

**Вологдин Э.И.**

**СТАНДАРТЫ И СИСТЕМЫ  
ЦИФРОВОЙ ЗВУКОЗАПИСИ**

**Санкт-Петербург  
2012**

## СОДЕРЖАНИЕ

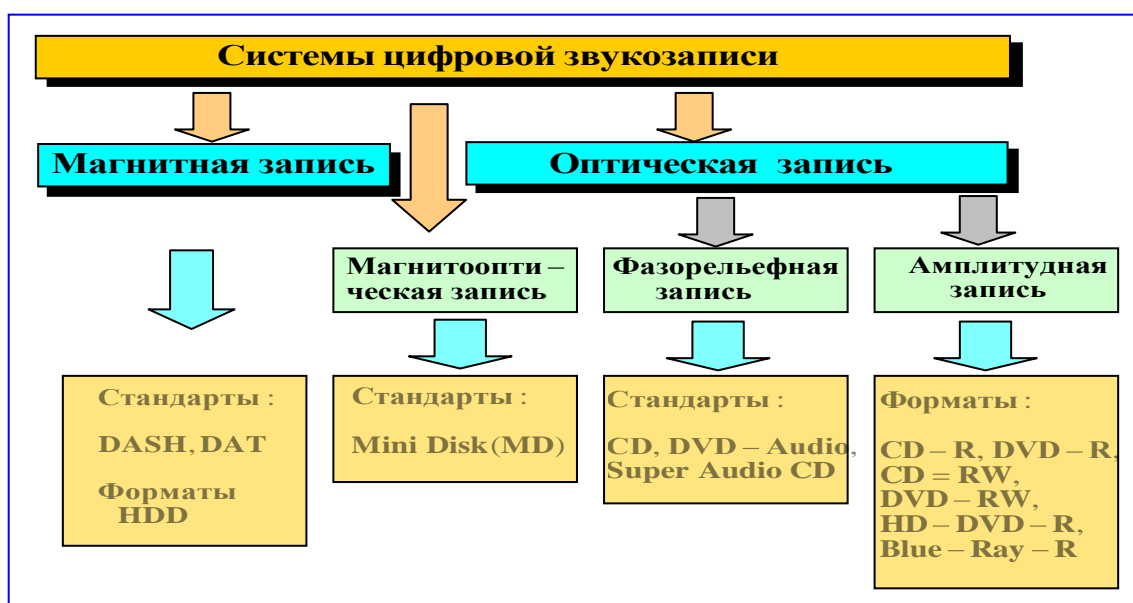
1. СИСТЕМЫ, СТАНДАРТЫ И ФОРМАТЫ ЦИФРОВОЙ ЗВУКОЗАПИСИ .....	4
2. ЦИФРОВОЙ ТРАКТ АППАРАТУРЫ ЗВУКОЗАПИСИ .....	8
2.1. Функциональная схема тракта.....	8
2.2. Общие характеристики цифровых трактов.....	10
3. ЛИНЕЙНАЯ ИМПУЛЬСНО-КОДОВАЯ МОДУЛЯЦИЯ.....	14
3.1. Назначение и принцип работы .....	14
3.2. Модуляция.....	14
3.3. Демодуляция .....	18
3.1. Назначение и принцип работы .....	21
3.2. Модуляция.....	21
3.3. Демодуляция .....	25
4. СПЕКТР И ЗВУЧАНИЕ ОШИБОК КВАНТОВАНИЯ .....	29
4.1. Спектр ошибок округления .....	29
4.2. Спектр ошибок квантования .....	30
4.3. Звучание ошибок квантования .....	37
5. СПЕЦИАЛЬНЫЕ АУДИО ТЕХНОЛОГИИ .....	41
5.1. Технология Dithering .....	41
5.2. Технология Noise Shaping.....	47
5.3. Технология Super Bit Mapping (SBM) .....	52
5.4. Технология Oversampling .....	53
6. СИГМА-ДЕЛЬТА МОДУЛЯЦИЯ.....	57
6.1. Технология Noise Shaping.....	57
6.2. Математические модели сигма-дельта модулятора .....	58
6.3. Одноразрядный АЦП на основе сигма-дельта модуляции.....	61
6.4. Одноразрядный ЦАП на основе сигма-дельта модуляции.....	65
6.5. Многоуровневые ЦАП .....	68
7. КОДИРОВАНИЕ ЗВУКОВЫХ СИГНАЛОВ .....	72
7.1. Корректирующее кодирование.....	72
7.2. Канальная модуляция .....	83
8. ОПТИЧЕСКАЯ ЗАПИСЬ ЗВУКА .....	89
8.1. Стандарты и форматы оптической записи .....	89
8.2. Физические основы оптической записи.....	91
8.3. Рельефно-фазовая запись.....	94
8.4. Амплитудная запись .....	99
8.5. Системы автоматического позиционирования .....	102
8.6. Стандарты корректирующего кодирования CD и DVD .....	107
8.7. Стандарты канальной модуляции CD и DVD.....	112
9. МАГНИТНАЯ ЗАПИСЬ ЗВУКА.....	116
9.1. Физические основы цифровой магнитной записи .....	116
9.2. Многодорожечные цифровые магнитофоны стандарта DASH.....	121
9.3. Стереофонические магнитофоны стандарта DAT .....	123
9.4. Многоканальные дисковые магнитофоны формата HDD .....	129
10. МАГНИТООПТИЧЕСКАЯ ЗАПИСЬ ЗВУКА .....	133

10.1. Физические основы магнитооптической записи .....	133
10.2. Стереофоническая звукозапись по стандарту MD .....	134
10.3. Форматы магнитооптической записи.....	135
11. ЦИФРОВЫЕ ЗВУКОВЫЕ ИНТЕРФЕЙСЫ.....	137
11.1. Общие сведения .....	137
11.2. Интерфейсы AES3, AES3id и SPDIF .....	139
11.3. Интерфейсы AES3 и AES/EBU.....	141
11.4. Интерфейсы SDIF-2, SDIF-3 и S/PDIF .....	144
11.5. Интерфейс AES10 .....	145
11.7 Интерфейс MIDI.....	149
СПИСОК ЛИТЕРАТУРЫ.....	154

## 1. Системы, стандарты и форматы цифровой звукозаписи

В радиовещании звукозапись производится в процессе подготовки, монтажа и архивации речевых и музыкальных программ. Звуковая программа может быть в виде одного канала (моно), это речь диктора, чтение стихов и любой разговорный жанр. Чаще используются два канала (стерео), это большинство музыкальных программ. В системах цифрового радиовещания число каналов может быть до 6..8, что обеспечивает пространственное звучание.

Нередко при первичной записи оркестровой музыки возникает необходимость одновременной записи отдельных музыкальных инструментов и тогда число записываемых звуковых каналов достигает 24..48, а иногда и больше. Звуковые каналы могут записываться по отдельным дорожкам или в одной с использованием временного



уплотнения.

Запись звуковых сигналов осуществляется на движущийся ленточный или дисковый носители, имеющие специальное покрытие для регистрации звуковой информации в цифровой форме. В процессе записи в покрытии носителя происходят химические, магнитные, оптические или другого рода физические изменения, формирующие дорожку фиксированной ширины которые несут информацию о звуке. Такая дорожка называется *цифровой фонограммой*.

В радиовещании используются *4 системы звукозаписи*, отличающиеся физическими принципами регистрации цифровой информации: *магнитная, оптическая и магнитооптическая*. Для профессиональных целей пока больше подходит магнитная или магнитооптическая запись звука. Рельефно-фазовая запись применяется только при массовом тиражировании цифровых звукозаписей на оптических дисках, которые являются готовыми источниками музыкальных программ. Амплитудная оптическая запись используется для архивации цифровых фонограмм.

*Стандарт цифровой звукозаписи* определяет совокупность данных, включающих в себя технические характеристики носителя с записью звуковой информации, используемые способы аналого-цифрового преобразования, принципы записи и считывания цифровых данных, а также методы помехоустойчивого кодирования и исправления ошибок. В настоящее время цифровая звукозапись производится в трех стандартах: CD-DA, SACD и DVD-Audio, основные характеристики которых приведены в [табл.1](#). Как видно, эти стандарты отличаются главным образом характеристиками аналого-цифрового преобразования и числом звуковых каналов. В профессиональной

звукозаписи используются два основных стандарта цифровой магнитной звукозаписи DASH и DAT.

**Формат цифровой звукозаписи** определяет совокупность данных, включающих в себя технические характеристики носителя с записью *любых цифровых данных* (звук, видео...), организацию и характеристики цифрового потока, принципы записи и считывания, а также методы помехоустойчивого кодирования и исправления ошибок. В профессиональной звукозаписи используется множество форматов записи: в магнитных системах –HDD, в оптических системах – CD-R, CD-RW, DVD-R, DVD-RW, Blue-Ray и HD-DVD. В этих форматах запись звука может производиться в разных стандартах.

**Магнитная запись** осуществляется магнитным полем, формируемом в очень узком зазоре записывающей головки. При этом вдоль дорожки происходит намагничивание магнитного покрытия НЗ в соответствии со звуковой информацией. Для воспроизведения такой записи используются *индукционные* или *магниторезистивные*

Стандарты	Модуляция	Разрядность	Частота выборки	Число каналов	Компрессия
CD-DA	ИКМ	16	44,1 кГц	2	нет
SACD	SDM	1	2822,4	2+5,1	Есть (без потерь)
DVD-A	ИКМ	16...24	48...192 кГц	5,1	Есть (с потерями)
DASH	ИКМ	16, 24	32, 48 кГц	2...48	нет
DAT	ИКМ	16, 24	48	2	нет

головки считывания. В качестве НЗ могут использоваться узкие и широкие магнитные ленты, а также жесткие диски компьютеров.

До настоящего времени этот вид записи является ведущим в радиовещании, так как он обеспечивает наибольшую длительность непрерывной записи звука и на его основе создан громадный парк магнитофонов самого различного назначения. Серьезным недостатком является относительно малое время хранения магнитной записи, которое обычно не превышает 10...30 лет.

В стандарте **DASH** осуществляется многодорожечная многоканальная запись звука на широкой магнитной ленте. В стандарте **DAT** используется наклонно-строчечная запись стереофонического сигнала на узкую магнитную ленту. В формате **HDD** производится многоканальная запись звука на вращающемся жестком диске компьютера в форме концентрических дорожек с секторной организацией данных.

**Оптическая запись** звука осуществляются на дисковых *оптических носителях* с основой из прозрачного полимера. Запись и считывание звуковой информации производится лазерным излучением, сфокусированным в пятно диаметром 0,3...1 мкм. Практически она реализуется в двух вариантах, отличающихся способами регистрации цифровой информации и назначением.

**Рельефно-фазовая запись** используется для массового изготовления дисков с записью музыкальных программ. Запись и тиражирование таких дисков производится принципиально также, как изготавливаются виниловые грампластинки, но на более высоком технологическом уровне штамповки. При этом используется механический способ записи, при котором информация хранится на диске в виде изменений рельефа его поверхности, как это делается на грампластинках. **Штампованные диски** выпускаются на специализированном производстве. На этих дисках цифровая звуковая информация регистрируется вдоль *спиральной* дорожки в виде последовательности *микро углублений* различной длины и, что очень важно, заданной глубины. Благодаря этому при считывании возникает *фазовая модуляция* отраженного светового потока. В радиовещании эти диски используются в качестве источника готовых музыкальных программ.

**Амплитудная оптическая запись** широко используется в профессиональной и бытовой звукозаписывающей аппаратуре. В этом случае цифровая информация регистрируется в виде локальных изменений коэффициентов отражения светочувствительного покрытия

диска. Для осуществления амплитудной записи массово тиражируются подложки этих дисков с «немой» канавки с нанесением покрытий с различными характеристиками. В зависимости от свойств покрытия запись может производиться только однократно или возможно стирание и повторная многократная запись.

