

14 травня 2015 р.

Аналоговая и цифровая обработка звукового сигнала. Мифы и реальность.

Попробуем начать с основ, на которых строятся принципы обработки, анализа и преобразований звука.

Итак, диапазон частот, которые может слышать и различать человек, находится в пределах от 15 Гц до 20 кГц. Человек может слышать только аналоговый звук. Т.е. вне зависимости от того как получен сигнал звука, для того, чтоб он был услышан, он всегда должен быть преобразован в аналоговую форму. Метод представления аналоговой формы выберем классический – waveform, т.е. это такая кривая, которая отображает изменение сигнала во времени. Сразу предположим, что мы имеем громкоговорители близкие к идеалу, и форма звукового давления, которое они создают, максимально повторяет форму сигнала, чтобы исключить фактор акустики, так как это отдельный разговор.

I. Природа тембра и окраса звука.

В звуке все крутится вокруг гармонических колебаний, т.е. колебаний синусоидальной формы. Собственно, абсолютно любая другая форма сигнала может быть представлена суммой различных гармонических колебаний – это и есть та самая основа для «тембра» и «окраса» звука.

Те, для кого это очевидно, могут с легкостью пропустить эту часть, однако, думаю найдутся те, кому это и будет интересно.

Откуда же берутся все эти понятия «теплый ламповый звук», «холодный транзисторный звук» и т.п? Каждый звуковой прибор, будь то обработка или

усилитель, вносят определенные изменения в форму сигнала, так как практически все элементы схем являются нелинейными. Представим, что мы имеем идеальный генератор синусоидального сигнала. Вот такого:

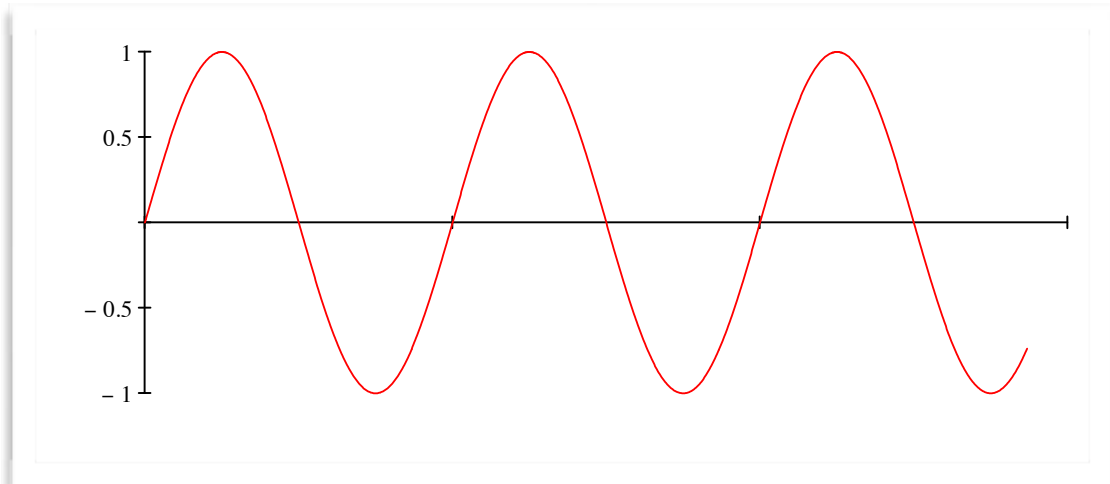


Рис 1. Идеальная синусоида

Подадим этот сигнал на вход устройства, которое, например, имеет нелинейную передаточную характеристику показанную на рисунке 2 синим

