плохо, однако разница очень заметна на слух. Недостаток метода - те же самые прыжки громкости. Однако механизм их возникновения отличается от скачков громкости возникающих при обрезании, а их характер отличен - они проявляются в основном при очень сильном усилении тихих звуков (а не при обрезании громких, как при обычном усилении). Чрезмерный уровень компрессии приводит к уплощению звуковой картины - все звуки стремятся к одинаковой громкости и невыразительности.

Сильное усиление тихих звуков может привести к тому, что станут слышны шумы записи. Поэтому в фильтре применен немного модифицированный алгоритм, чтобы уровень шумов поднимался меньше:



Т.е. на уровне громкости -50дБ происходит перегиб передаточной функции, и шумы будут усиливаться меньше (желтая линия). При отсутствии такого перегиба шумы будут значительно громче (серая линия). Такая простая модификация значительно снижает количество шумов даже при очень сильных уровнях сжатия (на рисунке - сжатие 1:5). Уровень “DRC” в фильтре задает уровень усиления для тихих звуков (на уровне -50дБ), т.о. уровень компрессии 1/5, показанный на рисунке, соответствует уровню +40дБ в настройках фильтра.



## 8 Микширование и громкость

Есть еще одна далеко неочевидная тонкость, которая может сильно повлиять на громкость при проигрывании многоканальных записей на стереосистеме (или в наушниках). Пусть мы имеем исходную запись в формате 5.1, а на выходе необходимо получить 2 канала. В каждый из выходных каналов микшируется один фронтальный канал, один тыловой, канал LFE, и часть центрального канала:

$$L' = L + 0.7 * C + SL + LFE$$

$$R' = R + 0.7 * C + SR + LFE$$

Допустим, что по всем входным каналам сразу запущен звук максимальной громкости (хотя это и не слишком вероятно). Тогда амплитуда в выходном канале будет превышать максимум в 3.7 раза (на 11дБ), т.е. произойдет сильнейшее переполнение. Чтобы переполнения не допустить формулы микширования должны быть переписаны следующим образом:

$$L' = (L + 0.7 * C + SL + LFE) / 3.7$$

$$R' = (R + 0.7 * C + SR + LFE) / 3.7$$

Это нормализованное микширование (не путать с нормализацией самой записи!) - переполнения гарантировано не происходит. Однако, смикшированный таким образом звук на целых 11дБ тише! Можно ли не проводить нормализацию? Можно. Однако в таком случае возможно переполнение и соответствующие искажения. Переполнение происходит только тогда, когда по всем входным каналам одновременно воспроизводится громкий звук. Это не характерно для фильмов (обычно тыловые каналы тише фронтальных, да и канал LFE не всегда задействован), но характерно для многоканальных музыкальных записей. Поэтому для фильмов можно оставить микширование ненормализованным, а для музыки нормализацию лучше включать. (В фильтре за нормализацию микширования отвечает опция **“Normalize matrix”**).

## 9 Рекомендации

Рекомендации по настройке сильно зависят от конкретных условий прослушивания и поставленной цели. Условно можно разделить все условия на следующие категории:

- Высококачественные домашние кинотеатры и хорошие условия прослушивания - отсутствуют посторонние шумы (уличный шум, соседи сверху, бегающие вокруг дети) и можно шуметь самим (отсутствие соседей). Для этой категории всю пост-обработку можно отключить (пресет standard).
- Хорошая акустика, но наличие внешнего шума. Можно применить небольшое усиление и небольшой уровень компрессии.

- Нельзя шуметь самим (ночь, соседи/дети) или посредственная акустика (дешевые наушники, слабые колонки). Необходимо сильный уровень компрессии плюс повышение уровня или однопроходная нормализация.

Наилучших результатов можно добиться только при комбинировании всех методов. В данном разделе будет рассматриваться только случай просмотра DVD-фильмов. Для других случаев (музыка, MPEG4-фильмы и пр.) характеристики звука слишком отличаются. Однако, освоив настройку фильтра для DVD-фильмов, уже несложно будет разобраться и с другими случаями.

Усиление (уровень **“Master”**). Скачки громкости, возникающие при переполнении незаметны только в случае, когда обрезаемый звук намного громче основного: в примере с телефоном звук телефона сравним по громкости со звуком голоса, однако если вместо телефона будет оглушительный взрыв, то звук голоса в любом случае будет неразличим. Поэтому ограничение самых громких звуков допустимо и практически незаметно. Очень громкими можно считать звуки от +15..+20дБ относительно уровня диалогов. Таким образом, при уровне диалогов равном -27дБ (де-факто для DVD) допустимый уровень усиления равен +7..+12дБ. В качестве еще одного ориентира можно обратиться к гистограммам усиления: усиление до +10дБ не сильно сказывается на форме гистограммы, в то время как при усилении +20дБ ограниченным оказывается очень большое количество звуков. Т.о. можно считать допустимым усиление до +10дБ. Его можно держать в этом положении практически всегда - на тихих фильмах это заметно поднимет громкость, а на громких - не будет сильно мешать. (Еще раз отметим, что рассуждения приведены только для случая просмотра DVD-фильмов - для музыки и большинства mpeg4-фильмов они неприменимы, т.к. характеристики гистограмм сильно отличаются).

Усиление также может применяться для ограничения громкости сигнала (например, для прослушивания ночью). При уровне диалогов равном -27дБ и усилении +17дБ уровень диалогов после усиления будет составлять -10дБ, а самого громкого звука, как всегда, 0дБ, что всего лишь на +10дБ выше уровня диалогов. Т.о. установив в фильтре усиление +17дБ, и отрегулировав системную громкость так, чтобы диалоги воспроизводились с приемлемой абсолютной громкостью, мы получим гарантию, что самые громкие звуки не будут превышать уровень диалогов больше, чем на 10дБ (хотя, конечно, в этом случае для громких звуков будет применено ограничение сигнала).

Компрессия (**“DRC”**). Границы применения компрессии определить намного сложнее, поскольку слышимость дефектов компрессии сильно зависит от акустической системы, условий прослушивания и самого слушателя. Если подойти к вопросу с точки зрения гистограмм, то верхний предел применимости компрессии можно оценить как +20..+30дБ (см. гистограммы). В этом случае громкие звуки еще остаются выделенными по громкости от слабых. Таким образом, уровень компрессии подбирается только на слух - до тех пор, пока не будет достигнута необходимая громкость, и дефекты будут оставаться незаметными.

Необходимо учитывать, что компрессия и усиление действуют одновременно, поэтому при уровне усиления +20дБ и компрессии в два раза (+25дБ) реальный уровень усиления будет составлять +10дБ. Это нормально, поскольку потребность в усилении также уменьшается с увеличением компрессии.

Однопроходная нормализация (**“One-pass normalize”**). Применима практически всегда. Для высококачественных систем с небольшим уровнем усиления позволит уменьшить количество переполнений, а в остальных случаях в комбинации с компрессией позволит добиться максимальной громкости опять же с минимумом искажений по переполнению. Начальный уровень усиления (уровень **“Master”**) для нормализации выбирается исходя из целей - если большого усиления не нужно - устанавливается желательное, а для достижения максимальной громкости можно установить +20дБ.