*Форматы дисков CD-R, DVD-R, Blue-Ray и HD-DVD* предназначены только для однократной записи. Они используются для архивной записи звука в студийных условиях в любых стандартах. Поверхность дисков этих форматов покрывается светочувствительным покрытием на основе органического красителя, у которого при записи под действием лазерного излучения меняется *коэффициент пропускания* в точном соответствии с записываемой цифровой информацией.

*Форматы CD-RW и DVD-RW* предназначены для многократной записи. Перезаписываемые диски отличаются тем, что запись на них можно стереть и произвести повторно почти бесконечное число раз. Для обеспечения возможности записи и стирания информации покрытие изготавливается из материалов, которые под действием нагревания от лазерного излучения могут переходить из кристаллического состояния в аморфное и обратно. В радиовещании такие оптические диски часто применяются в студиях для монтажа фонограмм вместо магнитных дисков.

*Магнитооптическая запись звука* осуществляется на дисковых носителях с основой из прозрачного полимера. Запись и считывание цифровой информации производится лазерным излучением, сфокусированным в пятно диаметром 0,5...1 мкм. Магнитооптическая запись осуществляется слабым магнитным полем при одновременном действии достаточно мощного сфокусированного лазерного излучения. Под влиянием этих двух физических процессов намагничивание осуществляется только в областях дорожки, где температура от лазерного излучения выше точки Кюри. При этом используется специальный магнитный НЗ, у которого при намагничивании меняется *угол поляризации* отраженного считывающего светового пучка. Магнитооптическую запись можно стирать с помощью лазерного излучения и повторно записывать, поэтому технические характеристики магнитных и *магнитооптических дисковых рекордеров* достаточно близки, и они используются в процессе монтажа многоканальных фонограмм большого объема.

По ориентировочной оценке длительность хранения записи на оптических и магнитооптических дисках профессионального назначения может достигать 100...250 лет.

*Развитие стандартов цифровой записи звука* происходит в направлениях повышения качества звучания. Для этого используется большее число разрядов, повышается частота дискретизации и происходит переход от стереофонической записи звука к 6-канальной, обеспечивающей более качественное пространственное звучание. Для увеличения непрерывной длительности записи в новых оптических форматах, таких как HD-DVD\*а также Blue-ray Disc, существенно уменьшена длина волны излучения лазера и повышена апертура фокусирующего объектива, благодаря чему уменьшен диаметр сфокусированного лазерного излучения до 0,3 мкм.

По своему назначению цифровая звукозаписывающая аппаратура разделяется на следующие группы:

- *студийная аппаратура* записи и воспроизведения звуковых сигналов (цифровые магнитофоны, студийные магнитофоны со стерео- и многоканальной записью звука, многоканальные дисковые магнитные рекордеры, магнитооптические накопители, лазерные рекордеры и профессиональные *проигрыватели оптических дисков различных форматов*);
- *репортерская аппаратура* звукозаписи (катушечные и кассетные цифровые стереомагнитофоны, магнитооптические рекордеры);
- *бытовая аппаратура* записи и воспроизведения звука (бытовые плееры оптических дисков, стереомагнитофоны и кассетные магнитные плееры).

Отличие этих 3 групп заключается, прежде всего, в технических возможностях аппаратуры.

Первичная студийная запись в аналоговой форме на современных магнитных лентах обеспечивает достаточно высокое качество, поэтому до сих пор она часто используется для создания цифровых дисков в формате CD-DA. Не редко цифровые записи на оптических дисках, изготовленные на основе аналоговых фонограмм, звучат хуже, чем

исходный оригинал

Абревиатура стандартов и форматов записи	
CD-DA	Compact Disc Digital Audio
CD-R	Compact Disc Recordable
CD-RW	Compact Disc Rewritable
SACD	Super Audio Compact Disc
DVD-A	Digital Versatile Disc Audio
DVD-R	Digital Versatile Disc Recordable
HD-DVD	Digital Versatile Disc Rewritable
Blue Ray	Blue Ray
HD-DVD	High Density DVD
HDD	Hard Disc Digital
DASH	Digital Audio Stationary Head
DAT	Digital Audio Tape Recorder
MO	Magneto-optical disks

Несмотря на значительные достижения в системах оптической звукозаписи, емкость цифровых магнитных ленточных и дисковых носителей *пока* значительно выше. Современные магнитные пластины жестких дисков компьютеров имеют емкость до 80 Гбайт. На такой пластине может размещаться до 200 минут записи 24 звуковых каналов при 24 разрядном аналого-цифровом преобразовании с частотой дискретизации 48 кГц. Компьютер с таким жестким диском фактически

является многоканальным цифровым магнитофоном.

Студийные цифровые многодорожечные магнитофоны обладают высокими техническими характеристиками, но крайне сложны как по устройству, так и в эксплуатации. Поэтому в студийной технике звукозаписи основным направлением развития является использование компьютерных технологий, как для монтажа фонограмм, так и для звукозаписи на основе оптических и магнитных дисковых носителей.

---

*Абревиатура* .\*: CD-DA (Compact Disc Digital Audio), SACD (Super Audio Compact Disc), DVD-A(Digital Versatile Disc Audio), CD-R (Compact Disc Recordable), CD-RW (Compact Disc Rewritable), DVD-R (Digital Versatile Disc Recordable), DVD-RW (Digital Versatile Disc Rewritable), HDD (Hard Disc Digital), DASH (Digital Audio Stationary Head), DAT (Digital Audio Tape Recorder), HD-DVD (High Density DVD), BD (Blue Ray), MO (Magnetic Optical), MD (Mini Disc).

## 2. Цифровой тракт аппаратуры звукозаписи

### 2.1. Функциональная схема тракта

В профессиональной аппаратуре звукозаписи используются цифровые тракты с двумя видами аналого-цифрового и цифро-аналогового преобразований (АЦП и ЦАП): на основе импульсно- кодовой модуляции (ИКМ) и на основе сигма-дельта модуляции ( $\Sigma\Delta$ -модуляция). Последняя стала широко использоваться только несколько лет назад. По техническим характеристикам в настоящее время эти два вида модуляции примерно равнозначны.

При ИКМ модуляции аналого-цифровое преобразование осуществляется путем дискретизации и квантования аналогового ЗС на относительно низкой частоте дискретизации  $f_s$ . С этой частотой цифровые звуковые данные передаются по тракту в виде непрерывной последовательности кодовых слов с достаточно большим числом разрядов  $q$  от 16 до 24. Выбор частоты дискретизации (частоты выборки) основывается на широко известном неравенстве из теоремы В.П.Котельникова. Для характеристики цифровых трактов часто используется частота Найквиста  $f_N$ , равная половине частоты дискретизации. Она определяет максимально допустимую частоту спектра ЗС на входе АЦП.

При  $\Sigma\Delta$ -модуляции дискретизация аналогового сигнала осуществляется на очень высокой частоте, во много раз превышающей требования теоремы В.П.Котельникова. Для преобразования звукового сигнала в цифровой код используется всего от 1 до 10 двоичных разрядов. При одноразрядном аналого-цифровом преобразовании формируется непрерывная последовательность двоичных данных, без разделения их на кодовые слова, символы или блоки.

Пока на большинстве радиовещательных студий используется аппаратура цифровой звукозаписи на основе ИКМ, поэтому ей будет уделено больше внимания. На [рис.2.1](#). приведена обобщенная функциональная схема цифрового тракта записи-воспроизведения звука, которая имеет  $N_{ch}$  входных и выходных звуковых каналов. Запись производится с



Рис.2.1. Обобщенная функциональная схема записи-воспроизведения звука с АЦП и ЦАП на основе ИКМ

мультиплексированием всех каналов в единый цифровой поток. Эта схема справедлива также для монофонической и стереофонической записи. Она в равной мере относится к магнитной, оптической и магнитооптической аппаратуре цифровой звукозаписи. В приведенной схеме носитель записи может быть ленточным или дисковым, головки записи и считывания могут быть фиксированными или вращающимися.

Функциональные схемы конкретных цифровых устройств записи-воспроизведения может несколько отличаться от приведенной схемы. Это зависит, прежде всего, от числа и типа используемых записывающих и считывающих головок. Они могут быть



оптическими, магнитными, магнитооптическими, индукционными и магниторезистивными. Головки могут предназначаться только для записи или воспроизведения или быть универсальными. Запись и воспроизведение могут выполняться раздельно или одновременно.

Входной блок функциональной схемы на рис.1.2. состоит  $N_{ch}$  АЦП, в выходной блок входит такое же число ЦАП. Все они работают на основе использования линейной импульсно-кодовой модуляции. АЦП преобразуют аналоговые сигналы звуковых каналов в последовательности кодовых слов. В ЦАПах производится обратное преобразование последовательности кодовых слов в аналоговые сигналы звуковых каналов.

Для исправления кодовых ошибок в звукозаписи используются только блочные корректирующие коды. Кодер корректирующего блочного кода может включать в себя несколько ступеней кодирования. В каждой ступени вычисляются проверочные символы, и осуществляется межблочное перемежение цифровых данных с целью исправления в декодере коротких и длинных пакетов ошибок. Выходным сигналом кодера является последовательность двоичных данных, организованных в блоки.

Канальный модулятор преобразует двоичную последовательность данных в сигнал записи в виде непрерывной однополярной последовательности импульсов и интервалов между ними дискретной длительности. В изменениях длительности этих временных и интервалов передается цифровая информация. Под влиянием импульсного тока головка записи формирует на носителе последовательность зон различной намагниченности или питы в виде углублений поверхности носителя, длина которых меняется дискретным образом в соответствии передаваемой цифровой информацией. Головка считывания преобразует эти изменения свойств поверхности носителя записи в электрический сигнал. При канальной демодуляции воссоздается исходная последовательность двоичных данных.

Декодирование корректирующего кода также производится в нескольких ступенях, как и в кодере. В каждой из них производится деперемежение цифровых данных, при котором восстанавливается исходная последовательность кодовых комбинаций. Затем осуществляется обнаружение ошибочных блоков, Если они есть, определяются местоположения ошибок и производится их исправление. Сигнал с декодера корректирующего кода через демультиплексор подается на блок из  $N_{ch}$  ЦАП, которые формируют аналоговые сигналы во всех звуковых каналах.

Рассмотренная функциональная схема может быть названа «классической». В ней по тракту не меняются частота дискретизация и число двоичных разрядов в кодовых словах. Поэтому ее проще анализировать и изучать. Функциональная схема

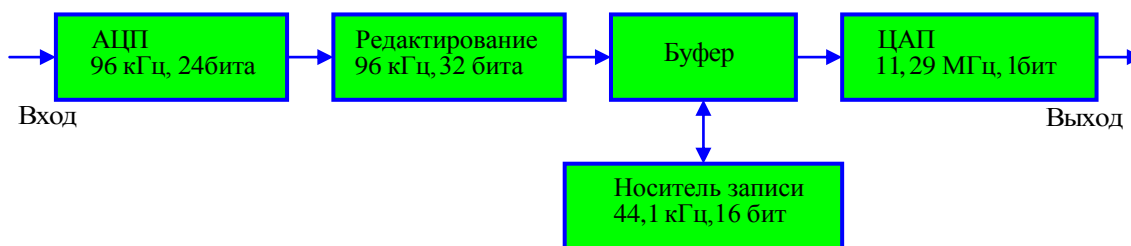


Рис.2.2. Комбинированный тракт на основе ИКМ и  $\Sigma\Delta$ -модуляции

современных цифровых трактов аппаратуры записи-воспроизведения звука значительно сложнее рассмотренной.

На рис.2.2. приведена упрощенная блок-схема одного из таких трактов. В этой схеме АЦП 24-х разрядный на основе ИКМ. ЦАП многоразрядный, выполнен на основе  $\Sigma\Delta$ -модуляции с использованием передискретизации частоты и преобразователя формы огибающей спектра шума квантования. Этот преобразователь известен под названием «Noise Shaping», который многократно увеличивает SNR, что эквивалентно увеличению числа разрядов до 20-24 бит. Целью формирования такого комбинированного тракта

является повышение качества звучания без увеличения емкости носителя записи. В этом случае имеется также возможность хранить в фонотеке отредактированную цифровую фонограмму очень высокого качества.