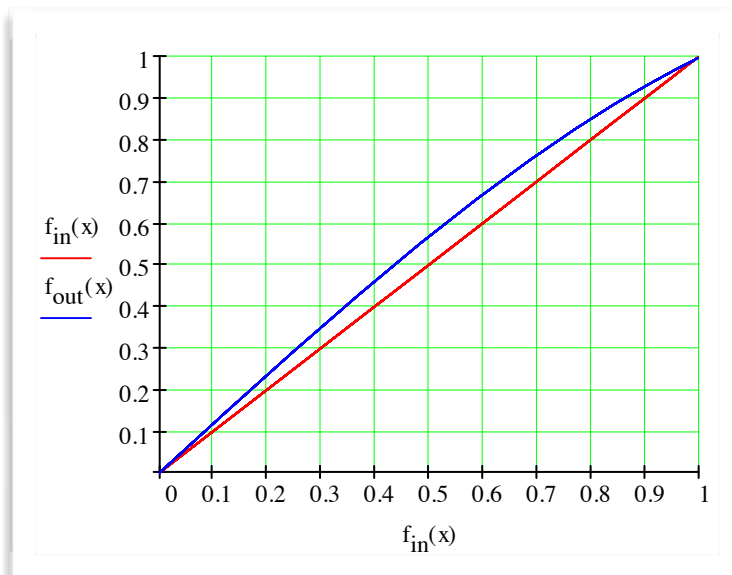


Рис 2. Передаточные функции

цветом, а красным цветом обозначена идеальная линейная передаточная функция (для сравнения).

На выходе устройства будет сигнал отображенный синим цветом на рисунке 3. Красным цветом там отображен

входной сигнал для сравнения.

Визуально форма сигнала практически не отличается. Если отобразить ее без сравнения с входной формой, то разница будет практически не заметна. А

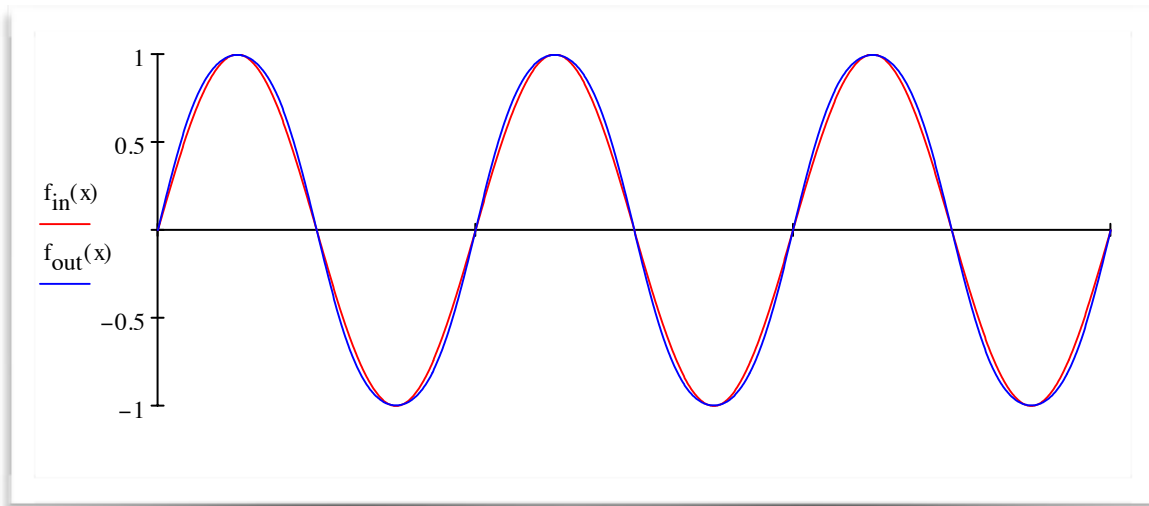


Рис 3. Входной и выходной сигнал

вот на слух они будут отличаться. Давайте посмотрим на спектр входного и выходного сигнала на рисунке 4, чтобы понять характер различий. Красным цветом обозначен спектр входного сигнала, на котором виден единственный пик на первой гармонике (основной частоте). А вот с выходным сигналом все уже