## 2.2. Общие характеристики цифровых трактов

Частота выборки и число двоичных разрядов в кодовом слове. В профессиональной аппаратуре звукозаписи число разрядов в слове  $q$  при ИКМ варьируется в пределах: 16, 20 и 24, а частота выборки выбирается из ряда: 44,1; 48,0; 88,2; 96,0; 176,4; и 192 кГц. В этом ряду присутствуют частоты бытового ряда, так как часто в радиовещании источником музыкальных программ являются компакт диски. В последних разработках АЦП импульсно- кодовая модуляция выполняется в формате 24/96 и отдается предпочтение варианту, приведенному на рис.1.3., где по тракту оптимальным образом меняются частота дискретизации и число разрядов в словах.

*Скорость цифрового потока аудиоданных*  $V_{ad}$  после мультиплексирования выходных данных АЦП определяется равенством

$$V_{ad} = N_{ch} \cdot f_s \cdot q, \text{ бит/с.}$$

В современной высококачественной цифровой аппаратуре звукозаписи число каналов может быть от 2 до 6, частота дискретизации – от 44,1 кГц до 192 кГц, а число разрядов - от 16 до 24. В многоканальных студийных магнитофонах и рекордерах число каналов может достигать 24....48. Поэтому величина  $V_{ad}$  в различных моделях аппаратуры звукозаписи меняется в очень широких пределах от 1,4 до 220 Мб/с.

*Скорость передачи цифровых данных.* При цифровой записи звука всегда производится корректирующее кодирование, при котором вводятся избыточные проверочные данные. Для этих кодов коэффициент избыточности может быть от 20 до 50%. Поэтому скорость цифрового потока передачи данных  $V_d$  в 1,2.....1.5 раза больше  $V_{ad}$ . Одновременно с музыкальной программой всегда записывается дополнительная информация для потребителя (номер и название альбома, номер фрагмента, название музыкального произведения, фамилии композитора и исполнителей, текущее время и др.). Это увеличивает скорость цифрового потока данных еще на 5...10%.

*Передискретизация* («oversampling»)–это повышение частоты дискретизации. Передискретизация может быть аналоговой, когда с повышенной частотой делаются выборки ЗС, или цифровой, когда между уже существующими выборками после квантования ЗС вставляются дополнительные, рассчитанные путем интерполяции. В АЦП используется аналоговая передискретизация, в ЦАП - цифровая. Применение передискретизации существенно повышает качество звучания цифровой аппаратуры. Это происходит за счет увеличения отношения сигнал-шум (SNR) и уменьшения частотных, фазовых и нелинейных искажений во входных и выходных ФНЧ. Применение передискретизации упрощает конструкцию этих фильтров и существенно уменьшает требования к их крутизне спада. АЦП с передискретизацией могут обходиться без схемы выборки – хранения.

*Технология «Dithering».* Суть этой технологии заключается в декорреляции ошибок квантования путем добавления к звуковому сигналу перед АЦП шума малого уровня, соизмеримого с шагом квантования. Благодаря этому дополнительному шуму детерминированные ошибки квантования преобразуются в шум квантования с равномерной спектральной плотностью мощности (белый шум). Что очень важно, при этом линеаризуется передаточная функция квантователя для малых уровней ЗС. Все это приводит к очень заметному улучшению качества звучания ЗС, поэтому такая операция производится во всех современных высококачественных АЦП.

*Технология «Noise Shaping».* Содержанием этой операции является преобразование формы огибающей спектральной плотности мощности шума квантования таким образом,

чтобы резко снизить уровень спектральных составляющих в звуковом диапазоне частот и перенести их в область высоких частот, где человек их не слышит. Такая технология используется при аналого-цифровом и цифро-аналоговом преобразованиях. Она является основой сигма дельта модуляции.

*Компрессия звуковых данных при записи.* В профессиональной технике, когда недостаточно емкости носителя, для обеспечения необходимого времени записи, производится компрессия скорости цифрового потока, при которой потерь данных не происходит. Таким путем, удастся снизить скорость цифрового потока при звукозаписи на 30 ... 50%. В соответствии со стандартами при записи дисков DVD-A применяется компрессия MLP (Meridian Lossless Packing), а при записи SACD дисков – DST (Direct Stream Transfer).

В бытовой цифровой звукозаписи для понижения скорости цифрового потока используются компрессирование с потерями (Lossy-Compressed). Благодаря такой компрессии скорость цифрового потока понижается от 4 до 10 раз. Такая компрессия производится при записи оптических дисков в форматах MP3 и MD, а также при записи звукового сопровождения кинофильмов на дисках DVD-V. Для этих целей разработано множество систем компрессирования, которые основываются на использовании особенностей слуха человека, такие как Musicam, DTS, AC-3 и другие. В этих системах звуковая информация, которую человек не воспринимает на слух, в тракт записи не передается. Пока такое компрессирование в профессиональной технике звукозаписи не используется.

*Корректирующее кодирование* предназначено для обнаружения и исправления кодовых ошибок как случайного характера, так и в виде длинных пакетов или выпадений, возникающих из-за шумов и дефектов носителя записи. Оно характеризуется исправляющей способностью, под которой понимается величина уменьшения вероятности ошибок на выходе декодера по отношению к входу. При современной технологии изготовления носителей вероятность ошибочных бит не превышает  $10^{-3} \dots 10^{-4}$ , а вероятность их на выходе декодера находится в пределах  $10^{-12} \dots 10^{-15}$ .

*Канальная модуляция* (кодирование) предназначена для преобразования двоичной последовательности в двухуровневый сигнал записи в виде последовательности временных интервалов дискретной длины и согласования спектра записываемого сигнала с частотной характеристикой тракта записи-воспроизведения. Наиболее важными характеристиками канального кода являются минимальный  $T_{\min}$  и максимальный  $T_{\max}$  временные интервалы, определяющие верхнюю и нижнюю границы спектра.

*Скорость записи*  $V_{rec}$  определяет минимальную длину элементов цифровой фонограммы и физическую плотность записи. Эти элементы в оптических системах при рельефно фазовой записи называются питами (pit – углубление), а при магнитной и оптической амплитудной записи – метками. Наиболее важным параметром такой фонограммы является минимальная длина пита (метки), от которой зависит требуемая разрешающая способность считывающей головки (магнитной или оптической).

$$l_{p \min} = T_{\min} \cdot V_{rec}$$

Для воспроизведения записи зазор считывающей магнитной головки или диаметр считывающего пятна света не должны превышать значения  $l_{p \min}$ . Чем выше скорость записи, тем больше длина пит, тем меньше требования к разрешающей способности головок, но соответственно уменьшается время записи.

*Линейная плотность записи* определяет емкость носителя записи. Она зависит от скорости цифрового потока данных  $V_d$ , скорости записи  $V_{rec}$  и коэффициента плотности канального кода  $K_d$

$$D_{lin} = \frac{V_d \cdot K_d}{V_{rec}}$$

В зависимости от выбора канального кода коэффициент плотности может быть от 0,8 до 1,5. Максимальная линейная плотность зависит только от разрешающей способности считывающей головки с зазором или пятном размером  $\delta$

$$D_{lin\max} = \frac{K_d}{\delta}$$

*Емкость дискового носителя*  $C_d$  определяет длительность записи. Она зависит от линейной плотности записи и длины спиральной дорожки  $L_{sp}$

$$C_d = D_{lin\max} \cdot L_{sp}$$

где

$$L_{sp} = \frac{\pi \cdot (R_{\max}^2 - R_{\min}^2)}{\Delta_{sp}},$$

$R_{\min}$  и  $R_{\max}$  – границы зоны записи,  $\Delta_{sp}$  – шаг спиральной дорожки. Фирмы изготовители приводят емкость дисков для аудиоданных, которая на 30...40% меньше значения  $C_d$ , с ее помощью определяется время записи

$$T_{rec} = \frac{C_a}{V_a}.$$

Наиболее важной характеристикой записи является поверхностная плотность  $D_s$ , которая позволяет сравнивать различные системы записи

$$D_s = \frac{D_{lin}}{\Delta_{sp}}.$$

Т а б л и ц а 2.1. Сравнительные характеристики стандартов и форматов записи						
Параметры записи	CD	DVD	BlueRay	DASH	DAT	HDD
Минимальная длина пита, мкм	0,83	0,4	0,16	0,66	0,416	0,05
Линейная плотность записи, Кбит/мм	1,69	3,75	8,75	1,514	2,4	30
Поверхностная плотность записи, Мбит/мм кв.	1,13	5,1	21,9	0,054	0,178	61,2
Шаг спиральной дорожки, мкм	1,5	0,74	0,32	-	-	0,49
Длина спиральной дорожки, км	5,74	11,8	21,5	-	-	10,7
Емкость носителя (данные), Гбайт	1,2	5,5	23,3	-	-	40
Емкость носителя (аудио), Гбайт	0,84	4,7	-	-	-	-

В табл.2.1. приведены основные сравнительные характеристики магнитной и оптической записи звука в различных форматах. Как видно, наиболее показательной характеристикой является поверхностная плотность записи. По этому показателю магнитная запись на жестких дисках пока значительно превосходит все другие форматы записи. Причем емкость 40 Гбайт достигается на одной стороне диска диаметром 3,5 дюйма.

*Шум носителя записи.* Шум носителя записи определяется, прежде всего неравномерностью чувствительности и толщины магнитного или оптического покрытий подложки. Сильное влияние оказывает на шум шероховатость подложки и наличие в ней внутренних напряжений. Запись данных на носитель производится всегда в канальном коде, который имеет значительную помехоустойчивость. Благодаря этому коду надежное воспроизведение цифровой записи осуществляется уже при SNR равным 20 дБ.

*Технические характеристики* цифровой аппаратуры записи по звуковому выходу определяются ее цифровым трактом. Они примерно одинаковы для всех устройств записи:

- частотный диапазон 20...20000 Гц с неравномерностью АЧХ равной 0,2 ...0,5 дБ,
- максимальный динамический диапазон 96...120,
- максимальное отношение сигнал/шум 96...130 дБ,
- нелинейные искажения 0,002... 0,007,
- разделение между каналами – более 90 дБ,

- длительность записи от 60 мин до 24 часов.

Надо заметить, что при одинаковых технических характеристиках качество звучания аппаратуры цифровой звукозаписи отличается очень сильно, это замечают и не профессионалы. Улучшение качества звучания уверенно замечается в системах с передискретизацией и использованием операций «Noise Shaping» и «Dithering». Причем, это в равной мере касается трактов с АЦП и ЦАП на основе ИКМ и  $\Sigma\Delta$ -модуляции.

### 3. Линейная импульсно-кодовая модуляция

#### 3.1. Назначение и принцип работы

Аналого-цифровое и цифро-аналоговое преобразования звука в высококачественных цифровых трактах наиболее часто осуществляются на основе линейной импульсно-кодовой модуляции (ИКМ). Она отличается тем, что цифровое преобразование звуковых сигналов (ЗС) малого и высокого уровня производится с одинаковой абсолютной точностью. При этом обеспечивается пока наилучшее качество звучания, но достигается это за счет слишком высокой скорости цифрового потока и дорогостоящей технологии изготовления АЦП и ЦАП.

*Принцип работы ИКМ.* Для преобразования ЗС в цифровой код производится его дискретизация, при которой с частотой  $f_s$  делаются очень короткие по времени выборки, так чтобы за время их длительности ЗС не успевал существенно измениться. Эти выборки запоминаются на время, необходимое для преобразования их в двоичный код. Частота дискретизации, в соответствии с теоремой В.П.Котельникова, должна, по крайней мере вдвое, превышать максимальную частоту спектра звукового сигнала  $F_{\max}$

$$f_s > 2F_{\max}.$$

Чем выше частота дискретизации и больше число разрядов, используемых для кодирования, тем лучше звучание при воспроизведении цифровой записи музыкальных программ.

#### 3.2. Модуляция

*Функциональная схема ИКМ модулятора* схема приведена на рис.3.1. В этой схеме антиэлайзинговый фильтр АФНЧ в соответствии с требованиями теоремы Котельникова В.П. ограничивает спектр входного ЗС частотой Найквиста, равной половине частоты

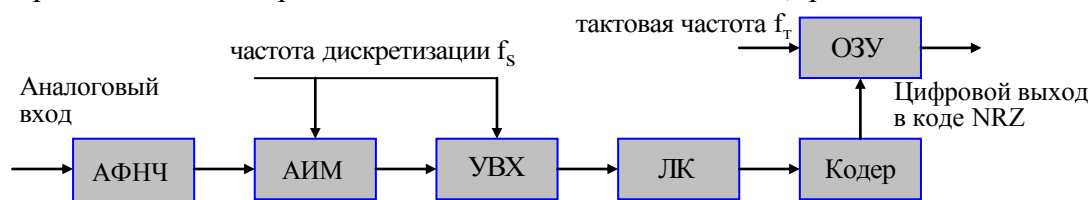


Рис.3.1. Функциональная схема ИКМ модулятора

дискретизации. Дискретизация осуществляется с помощью амплитудно-импульсной модуляции (АИМ). *Линейный квантователь (ЛК) и двоичный кодер* осуществляют аналого-цифровое преобразование АИМ сигнала, при котором формируется последовательность двоичных  $q$  – разрядных кодовых слов с частотой выборки  $f_s$ . При этой операции устанавливается полярность выборки, ее величина и осуществляется отсчет в двоичном коде. *Оперативное запоминающее устройство (ОЗУ)* предназначено для формирования выходного цифрового сигнала в канальном коде NRZ (Non Return to Zero). В этом коде осуществляется передача двоичной информации между элементами цифровой схемотехники.

*Антиэлайзинговая фильтрация* в схеме модулятора предназначена для исключения возможности перекрытия спектров звукового и АИМ сигналов. В спектре такого сигнала около каждой гармоники частоты дискретизации возникают верхняя и нижняя боковые полосы. Огибающие спектров этих полос являются соответственно прямым и зеркальным отражениями огибающей спектра ЗС (рис.3.2.). Антиэлайзинговый ФНЧ подавляет все продукты модуляции выше частоты Найквиста  $f_N$ . В случае отсутствия АФНЧ и когда он

плохо спроектирован из-за перекрытия спектров невозможно восстановить исходный ЗС при ИКМ демодуляции и возникают очень сильные искажения.

Звуки речи и музыкальных инструментов не превышают частоты 15...20 кГц. Особая необходимость подавлять более высокие частоты, возникает при использовании как источника музыкальных программ аналоговых магнитофонов, у которых вместе со звуком

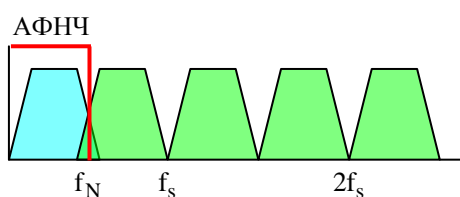


Рис.3.2. Перекрытие АИМ спектров

присутствует сигнал высокочастотного подмагничивания с частотой 50...250 кГц. Частоты выше 20 кГц могут быть также в шуме, синтетической музыке и звуках, издаваемыми животными и птицами. Для надежного подавления спектральных составляющих выше частоты Найквиста антиэлайзинговый ФНЧ должен иметь частоту среза несколько ниже частоты Найквиста и очень сильно (не меньше 90 дБ) подавлять оставшиеся спектра сигнала на этой частоте. Обычно это эллиптические фильтры 7...9 порядков.

Высокая крутизна среза ФНЧ приводит к специфическим искажениям в виде «звона». Это связано с тем, что отклик на импульсное воздействие такого фильтра описывается осциллирующей функцией вида  $\text{sinc}(\omega)$ . Чем выше крутизна среза ФНЧ, тем медленнее происходит затухание звуковых колебаний. Единственным способом борьбы с этими искажениями является повышение частоты выборок. Это позволяет уменьшить крутизну спада антиэлайзингового ФНЧ без уменьшения эффективности подавления спектральных составляющих выше частоты Найквиста.

**Дискретизация** заключается в преобразование аналогового ЗС в дискретную последовательность коротких выборок положительной и отрицательной полярности длительностью  $\tau$ . Такая операция выполняется с помощью балансной АИМ. На один вход

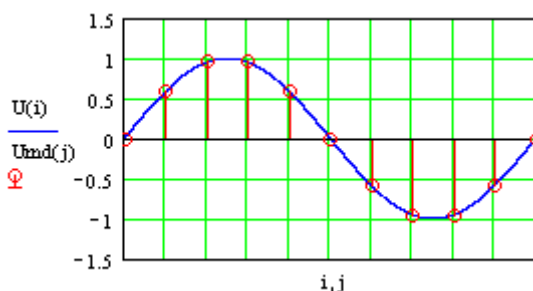


Рис.3.3. АИМ сигнал

модулятора подается ЗС, а на второй - управляющий сигнал, который представляет собой последовательность коротких импульсов с частотой дискретизации. На выходе модулятора формируется АИМ сигнал  $U_d(j)$  (рис.3.3.). Чем меньше длительность выборки, тем более точно выполняется операция дискретизации и меньше нелинейные искажения.

В УВХ, часто включаемой в состав дискретизатора, значения выборок запоминаются на время  $T_s$  и дискретные выборки  $U_d(j)$  преобразуются в непрерывный аналоговый сигнал ступенчатой формы  $U_{dh}(i)$ . Аналоговая память в УВХ реализуется с помощью конденсатора, заряд и разряд которого осуществляется через управляемый ключ. В этом устройстве очень малая постоянная времени заряда и очень большая постоянная времени разряда.

АИМ сигнал является периодическим только при условии, что периоды звукового сигнала  $T_F$  сигнала дискретизации  $T_s$  находятся в целочисленном отношении, то есть когда в  $x$  периодов ЗС укладывается точно  $y$  периодов сигнала дискретизации. Это значит, что частота ЗС  $F$  всегда кратна или субкратна частоте дискретизации  $f_s$ . Коэффициент кратности этих частот определяется отношением простых чисел  $y/x$

$$k = \frac{T_F}{T_s} = \frac{f_s}{F} = \frac{y}{x} > 2,$$

например,  $k = 3/1, 4/1, \dots, 5/2, 7/3, 9/4, \dots, 973/127, \dots$

Таким образом, если  $k$  целое число – частоты  $f_s$  и  $F$  кратны, если дробное – субкратны. Период АИМ сигнала  $T_m$  определяется равенством

$$T_m = \frac{x}{F}$$

Это значит, что, на субкратных частотах нижняя граничная частота спектра АИМ сигнала в  $x$  раз ниже частоты ЗС. В качестве иллюстрации на рис.3.4. приведен АИМ сигнал с периодом, когда в нем укладывается 5 периодов ЗС и 11 периодов сигнала дискретизации, при этом  $k = 11/5$ .

При амплитудной модуляции расчетные формулы для сигналов и ошибок удобно представлять в функции дискретного времени  $i$  тогда аналитическое выражение для ЗС синусоидальной формы принимает вид

$$U(i) = A \cdot \sin(\Delta\varphi \cdot i), \quad \Delta\varphi = \frac{2\pi}{K_f}, \quad A = \frac{A_m}{Q},$$

где  $\Delta\varphi$  - изменение фазы синусоидального сигнала между выборками,  $A$  - нормализованное значение амплитуды ЗС  $A_m$  в квантах,  $Q$  - шаг квантования по уровню,  $i = 0, 1/h, \dots, y$ ,  $h$  - число точек отсчета на интервале между выборками.

Дискретизированный сигнал задается только в моменты выборок, поэтому он рассчитывается в соответствии с равенствами:

$$U_d(j) = A \cdot \sin(\Delta\varphi \cdot j), \quad \text{или} \quad U_d(j) = A \cdot \sin(\Delta\varphi \cdot \text{floor}(i)),$$

где  $j = 0, 1, 2, \dots, y$ ,  $\text{floor}(\cdot)$  - функция, равная целому значению числа в скобках.

Частота выборок имеет стандартные значения: для профессиональной аппаратуры звукозаписи - 48 кГц, для бытовой - 44,1 кГц. В последнее время для повышения качества звучания такой аппаратуры стали использовать более высокие частоты, превышающие стандартные значения в 2 или 4 раза. Так, в формате DVD-Audio предусмотрена возможность осуществлять запись с частотами дискретизации 96 и 192 кГц.

**Линейное квантование** выборок осуществляется с помощью линейной шкалы, разделенной на множество уровней квантования с одинаковым шагом  $Q$ . В процессе квантования производится округление значений выборок до

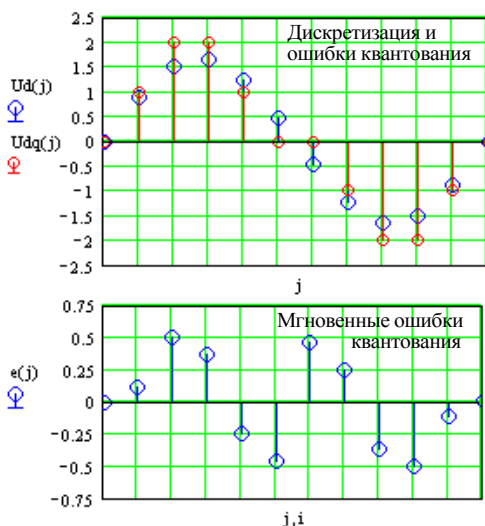


Рис.3.5. Квантование

целого значения шкалы. При этом определяется полярность выборок и число уровней квантования ЗС в десятичной форме  $N_q$ . Процесс квантования последовательности выборок  $U_d(j)$  и возникновение ошибок квантования в моменты выборок поясняется рис.3.5.

Максимальное число уровней квантования шкалы для каждой полярности ЗС  $N_{q\max}$  зависит от числа разрядов в кодовом слове  $q$

$$N_{q\max} = 2^{(q-1)}, \quad q > 1.$$

Чем больше  $q$ , тем на большее число уровней разбивается шкала и тем с большей точностью производится аналого-цифровое преобразование. Если задано максимальное число уровней квантования  $N_{q\max}$ , то необходимое число разрядов в кодовом слове рассчитывается по формуле

$$q = \log_2(N_{q\max}) + 1.$$

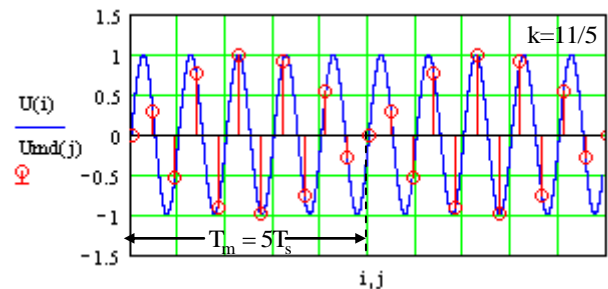


Рис.3.4. АИМ сигнал на субкратной частоте



Шаг квантования  $Q$  шкалы (в единицах напряжения: мВ, мкВ...) определяется максимально допустимой амплитудой входного сигнала  $A_{m\max}$  и максимальным числом уровней шкалы квантования

$$Q = \frac{A_{m\max}}{N_{q\max}} .$$

Входное напряжение квантователя не должно превышать значения  $A_{m\max}$ , в противном случае возникает перегрузка по входу и появляется отсечка (сверху и снизу) в выходном сигнале ИКМ демодулятора и сильные искажения звука.