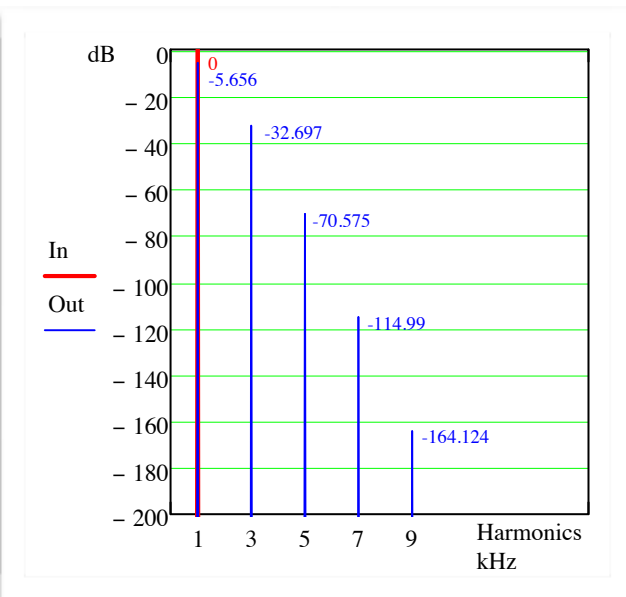


Рис 4. Входной и выходной сигнал

гораздо интересней, там добавились нечетные гармоники (уровень каждой указан на рисунке). Т.е., если мы возьмем 5-ти канальный генератор с частотами кратным гармоникам, и правильно установленными фазами и уровнями и просуммируем их, то мы получим точно такой сигнал как на

выходе. И вот эти гармоники и придают тот самый тембр звуку. Собственно, если записать вокалиста берущим какую-то ноту, то мы получим чем-то похожую картину спектра, где будет ярко выражена основная частота ноты и менее

выраженные частоты четных и нечетных гармоник. Это и есть то, что называется тембром и окрасом звука.

Т.е. любая периодичная форма сигнала отличная от синусоидальной состоит из суммы синусоидальных колебаний частот кратных основной частоте. И из этого следует то, что для частот 10 кГц и выше, вторая гармоника будет 20 кГц и выше, третья 30 кГц и т.д., но человек не слышит частоты выше 20 кГц, а это означает, что человеческое ухо не в состоянии различить форму сигнала на частотах 10÷20 кГц. Эти факты нам еще пригодятся, при рассмотрении цифровых методов обработки сигнала немного позже.

II. Аналог против цифры – что же круче?

Сразу следует сказать, что вопрос сам по себе некорректный, но часто он или его производные становятся предметами споров. Бытует мнение, что «аналог» – это круто и дорого (ну или круто, если дорого), а вот цифра – это такой себе бюджетный ширпотреб. Попробуем разобраться почему существуют такие мнения...

Что есть у аналоговых приборов, чего нет у цифровых или «хуже»:

1. Управление. Привычное (часто считается, что более удобное) управление. Т.е. на каждом аналоговом приборе есть подписанные ручки, кнопки, переключатели и т.д. которые доступны сразу и чаще всего мгновенно оказывают влияние на сигнал. Другими словами у Вас есть большая рэковая стойка с кучей приборов и к каждому регулируемому параметру Вы имеете мгновенный доступ. У цифровых приборов/программ обычно не все так очевидно и есть меню, кнопки и т.д. и часто, чтобы добраться до регулировки какого-то параметра нужно предварить это

несколькими дополнительными манипуляциями, при том, что сами приборы обычно компактней и более функциональны.

2. Время. Аналоговые приборы обрабатывают звук исключительно при помощи физических процессов. И, как результат, время обработки сигнала находится в порядках скорости света разделенной на физическую длину прохождения сигнала. Т.е. в абсолютном большинстве случаев можно считать его пренебрежимо малым или 0. У цифровых приборов здесь неоспоримая брешь в этой сфере. Как известно, цифровая обработка звука начинается с АЦП (аналогово-цифрового преобразователя) и заканчивается ЦАП (цифро-

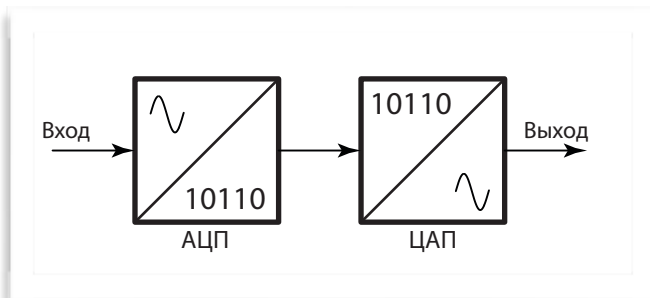


Рис 5. Простейший цифровой прибор

аналоговым преобразователем) между которыми есть «математика». Но даже, если представить, что математики у