Важной характеристикой квантователя является его передаточная функция. Возможны 3 варианта ее реализации, отличающиеся нормализованными (в квантах) значениями порогов квантования  $z = 0,5$ ;  $z = 0$  и  $z = 1$ . В первом варианте округление производится к ближайшему значению шкалы квантователя, во втором – к меньшему, а в третьем – к большему значению. Наибольшее распространение получил первый вариант, при котором ошибка квантования не превышает половины шага квантования  $\pm Q/2$ .

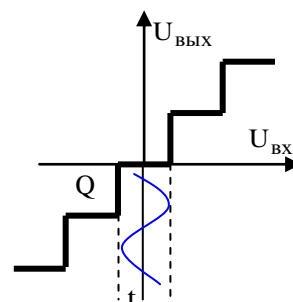


Рис.3.6. Передаточная функция квантователя

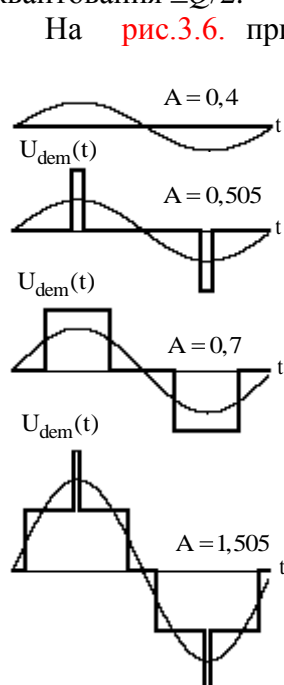


Рис.3.7. Выходной сигнал демодулятора

моменты выборки

На рис.3.6. приведен график передаточной функции квантователя при  $z = 0,5$ . Она имеет вид лестницы, у которой ступени равны шагу квантования  $Q$ . Как видно, пока пиковое значение входного сигнала не превышает 0,5 кванта, выходной сигнал квантователя равен нулю, т.е. имеет место центральная отсечка. Это приводит к нелинейности преобразования и возникновению больших искажений при малых амплитудах ЗС. Пока значение  $A \leq 1,5$  выходной сигнал демодулятора представляет собой последовательность прямоугольных импульсов, у которых с изменением уровня ЗС меняется длительность (рис.3.7). Это обусловлено ошибками квантования, которые соизмеримы с амплитудой входного сигнала. Искажения начинают заметно уменьшаться только при  $A > 2$ .

В ИКМ модуляторе квантуется дискретизированный сигнал в виде последовательности выборок. Функция выходного сигнала квантователя определяется равенством

$$U_{dq}(j) = \text{floor}[A \cdot \sin(\Delta\varphi \cdot j) + z] .$$

Ошибкой квантования называется разность между мгновенными значениями выходного и входного сигналов квантователя в

$$e(j) = \text{floor}[A \cdot \sin(\Delta\varphi \cdot j) + z] - A \cdot \sin(\Delta\varphi \cdot j) .$$

**Двоичный кодер** в схеме на рис.3.1. определяет полярность выборки и осуществляет преобразование положительного десятичного значения уровня квантования в двоичное. Это производится в соответствии с известными правилами с подбором весовых коэффициентов или с помощью таблицы. Например, если  $q = 8$  и  $N_q = 85$ , то это соответствует двоичному числу  $a(x) = 1010101$ . Существуют несколько способов двоичного кодирования, но всегда старший разряд в кодовом слове является знаковым и определяет его полярность.

Для примера в табл.3.1 приведены два наиболее распространенных способа двоичного представления положительных и отрицательных десятичных чисел, известные

под названиями «2's complement format» (двоичное дополнение) и «offset binary» (двоичный сдвиг). В обоих форматах отрицательные десятичные значения преобразуются в двоичные путем инверсии положительной значащей части двоичной кодовой слова того же уровня и прибавления «1» в младшем разряде.

**Оперативное ЗУ** непосредственно относится к двоичному кодеру, являясь его элементом цифровой памяти. Значения кодовых слов в двоичном виде вводятся в ОЗУ в параллельном коде с частотой выборки и с этой же частотой выдаются на параллельном

Таблица 3.1. Двоичные коды

Десятичное число	Двоичный сдвиг	Двоичное дополнение
+5	1101	0101
+4	1100	0100
+3	1011	0011
+2	1010	0010
+1	1001	0001
0	1000	0000
-1	0111	1111
-2	0110	1110
-3	0101	1101
-4	0100	1100
-5	0011	1011

выходе. С последовательного выхода двоичные данные поступают с тактовой частотой  $f_T$  на кодер коррекции ошибок или на вход мультиплексора, в зависимости от схемы построения цифрового тракта.

Вся цифровая схемотехника строится на передачи двоичных данных в виде двух уровней сигнала – высокого и низкого. Первый соответствует логической «1», второй – логическому «0». Это значит, что выходной сигнал ОЗУ представляет собой последовательность импульсов одной полярности с тактовой частотой  $f_T$ , равной

$$f_T = f_s \cdot q.$$

Длительности импульсов и интервалы между ними кратны периоду тактовой частоты. Переход от одного уровня к другому происходит только при изменении последовательности двоичных данных, когда после «0» следует «1» или наоборот. Представление двоичной последовательности в таком виде называется кодом NRZ. Запись цифрового сигнала производится всегда с использованием самосинхронизирующегося канального кода.

### 3.3. Демодуляция

На рис.3.8. приведена функциональная схема ИКМ демодулятора. В ней сигналы с тактовой частотой  $f_T$  и с частотой  $f_s$  выделяется из канального кода и используются для демодуляции. Входной сигнал в коде NRZ подается на ОЗУ демодулятора, которое

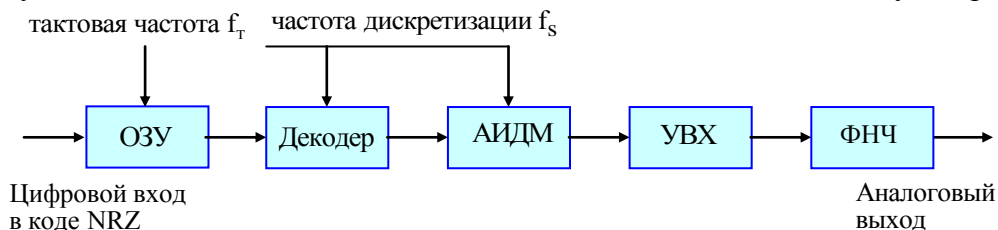


Рис.3.8. Функциональная схема демодулятора

формирует последовательность  $q$ -разрядных кодовых слов в параллельном двоичном коде с частотой  $f_s$ . Двоичный декодер восстанавливает квантованные выборки в виде положительных и отрицательных дискретных значений  $U_{dec}(j)$ . Они имеют точно такой же вид, как и на выходе квантователя модулятора и определяются равенством

$$U_{dec}(j) = \text{floor}[A \cdot \sin(\Delta\varphi \cdot j) + z].$$

УВХ осуществляет аппроксимацию сигнала двоичного декодера с помощью интерполяции нулевого порядка. При такой интерполяции запоминаются значения выборок  $U_{dec}(j)$  на время от одной выборки до другой. В результате на выходе ИКМ демодулятора (до ФНЧ) формируется непрерывный по времени ступенчатый сигнал  $U_{dem}(i)$ . Математическая реализация процесса интерполяции выполняется функцией  $\text{floor}(\cdot)$  в соответствии с равенством

$$U_{dem}(i) = \text{floor}[A \cdot \sin(\Delta\varphi \cdot \text{floor}(i)) + z].$$

*Аппроксимация* предназначена для увеличения уровня спектральных составляющих звукового диапазона и частичного подавления высокочастотных продуктов модуляции еще до ФНЧ и для ее осуществления может использоваться интерполяция высоких порядков. Вместе с сигналом аппроксимируются и ошибки квантования, возникшие в ИКМ модуляторе. Они становятся непрерывной функцией времени в виде последовательности прямоугольных импульсов различной амплитуды и длительности, расчетное соотношение для них принимает вид

$$e_h(i) = U_{dem}(i) - U_{dh}(i).$$

Спектр такой последовательности ограничен частотой Найквиста, поэтому ошибки квантования фильтром нижних частот на выходе демодулятора не ослабляются.

Спектры АИМ сигналов на выходах модулятора и демодулятора существенно отличаются (рис.3.9), так как передаточная функция УВХ демодулятора является частотно-зависимой и определяется равенством

$$K(\eta)_{dem} = |\mu \text{sinc } \pi\mu\eta|, \quad \eta = \frac{f}{f_s}, \quad \mu = \frac{T_{sh}}{T_s}$$

где  $\eta$  - нормализованная частота,  $T_{sh}$  – время удержания УВХ.

На рис.3.9. приведены графики спектра широкополосного ЗС, рассчитанные с использованием этой формулы при  $\mu=1$ . Из них видно, что УВХ вносит достаточно

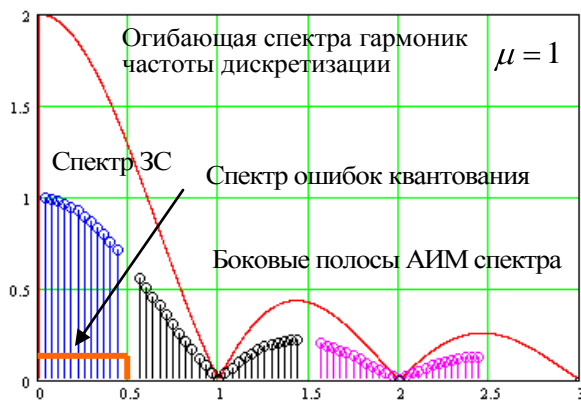


Рис.3.9. Спектр сигнала на выходе УВХ демодулятора

сильные частотные искажения. На частотах выше частоты Найквиста амплитуда побочных продуктов модуляции быстро затухает и это упрощает требования к ФНЧ. В тоже время в звуковом диапазоне возникает спад высоких частот, который на частоте Найквиста достигает минус 3 дБ. Поэтому после ФНЧ включают частотный корректор, называемый аппретурным коррелятором, который компенсирует спад частотной характеристики. К сожалению, этот корректор вносит существенные фазовые искажения, ухудшающие качество звучания. Частотные искажения уменьшаются с уменьшением времени

удержания, но при этом сильно падает уровень ЗС.

*ФНЧ на выходе демодулятора* осуществляет демодуляцию ИКМ сигнала, при которой он подавляет высокочастотные продукты модуляции выше частоты Найквиста и выделяет составляющие спектра в звуковом диапазоне частот. Для реконструкции ЗС в аналоговой форме особой необходимости применения фильтра для ослабления и подавления этих частот нет, так как на слух они не воспринимаются. Однако, необходимо учитывать, что последующие каскады усилителей на транзисторах и акустическая аппаратура из-за неизбежной их нелинейности могут создавать комбинационные искажения попадающие в полосу звуковых частот. Если, например, частота выборок равна 44,1 кГц, а частота ЗС – 20 кГц, то нижняя частота зеркального спектра 24,1 кГц может создавать хорошо слышимую разностную частоту 4,1 кГц. По этой причине подавление выше частоты Найквиста должно быть не меньше 40...60 дБ.

Не очень высокие требования к ФНЧ ИКМ демодулятора обусловлены следующими причинами:

- звуковые акустические системы имеют малую чувствительность на высоких частотах,

- музыкальный сигнал с множеством обертонов создает значительную спектральную маскировку,
- переходные искажения порождают временное маскирование.

В демодуляторе на вход ФНЧ с частотой дискретизации подаются квантованные выборки ЗС сигнала в виде последовательности прямоугольных импульсов различной амплитуды и полярности. Выходной отклик идеального ФНЧ с частотой среза равной частоте Найквиста при каждом импульсном воздействии описывается функцией  $\sin x/x$ , у которой пиковое значение которой равно амплитуде входного импульса (рис.3.10), а затухающий колебательный процесс происходит с периодом дискретизации.

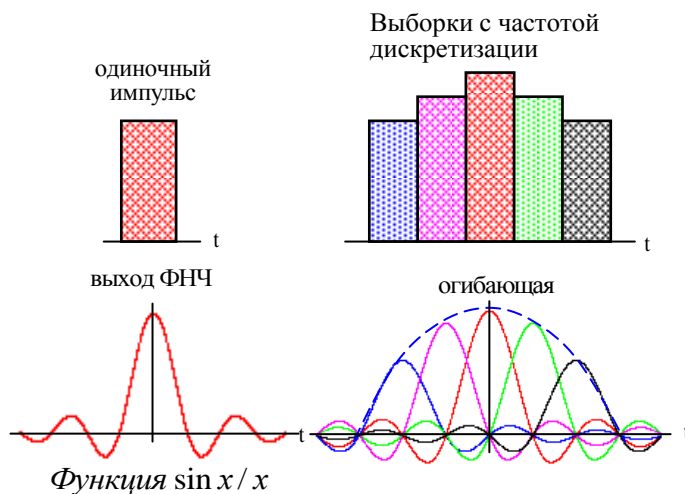


Рис. 3.10. Реконструкция ЗС

В случае непрерывной последовательности выборок происходит перекрытие откликов от отдельных выборок. При этом пиковое значение каждого отклика равно амплитуде входного импульса, и оно не зависит от соседних откликов. Значение восстановленного ЗС в каждой точке между выборками определяется суммой огромного числа откликов, возникших как раньше, так и позже рассматриваемого момента времени. Особо большую роль при этом играют медленно затухающие колебания.

Реконструкция ЗС осуществляется ФНЧ по отдельным выборкам путем их интерполяции. Для восстановления звукового сигнала с широким спектром от 20 Гц до 20 кГц необходимо накопление примерно 2000 выборок, а более точно оно определяется известным равенством.

$$N_{\text{выб}} = F_{\text{max}} / F_{\text{min}} = 4000 \text{ выборок.}$$

Это очень жесткий критерий. Считается, что для музыкальных программ при частоте дискретизации 44100 Гц вполне достаточно 2200 выборок. Для того чтобы обеспечивалось такое число интерполируемых выборок *время задержки ФНЧ* должно быть больше разности периодов граничных спектральных составляющих широкополосного ЗС со спектром прямоугольной формы. В противном случае возникают нелинейные интермодуляционные искажения, которые на слух проявляются в виде диссонансов, «грохота» и «хвостов». Это значит, что для частотного диапазона 20...20000 Гц время задержки ФНЧ должно быть 50 мс. Для сигнала с полосой от 10 Гц до 20 кГц, время задержки фильтра должна быть уже 100 мс и т.д. Такие фильтры создать очень сложно, поэтому надо ограничивать нижний диапазон частот.

Чем большее число выборок используется в реконструкции ЗС, тем ближе восстановленный сигнал к исходной форме. При этом восстанавливаются без ошибки не только выборки, но все промежуточные значения между ними. Число точек, по которым осуществляется восстановление звукового сигнала, определяется крутизной среза ФНЧ. Чем ближе реальный ФНЧ к идеальному фильтру, с крутизной среза близкой к бесконечности, тем больше время задержки сигнала в этом фильтре и тем больше выборок связываются в этом фильтре друг с другом. Переходные процессы в фильтре дополняют восстанавливаемый сигнал между значениями соседних выборок.

На практике время задержки ФНЧ, рассчитанного на пропускание спектра от 20 Гц до 20 кГц составляет всего 5...10 мс. Это значит, что при подаче цифрового сигнала на вход ФНЧ ЗС на выходе он появится через указанное время. С такой маленькой задержкой при формировании ЗС участвуют только 220...440 выборок и поэтому при

реконструкции широкополосных музыкальных сигналах неизбежно возникают нелинейные искажения.

В современной аппаратуре для реконструкции ЗС в аналоговой форме ФНЧ с катушками индуктивности и конденсаторами, а также активные транзисторные РС-фильтры практически не применяются. Для этих целей более целесообразно использовать цифровые интерполяторы 15...16 порядка или производить вычисление выходного сигнала по формуле из теоремы В.П.Котельникова

$$U_{фнч}(i) = \sum_{j=0}^{j_{\max}} U_{дем}(j) \cdot \text{sinc}[(i-j) \cdot \pi]$$

где  $j_{\max}$  - число выборок, по которым производится интерполяция. По этой формуле правильно рассчитываются значения ЗС не только в моменты выборок, но и во всех точках между выборками. В обоих случаях требуется высокоскоростной процессор и достаточно большая память.

### 3.1. Назначение и принцип работы

Аналого-цифровое и цифро-аналоговое преобразования звука в высококачественных цифровых трактах наиболее часто осуществляются на основе линейной импульсно-кодовой модуляции (ИКМ). Она отличается тем, что цифровое преобразование звуковых сигналов (ЗС) малого и высокого уровня производится с одинаковой абсолютной точностью. При этом обеспечивается пока наилучшее качество звучания, но достигается это за счет слишком высокой скорости цифрового потока и дорогостоящей технологии изготовления АЦП и ЦАП.

*Принцип работы ИКМ.* Для преобразования ЗС в цифровой код производится его дискретизация, при которой с частотой  $f_s$  делаются очень короткие по времени выборки, так чтобы за время их длительности ЗС не успевал существенно измениться. Эти выборки запоминаются на время, необходимое для преобразования их в двоичный код. Частота дискретизации, в соответствии с теоремой В.П.Котельникова, должна, по крайней мере вдвое, превышать максимальную частоту спектра звукового сигнала  $F_{\max}$

$$f_s > 2F_{\max} \cdot$$

Чем выше частота дискретизации и больше число разрядов, используемых для кодирования, тем лучше звучание при воспроизведении цифровой записи музыкальных программ.

### 3.2. Модуляция

*Функциональная схема ИКМ модулятора* схема приведена на [рис.3.1](#). В этой схеме *антиэлайзинговый фильтр* АФНЧ в соответствии с требованиями теоремы Котельникова В.П. ограничивает спектр входного ЗС частотой Найквиста, равной половине частоты дискретизации. Дискретизация осуществляется с помощью амплитудно-импульсной модуляции (АИМ). *Линейный квантователь* (ЛК) и *двоичный кодер* осуществляют аналого-цифровое преобразование АИМ сигнала, при котором формируется последовательность двоичных  $q$  – разрядных кодовых слов с частотой выборки  $f_s$ . При этой операции устанавливается полярность выборки, ее величина и осуществляется отсчет в двоичном коде. *Оперативное запоминающее устройство* (ОЗУ) предназначено для формирования выходного цифрового сигнала в канальном коде NRZ (Non Return to Zero). В этом коде осуществляется передача двоичной информации между элементами цифровой схемотехники.

*Антиэлайзинговая фильтрация* в схеме модулятора предназначена для исключения возможности перекрытия спектров звукового и АИМ сигналов. В спектре такого сигнала

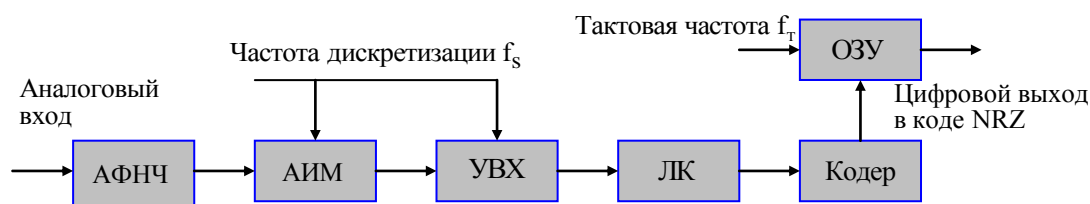


Рис.3.1. Функциональная схема ИКМ модулятора

около каждой гармонике частоты дискретизации возникают верхняя и нижняя боковые полосы. Огибающие спектров этих полос являются соответственно прямым и зеркальным отражениями огибающей спектра ЗС (рис.3.2.). Антиэлайзинговый ФНЧ подавляет все продукты модуляции выше частоты Найквиста  $f_N$ . В случае отсутствия АФНЧ и когда он плохо спроектирован из-за перекрытия спектров невозможно восстановить исходный ЗС при ИКМ демодуляции и возникают очень сильные искажения.

Звуки речи и музыкальных инструментов не превышают частоты 15...20 кГц. Особая необходимость подавлять более высокие частоты, возникает при использовании как источника музыкальных программ аналоговых магнитофонов, у которых вместе со звуком присутствует сигнал высокочастотного подмагничивания с частотой 50...250 кГц.

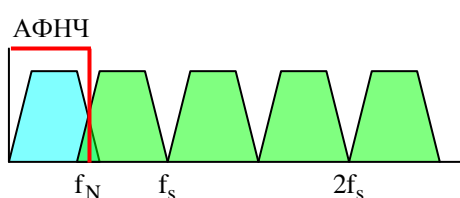


Рис.3.2. Перекрытие АИМ спектров

Частоты выше 20 кГц могут быть также в шуме, синтетической музыке и звуках, издаваемыми животными и птицами.

Для надежного подавления спектральных составляющих выше частоты Найквиста антиэлайзинговый ФНЧ должен иметь частоту среза несколько ниже частоты Найквиста и очень сильно (не меньше 90 дБ) подавлять оставшиеся спектра сигнала на этой частоте. Обычно это эллиптические фильтры 7...9 порядков.

Высокая крутизна среза ФНЧ приводит к специфическим искажениям в виде «звона». Это связано с тем, что отклик на импульсное воздействие такого фильтра описывается осциллирующей функцией вида  $\text{sinc}(\omega)$ . Чем выше крутизна среза ФНЧ, тем медленнее происходит затухание звуковых колебаний. Единственным способом борьбы с этими искажениями является повышение частоты выборок, Это позволяет уменьшить крутизну спада антиэлайзингового ФНЧ без уменьшения эффективности подавления спектральных составляющих выше частоты Найквиста.

*Дискретизация* заключается в преобразование *аналогового ЗС* в дискретную последовательность коротких выборок положительной и отрицательной полярности

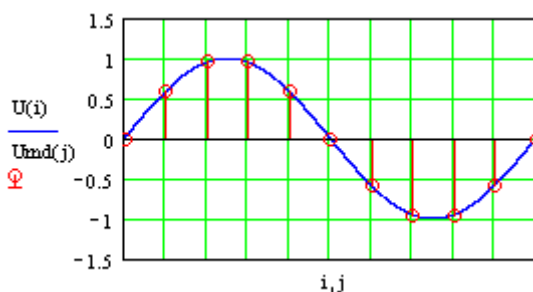


Рис.3.3. АИМ сигнал

длительностью  $\tau$ . Такая операция выполняется с помощью *балансной АИМ*. На один вход модулятора подается ЗС, а на второй - управляющий сигнал, который представляет собой последовательность коротких импульсов с частотой дискретизации. На выходе модулятора формируется АИМ сигнал  $U_d(j)$  (рис.3.3.). Чем меньше длительность выборки, тем более точно выполняется операция дискретизации и меньше нелинейные искажения.

В УВХ, часто включаемой в состав дискретизатора, значения выборок запоминаются на время  $T_s$  и дискретные выборки  $U_d(j)$  преобразуются в непрерывный аналоговый сигнал ступенчатой формы  $U_{dh}(i)$ . Аналоговая память в УВХ реализуется с помощью конденсатора, заряд и разряд которого осуществляется через управляемый ключ. В этом

устройстве очень малая постоянная времени заряда и очень большая постоянная времени разряда.