нас нет, то прибор с блок-схемой на рисунке 5 будет обеспечивать гарантированную задержку сигнала равную минимум 2 тактам частоты дискретизации (один для АЦП и один для ЦАП). Например, если частота дискретизации 48 кГц, то время задержки будет $\frac{1}{48000} \cdot 2 = 4.1667 \cdot 10^{-5}$ или

41.667 микросекунд (соответственно, при повышении частоты дискретизации время будет уменьшаться обратно-пропорционально). И это минимальная задержка только на преобразование сигнала, а его «математическая» обработка в «середине» прибора – это еще дополнительное время. Обычно минимальное время на обработку – 32 сэмпла (такта частоты дискретизации), что составит – 708.339 микросекунды. Это то, от чего цифровым приборам

вряд ли удастся избавиться в ближайшее время, если вообще удастся. Однако критическое значение это имеет в большинстве случаев только при обработке звука в реальном времени, например живой концерт. (Для справки. При нормальном атмосферном давлении скорость звука составляет 331 м/с, т.е. задержка звука на 41.667 микросекунд будет эквивалентна задержке от звуковых излучателей находящихся примерно на расстоянии 1.4 см от слушателя, а задержка в 708.339 микросекунды – расстоянию 23 см)

На этом реальные и объективные преимущества аналоговых приборов заканчиваются. Маркетинговую и психологическую составляющую мы не учитываем, хотя зачастую они играют решающее значение :)

Теперь про цифровой метод. Цифровые приборы могут абсолютно все тоже самое, что могут аналоговые (не учитывая задержку о которой шла речь немного выше). А вот аналоговые могут далеко не все то, что могут цифровые. Например, аналоговым методом (без привлечения механических составляющих) практически невозможно организовать существенную задержку сигнала, как то реверберация или delay, точный pitch и т.п. Аналоговым методом невозможно осуществлять фильтрацию сигнала не внося в него искажения и шум. Любой аналоговый фильтр вызывает сдвиг фазы сигнала. Собственно последняя «проблема» очень часто как раз и является тем самым характером прибора, то что многие называют «звучит». Любой компонент (даже пассивный) аналоговой схемы вносит нелинейные искажения, а именно: электронный шум, добавляет гармоники и сдвиг фазы сигнала. Сумма всех этих искажений вызванная конкретным схемотехническим решением во многих случаях и обеспечивает «фирмовость» прибора при одинаковой функциональности с аналогичными. Когда просто пропустил сигнал через эквалайзер, немного его настроил и все

«запело», а в другом аналоговом или цифровом устройстве – повторил все настройки, а оно «сухо» и «никак». Можно ли повторить этот эффект в «цифре»?

Да, можно. Математические алгоритмы и современная скорость их обработки позволяют повторять любое поведение сигнала от алгоритмического (когда мы знаем какие трансформации происходят с сигналом и создаем их математическую модель) вплоть до полной имитации работы аналоговой схемы (с помощью таких моделей как SPICE, например). Т.е. разработчик цифрового модуля может заложить и учесть все необходимые сдвиги фаз на разных частотах, добавление гармоник и т.д. вплоть до имитации работы транзистора в каком-то запредельном режиме, в котором реальный транзистор может «прожить» пару минут выдавая особое звучание, или которое сложно реализовать, например, температура охлаждения -10°C . Многие современные цифровые модули (будь то приборы или программы, плагины) делают по принципу аналоговых. Другими словами эквалайзер занимается не просто голой фильтрацией по полосам, а и добавлением своих нюансов, как его аналоговый брат.