АИМ сигнал является периодическим только при условии, что периоды звукового сигнала  $T_F$  сигнала дискретизации  $T_s$  находятся в целочисленном отношении, то есть когда в  $x$  периодов ЗС укладывается точно  $y$  периодов сигнала дискретизации. Это значит, что частота ЗС  $F$  всегда кратна или субкратна частоте дискретизации  $f_s$ . Коэффициент кратности этих частот определяется отношением простых чисел  $y/x$

$$k = \frac{T_F}{T_s} = \frac{f_s}{F} = \frac{y}{x} > 2,$$

например,  $k = 3/1, 4/1, \dots, 5/2, 7/3, 9/4, \dots, 973/127, \dots$

Таким образом, если  $k$  целое число – частоты  $f_s$  и  $F$  кратны, если дробное – субкратны. Период АИМ сигнала  $T_m$  определяется равенством

$$T_m = \frac{x}{F}.$$

Это значит, что, на *субкратных частотах нижняя граничная частота спектра АИМ сигнала в  $x$  раз ниже частоты ЗС*. В качестве иллюстрации на рис.3.4. приведен АИМ сигнал с периодом, когда в нем укладывается 5 периодов ЗС и 11 периодов сигнала дискретизации, при этом  $k = 11/5$ .

При амплитудной модуляции расчетные формулы для сигналов и ошибок удобно представлять в функции дискретного времени  $i$  тогда аналитическое выражение для ЗС синусоидальной формы принимает вид

$$U(i) = A \cdot \sin(\Delta\varphi \cdot i), \quad \Delta\varphi = \frac{2\pi}{K_f}, \quad A = \frac{A_m}{Q},$$

где  $\Delta\varphi$  - изменение фазы синусоидального сигнала между выборками,  $A$  - нормализованное значение амплитуды ЗС  $A_m$  в квантах,

$Q$  - шаг квантования по уровню,  $i = 0, 1/h, \dots, y$ ,  $h$  - число точек отсчета на интервале между выборками.

Дискретизированный сигнал задается только в моменты выборок, поэтому он рассчитывается в соответствии с равенствами:

$$U_d(j) = A \cdot \sin(\Delta\varphi \cdot j), \quad \text{или}$$

$$U_d(j) = A \cdot \sin(\Delta\varphi \cdot \text{floor}(i)),$$

где  $j = 0, 1, 2, \dots, y$ ,  $\text{floor}(\cdot)$  - функция, равная целому значению числа в скобках.

Частота выборок имеет стандартные значения: для профессиональной аппаратуры звукозаписи – 48 кГц, для бытовой – 44,1 кГц. В последнее время для повышения качества звучания такой аппаратуры стали использовать более высокие частоты, превышающие стандартные значения в 2 или 4 раза. Так, в формате DVD-Audio предусмотрена возможность осуществлять запись с частотами дискретизации 96 и 192 кГц.

*Линейное квантование* выборок осуществляется с помощью линейной шкалы, разделенной на множество уровней квантования с одинаковым шагом  $Q$ . В процессе квантования производится округление значений выборок до целого значения шкалы. При этом определяется полярность выборок и число уровней квантования ЗС в

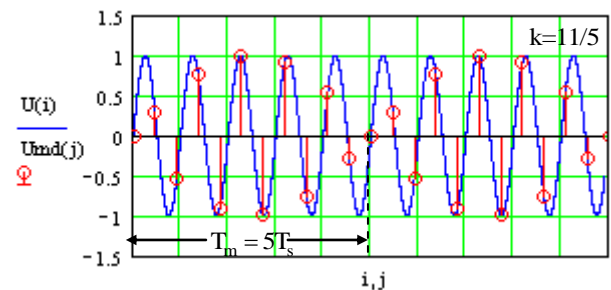


Рис.3.4. АИМ сигнал на субкратной частоте

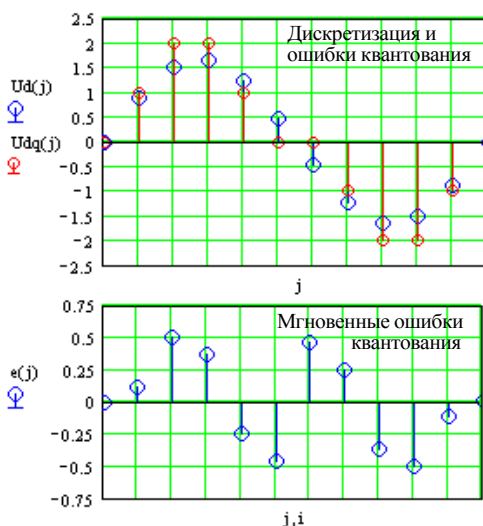


Рис.3.5. Квантование

десятичной форме  $N_q$ . Процесс квантования последовательности выборок  $U_d(j)$  и возникновение ошибок квантования в моменты выборок поясняется [рис.3.5](#).

Максимальное число уровней квантования шкалы для каждой полярности ЗС  $N_{q\max}$  зависит от числа разрядов в кодовом слове  $q$

$$N_{q\max} = 2^{(q-1)}, \quad q > 1.$$

Чем больше  $q$ , тем на большее число уровней разбивается шкала и тем с большей точностью производится аналого-цифровое преобразование. Если задано максимальное число уровней квантования  $N_{q\max}$ , то необходимое число разрядов в кодовом слове рассчитывается по формуле

$$q = \log_2(N_{q\max}) + 1.$$

Шаг квантования  $Q$  шкалы (в единицах напряжения: мВ, мкВ...) определяется максимально допустимой амплитудой входного сигнала  $A_{m\max}$  и максимальным числом уровней шкалы квантования

$$Q = \frac{A_{m\max}}{N_{q\max}}.$$

Входное напряжение квантователя не должно превышать значения  $A_{m\max}$ , в противном случае возникает перегрузка по входу и появляется отсечка (сверху и снизу) в выходном сигнале ИКМ демодулятора и сильные искажения звука.

Важной характеристикой квантователя является его передаточная функция. Возможны 3 варианта ее реализации, отличающиеся нормализованными (в квантах) значениями порогов квантования  $z = 0,5$ ;  $z = 0$  и  $z = 1$ . В первом варианте

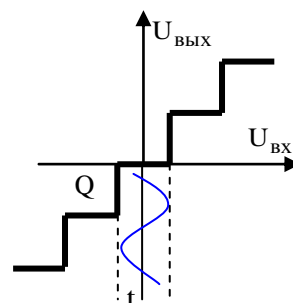


Рис.3.6. Передаточная функция квантователя

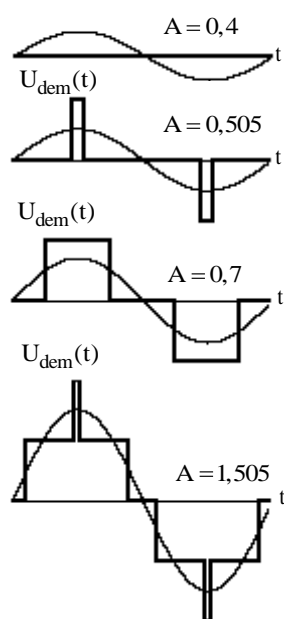


Рис.3.7. Выходной сигнал демодулятора

округление производится к ближайшему значению шкалы квантователя, во втором – к меньшему, а в третьем – к большему значению. Наибольшее распространение получил первый вариант, при котором ошибка квантования не превышает половины шага квантования  $\pm Q/2$ .

На [рис.3.6](#) приведен график передаточной функции квантователя при  $z = 0,5$ . Она имеет вид лестницы, у которой ступени равны шагу квантования  $Q$ . Как видно, пока пиковое значение входного сигнала не превышает 0,5 кванта, выходной сигнал квантователя равен нулю, т.е. имеет место центральная отсечка. Это приводит к нелинейности преобразования и возникновению больших искажений при малых амплитудах ЗС. Пока значение  $A \leq 1,5$  выходной сигнал демодулятора представляет собой последовательность прямоугольных импульсов, у которых с изменением уровня ЗС меняется длительность ([рис.3.7](#)). Это обусловлено ошибками квантования, которые соизмеримы с амплитудой входного сигнала. Искажения начинают заметно уменьшаться только при  $A > 2$ .

В ИКМ модуляторе квантуется дискретизированный сигнал в виде последовательности выборок. Функция выходного сигнала квантователя определяется равенством

$$U_{dq}(j) = \text{floor}[A \cdot \sin(\Delta\phi \cdot j) + z].$$

Ошибкой квантования называется разность между мгновенными значениями выходного и входного сигналов квантователя в моменты выборок



$$e(j) = \text{floor}[A \cdot \sin(\Delta\phi \cdot j) + z] - A \cdot \sin(\Delta\phi \cdot j) .$$

**Двоичный кодер** в схеме на **рис.3.1.** определяет полярность выборки и осуществляет преобразование положительного десятичного значения уровня квантования в двоичное. Это производится в соответствии с известными правилами с подбором весовых коэффициентов или с помощью таблицы. Например, если  $q = 8$  и  $N_q = 85$ , то это соответствует двоичному числу  $a(x) = 1010101$ . Существуют несколько способов двоичного кодирования, но всегда старший разряд в кодовом слове является знаковым и определяет его полярность.

Для примера в **табл.3.1** приведены два наиболее распространенных способа двоичного представления положительных и отрицательных десятичных чисел, известные под названиями «2's complement format» (двоичное дополнение) и «offset binary» (двоичный сдвиг). В обоих форматах отрицательные десятичные значения преобразуются в двоичные путем инверсии положительной значащей части двоичной кодового слова того же уровня и прибавления «1» в младшем разряде.

**Оперативное ЗУ** непосредственно относится к двоичному кодеру, являясь его элементом цифровой памяти. Значения кодовых слов в двоичном виде вводятся в ОЗУ в параллельном коде с частотой выборок и с этой же частотой выдаются на параллельном выходе.

**Таблица 3.1. Двоичные коды**

Десятичное число	Двоичный сдвиг	Двоичное дополнение
+5	1101	0101
+4	1100	0100
+3	1011	0011
+2	1010	0010
+1	1001	0001
0	1000	0000
-1	0111	1111
-2	0110	1110
-3	0101	1101
-4	0100	1100
-5	0011	1011

С последовательного выхода двоичные данные поступают с тактовой частотой  $f_T$  на кодер коррекции ошибок или на вход мультиплексора, в зависимости от схемы построения цифрового тракта.

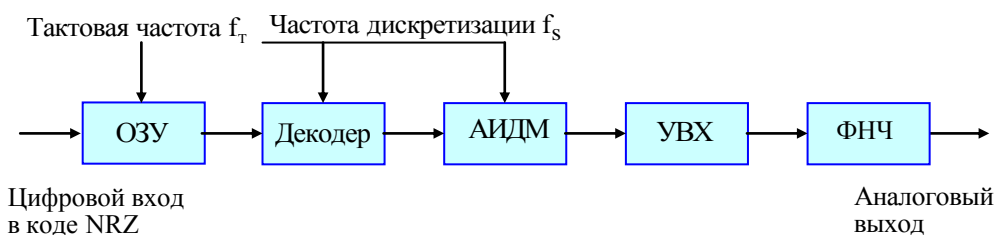
Вся цифровая схемотехника строится на передачи двоичных данных в виде двух уровней сигнала – высокого и низкого. Первый соответствует логической «1», второй – логическому «0». Это значит, что выходной сигнал ОЗУ представляет собой последовательность импульсов одной полярности с тактовой частотой  $f_T$ , равной

$$f_T = f_s \cdot q .$$

Длительности импульсов и интервалы между ними кратны периоду тактовой частоты. Переход от одного уровня к другому происходит только при изменении последовательности двоичных данных, когда после «0» следует «1» или наоборот. Представление двоичной последовательности в таком виде называется кодом NRZ. Запись цифрового сигнала производится всегда с использованием самосинхронизирующегося канального кода.