Хорошие цифровые плагины стоят не копейки и вот здесь снова часто срабатывает психология материалистов: отдать пачку денег за красивый, увесистый прибор с замечательной табличкой и «звучащий» – это понятно и приемлемо, а отдать не мало, хоть и меньше денег за ссылку по которой можно скачать набор байтов – это вот как-то непонятно. :) Отсюда и произрастает часть мифов про «скупость» звучания «цифры»... Многие, устанавливая программы, для работы со звуком пользуются базовыми инструментами обработки, которые зачастую и не позиционируются как что-то особенное, а часто это как раз

инструменты для минимально необходимых манипуляций, как скотч и канцелярский нож в офисе.

III. Как работает «цифра» ?

Теперь попробуем заглянуть немного внутрь цифровой обработки сигнала, чтоб иметь хотя бы общее представление что да как.

Как уже говорилось выше «путь» цифровой обработки сигнала начинается с АЦП (ADC). Это одно из ключевых устройств, которое имеет очень сильное влияние на качество процесса и является «необратимым» устройством. Т.е. все искажения которые будут внесены этим устройством на этапе преобразования будут необратимы. Какую-то часть как-то можно компенсировать алгоритмами обработки, однако в целом нужна максимальная точность преобразования.

Есть несколько типов АЦП устройств принципиально различающихся

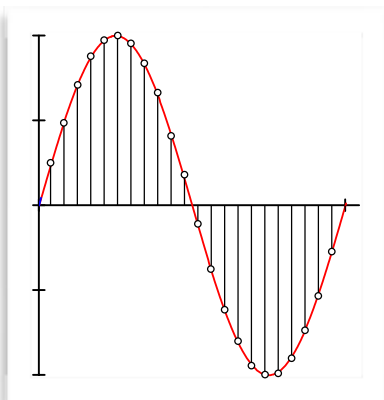


Рис 6. PCM преобразование

типом выхода цифровых данных, такие как PCM, DSD, PWM и т.д. Мы будем говорить о типе PCM, так как это единственный тип цифровых данных подходящий для полноценной обработки звука, все иные типы преобразовываются в тип PCM для полноценной обработки.

PCM (Pulse code modulation) – это импульсно-кодовая

модуляция. Ее смысл заключается в том, что через равные определенные

промежутки времени происходит «измерение» (сэмплирование) уровня сигнала и

передача цифрового значения этого «замера». Ключевым параметром здесь

является **частота дискретизации** – это частота «замеров» (квантизация).

Частота дискретизации определяет предельно возможную частоту

преобразовываемого сигнала – она называется **частотой Найквиста** и равна

ровно половине частоты дискретизации. Частота дискретизации определяет возможный частотный диапазон. Стандартными базовыми частотами дискретизации в обработке звука являются 44.1 кГц и 48 кГц, выше кратные им. Наивысшей стандартизированной частотой дискретизации используемой в обработке звука на сейчас является 352.8 кГц.

И вот здесь возникает вопрос: если предел слышимой частоты – 20 кГц, а частотный диапазон при дискретизации в 44.1 кГц выше слышимого и находится на отметке в 22.05 кГц, то зачем нужна частота дискретизации, например, в 192 кГц? Первый логичный ответ который напрашивается из рисунка 5 – это то, что мы сможем иметь больше «точек» описывающих кривую на высоких частотах и получить большую детализацию сигнала и как результат будем ближе к истинному сигналу. Это объяснение весьма логично, однако не имеет особого практического смысла, так как источники звука выдают сигнал в диапазоне до 20 кГц, а это означает, что в диапазоне 10÷20 кГц присутствуют только синусоидальные составляющие и зная это описывать подробно их форму не имеет особого смысла. На самом деле здесь мяч находится на аналоговой стороне АЦП (извините за тавтологию). Так вот, принципы работы АЦП РСМ типа требуют, чтоб входной сигнал находился строго в диапазоне от 0 Гц до частоты Найквиста, иначе в преобразовании будут ошибки, а это в свою очередь означает, что вход АЦП должен быть оборудован фильтром высоких частот, который будет срезать все, что выше частоты Найквиста, но при этом максимально ровно пропускать все что до нее. Т.е. другими словами, при частоте дискретизации 44.1 кГц необходим фильтр который будет иметь пропускную способность 0 дБ на частоте 20 кГц и -150 дБ (для 24 бит) на частоте 22.05 кГц (рисунок 7). Кто хоть немного знаком со схемотехникой понимает, что построение

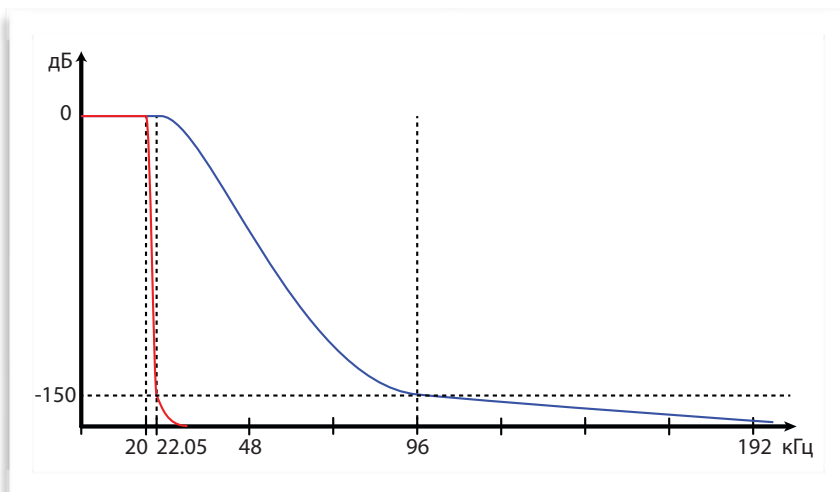


Рис 7. Характеристики фільтрів для різних частот Найквіста (22.05 кГц і 96 кГц)

такого фільтра викликає немало труднощів і приличну його ціну. А ось при частоті дискретизації в 192 кГц і частоті Найквіста 96 кГц фільтр повинен уже мати куди менше

крутої характеристики і

відповідно буде простіше і дешевше в реалізації. Тому при удешевленні технології виробництва цифрових схем, з часом, стало простіше і дешевше підвищувати частоту дискретизації, спрощаючи при цьому фільтр і підвищуючи якість.

Тоді виникає протилежний запитання, а чому не використовувати частоти дискретизації ще вище? Відповідь дуже проста – прироста якості це вже не дасть, а буде підвищувати ціну обробки звуку збільшенням потоку обробок на процесор(и), удорожанням перетворюючих чипів, великим обсягом місця для зберігання (на диску)...

Так само буває думка про те, що при більшій частоті дискретизації краще відбувається *pitching*. Насправді це не зовсім вірне твердження, тому що існують математичні алгоритми, які дозволяють достатньо точно відновлювати проміжні значення вхідного сигналу між зразками, знаючи частоту дискретизації, але цей алгоритм створює додаткову навантаження, в той час як при високій частоті

дискретизации не редко просто берут ближайшее необходимое значение, что значительно проще.

Еще один параметр АЦП – это его разрядность. Разрядность в данном случае – это прямой показатель возможного динамического диапазона обрабатываемого сигнала. Таким образом, разрядность в 8 бит обеспечивает динамический диапазон примерно в 48 дБ, 12 бит – 72 дБ, 16 бит – 96 дБ, 24 бита – 144 дБ, 32 бита – 192 дБ. Отношение сигнал/шум для большинства высококачественных источников звука находится на уровне 110-120 дБ, т.е. разрядность в 24 бита перекрывает с лихвой этот диапазон.

Все АЦП PCM типа обладают таким неприятным явлением как «шум квантизации». Это явление вызвано тем, что АЦП может воспринимать дискретные (конкретные) уровни сигнала. Рассмотрим на примере 4 битного преобразователя для простоты. Например, амплитуда номинального входного напряжения 1В – это означает, что такой АЦП может выдать значение входного напряжения с интервалом в 62.5 мВ, но что произойдет, если уровень входного напряжения в момент квантизации («замера»), например, 67 мВ? В зависимости от конкретного АЦП он воспримет этот уровень как один из граничных между которыми находится. В данном случае между 62.5 мВ и 125 мВ. Предположим, что рассматриваемым АЦП будет взято ближайшее значение, тогда при обратном преобразовании мы получим 62.5 мВ вместо 67 и, соответственно, ошибку в 5.5 мВ. При следующем преобразовании получим новую ошибку и т.д. Величина этой ошибки будет всегда не больше 62.5 мВ в данном случае или эквивалента младшего разряда в общем случае. Однако, не все так страшно, при разрядности в 24 бита получается, что шум квантизации будет находится на

уровне -144 дБ, что за пределами сигнал/шум многих источников звука и не будет оказывать влияния на полезный сигнал.

Мы подошли к очередному мифу из серии «аналог vs цифра», что аналоговый сигнал он такой весь гладенький и непрерывный, а цифровой сигнал весь в клеточку – побитый на интервалы во времени частотой дискретизации и амплитуде разрядностью. Так вот, качественно преобразованный цифровой сигнал тоже непрерывный и гладкий на столько, что Вы не заметите эти «клеточки». Например, источник сигнала с соотношением сигнал/шум 120 дБ равнозначен 20 битному звуку, поэтому все, что будет находится между дискретными интервалами в 20 бит – это будет просто шум, таким точно шумом можно заполнить младшие 4 бит в 24 битном звуке и вы не увидите и не услышите никаких «клеточек» и ступенек в сравнении с таким аналоговым источником звука.

Для справки. Обычные качественные компакт-кассеты обладают эквивалентом 5-6 битного диапазона, компакт-кассеты с металлической пленкой 8-10 битного как и катушечные магнитофоны, профессиональные катушечные магнитофоны с профессиональной пленкой 12-14 битного. Именно поэтому многие любят до сих пор использовать пленку как естественный компрессор. К тому же, если Вы посмотрите на АЧХ характеристику пленки, то увидите, что частота там далеко не бесконечна, а заваливается после 12-14 кГц. Это происходит потому, что на пленочных носителях она (частота) тоже дискретизирована. :) Только не «цифровыми выборками», а доменами магнитного материала из которого изготовлена пленка и которые имеют конечную и вполне определенную длину и сохраняют одно положение для магнитного потока, т.е. другими словами, можно взять пленку и абсолютно так

же поділити її на відрізки як малюнок 5. І абсолютно той же самий ефект і принцип дискретизації відбувається при русі плівки вздовж головки магнітофона. Іменно тому збільшуючи швидкість на катушковому магнітофоні ми отримуємо кращу передачу високих частот – ми збільшуємо частоту дискретизації.

Після перетворення звуку в цифровий вигляд, починається найчистіша математика і це неймовірний обсяг інформації, освітлювати який не входить в плани цієї статті.

В кінці оброблений звук, частіше за все перетворюється в аналоговий за допомогою ЦАП, за принципом протилежному АЦП. Т.е. ЦАП отримує значення рівня сигналу з частотою дискретизації і формує на виході відповідний рівень напруги. І ось тут у багатьох виникає в голові стереотипна картинка з сходишками :) На справді, якщо Ви підключите осцилограф до якісного ЦАП – Ви не побачите ніяких сходинок, а побачите гладеньку, рівеньку аналогову криву, да ще і з мінімальним рівнем шумів. Чому це відбувається? Тому що на виході ЦАП так само обладнані фільтрами високих частот, які зрізають все вище потрібного для звуку діапазону, але, якщо Ви згадаєте початок статті, то там говорилося, що для різких фронтів (прямокутного) сигналу потрібне наявність n -кратних гармонік до нескінченності, а наша основна гармоніка «сходинок» – це частота дискретизації, т.е. потрібні частоти від частоти дискретизації і вище кратні частоті дискретизації, які благополучно «режуться» згладжуючи сходишки в красивий і чистий звук :)

Все описане поширюється і на аналогові з псевдоаналоговими синтезаторами за якими багато так люблять. Як раритет, експонат,

моральное наслаждение и кайф на ощупь – да, как источник звука – не сложно полноценно заменяется цифровыми средствами.