### 3.3. Демодуляция

На **рис.3.8.** приведена функциональная схема ИКМ демодулятора. В ней сигналы с



**Рис.3.8.** Функциональная схема демодулятора

тактовой частотой  $f_T$  и с частотой  $f_s$  выделяется из канального кода и используются

для демодуляции. Входной сигнал в коде NRZ подается на ОЗУ демодулятора, которое формирует последовательность  $q$ -разрядных кодовых слов в параллельном двоичном коде с частотой  $f_s$ . Двоичный декодер восстанавливает квантованные выборки в виде положительных и отрицательных дискретных значений  $U_{dec}(j)$ , Они имеют точно такой же вид, как и на выходе квантователя модулятора и определяются равенством

$$U_{dec}(j) = \text{floor}[A \cdot \sin(\Delta\varphi \cdot j) + z].$$

УВХ осуществляет *аппроксимацию* сигнала двоичного декодера с помощью интерполяции нулевого порядка. При такой интерполяции запоминаются значения выборок  $U_{dec}(j)$  на время от одной выборки до другой. В результате на выходе ИКМ демодулятора (до ФНЧ) формируется непрерывный по времени *ступенчатый сигнал*  $U_{dem}(i)$ . Математически реализация процесса интерполяции выполняется функцией  $\text{floor}(\cdot)$  в соответствии с равенством

$$U_{dem}(i) = \text{floor}[A \cdot \sin(\Delta\varphi \cdot \text{floor}(i)) + z].$$

*Аппроксимация* предназначена для увеличения уровня спектральных составляющих звукового диапазона и частичного подавления высокочастотных продуктов модуляции еще до ФНЧ и для ее осуществления может использоваться интерполяция высоких порядков. Вместе с сигналом аппроксимируются и ошибки квантования, возникшие в ИКМ модуляторе. Они становятся непрерывной функцией времени в виде последовательности прямоугольных импульсов различной амплитуды и длительности, расчетное соотношение для них принимает вид

$$e_h(i) = U_{dem}(i) - U_{dh}(i).$$

Спектр такой последовательности ограничен частотой Найквиста, поэтому ошибки квантования фильтром нижних частот на выходе демодулятора не ослабляются.

Спектры АИМ сигналов на выходах модулятора и демодулятора существенно отличаются (рис.3.9.), так как передаточная функция УВХ демодулятора является частотно-зависимой и определяется равенством

$$K(\eta)_{dem} = |\mu \text{sinc } \pi\mu\eta|, \quad \eta = \frac{f}{f_s}, \quad \mu = \frac{T_{sh}}{T_s}$$

где  $\eta$  - нормализованная частота,  $T_{sh}$  - время удержания УВХ.

На рис.3.9. приведены графики спектра широкополосного ЗС, рассчитанные с использованием этой формулы при  $\mu=1$ . Из них видно, что УВХ вносит достаточно

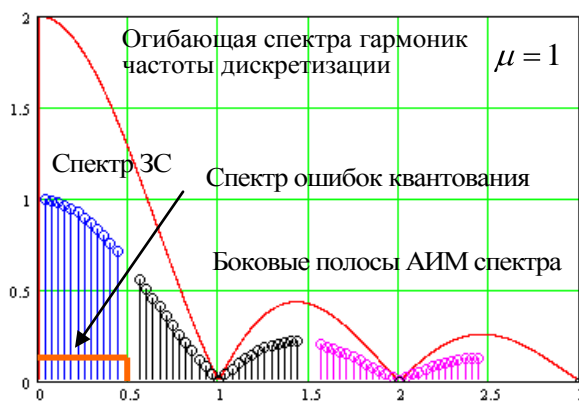


Рис.3.9. Спектр сигнала на выходе УВХ демодулятора

сильные частотные искажения. На частотах выше частоты Найквиста амплитуда побочных продуктов модуляции быстро затухает и это упрощает требования к ФНЧ. В тоже время в звуковом диапазоне возникает спад высоких частот, который на частоте Найквиста достигает минус 3 дБ. Поэтому после ФНЧ включают частотный корректор, называемый аппретурным коррелятором, который компенсирует спад частотной характеристики. К сожалению, этот корректор вносит существенные фазовые искажения, ухудшающие качество звучания. Частотные искажения уменьшаются с уменьшением времени

удержания, но при этом сильно падает уровень ЗС.

*ФНЧ на выходе демодулятора* осуществляет демодуляцию ИКМ сигнала, при которой он подавляет высокочастотные продукты модуляции выше частоты Найквиста и выделяет составляющие спектра в звуковом диапазоне частот. Для реконструкции ЗС в

аналоговой форме особой необходимости применения фильтра для ослабления и подавления этих частот нет, так как на слух они не воспринимаются. Однако, необходимо учитывать, что последующие каскады усилителей на транзисторах и акустическая аппаратура из-за неизбежной их нелинейности могут создавать комбинационные искажения попадающие в полосу звуковых частот. Если, например, частота выборок равна 44,1 кГц, а частота ЗС – 20 кГц, то нижняя частота зеркального спектра 24,1 кГц может создавать хорошо слышимую разностную частоту 4,1 кГц. По этой причине подавление выше частоты Найквиста должно быть не меньше 40...60 дБ.

Не очень высокие требования к ФНЧ ИКМ демодулятора обусловлены следующими причинами:

- звуковые акустические системы имеют малую чувствительность на высоких частотах,
- музыкальный сигнал с множеством обертонов создает значительную спектральную маскировку,
- переходные искажения порождают временное маскирование.

В демодуляторе на вход ФНЧ с частотой дискретизации подаются квантованные выборки ЗС сигнала в виде последовательности прямоугольных импульсов различной амплитуды и полярности. Выходной отклик идеального ФНЧ с частотой среза равной частоте Найквиста при каждом импульсном воздействии описывается функцией  $\sin x/x$ , у которой пиковое значение которой равно амплитуде входного импульса (рис.3.10), а затухающий колебательный процесс происходит с периодом дискретизации.

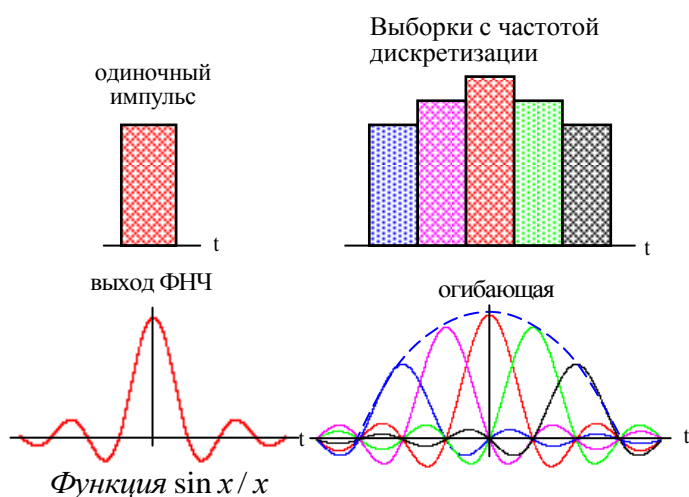


Рис. 3.10. Реконструкция ЗС

Реконструкция ЗС осуществляется ФНЧ по отдельным выборкам путем их интерполяции. Для восстановления звукового сигнала с широким спектром от 20 Гц до 20 кГц необходимо накопление примерно 2000 выборок, а более точно оно определяется известным равенством.

$$N_{\text{выб}} = F_{\text{max}} / F_{\text{min}} = 4000 \text{ выборок.}$$

Это очень жесткий критерий. Считается, что для музыкальных программ при частоте дискретизации 44100 Гц вполне достаточно 2200 выборок. Для того чтобы обеспечивалось такое число интерполируемых выборок *время задержки ФНЧ* должно быть больше разности периодов граничных спектральных составляющих широкополосного ЗС со спектром прямоугольной формы. В противном случае возникают нелинейные интермодуляционных искажения, которые на слух проявляются в виде диссонансов, «грохота» и «хвостов». Это значит, что для частотного диапазона 20...20000 Гц время задержки ФНЧ должно быть 50 мс. Для сигнала с полосой от 10 Гц до 20 кГц, время задержки фильтра должна быть уже 100 мс и т.д. Такие фильтры создать очень сложно, поэтому надо ограничивать нижний диапазон частот.

Чем большее число выборок используется в реконструкции ЗС, тем ближе восстановленный сигнал к исходной форме. При этом восстанавливаются без ошибки не только выборки, но все промежуточные значения между ними. Число точек, по которым осуществляется восстановление звукового сигнала, определяется крутизной среза ФНЧ. Чем ближе реальный ФНЧ к идеальному фильтру, с крутизной среза близкой к бесконечности, тем больше время задержки сигнала в этом фильтре и тем больше выборок связываются в этом фильтре друг с другом. Переходные процессы в фильтре дополняют восстанавливаемый сигнал между значениями соседних выборок.

На практике время задержки ФНЧ, рассчитанного на пропускание спектра от 20 Гц до 20 кГц составляет всего 5...10 мс. Это значит, что при подаче цифрового сигнала на вход ФНЧ ЗС на выходе он появится через указанное время. С такой маленькой задержкой при формировании ЗС участвуют только 220...440 выборок и поэтому при реконструкции широкополосных музыкальных сигналах неизбежно возникают нелинейные искажения.

В современной аппаратуре для реконструкции ЗС в аналоговой форме ФНЧ с катушками индуктивности и конденсаторами, а также активные транзисторные РС-фильтры практически не применяются. Для этих целей более целесообразно использовать цифровые интерполяторы 15...16 порядка или производить вычисление выходного сигнала по формуле из теоремы В.П.Котельникова

$$U_{фнч}(i) = \sum_{j=0}^{j_{\max}} U_{dem}(j) \cdot \text{sinc}[(i-j) \cdot \pi]$$

где  $j_{\max}$  - число выборок, по которым производится интерполяция. По этой формуле правильно рассчитываются значения ЗС не только в моменты выборок, но и во всех точках между выборками. В обоих случаях требуется высокоскоростной процессор и достаточно большая память.

## 4. Спектр и звучание ошибок квантования

### 4.1. Спектр ошибок округления

Для расчета спектра ошибок квантования обычно применяется быстрое преобразование Фурье квантованного сигнала на выходе интерполятора ИКМ демодулятора. Когда основным интересом является исследование звучания ошибок квантования, более важным становится сам процесс формирования этого спектра.

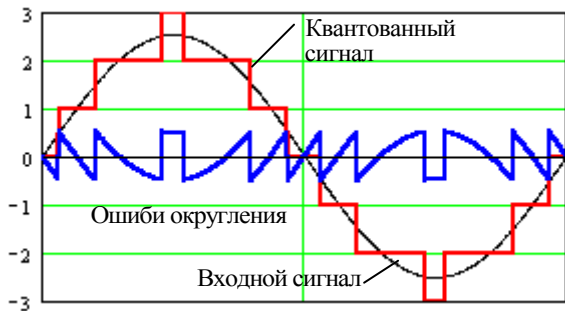


Рис. 4.1. Квантование синусоидального сигнала. Ошибки округления.

Поэтому для расчета спектра воспользуемся особым математическим приемом, который заключается в изменении порядка выполнения математических операций в ИКМ модуляторе. Для этого в схеме модулятора квантователь переносится на место перед амплитудно-импульсным модулятором.

При таком подходе на входе квантователя действует синусоидальный входной сигнал. Его выходной сигнал  $U_r(i)$  представляет собой ступенчатую функцию времени

(рис. 4.1.). Высота всех ступенек равна шагу квантования и не зависит от скорости изменения ЗС, а вот их длина от этой скорости зависит. Чем скорость изменения ЗС выше, тем меньше ширина ступенек. Аналитически квантованный сигнал записывается в виде:

$$U_r(i) = \text{floor}[A \cdot \sin(\Delta\varphi \cdot i) + z],$$

Разность между входным и выходным сигналами квантователя формально является ошибкой квантования, так ее и называют в большинстве публикаций по аудиотехнике. Однако, эта ошибка является результатом выполнения только одной операции – квантования, поэтому ее спектр не ограничен сверху частотой Найквиста. Интересующая нас *ошибка квантования* всегда является результатом выполнения двух операций: *квантования и дискретизации*. Чтобы не создавалась путаница в названиях, будем называть ошибку, возникающую *только при квантовании* синусоидального сигнала *ошибкой округления* и обозначать ее как  $e_r(i)$ . Это более соответствует ее физической природе. Ошибка округления рассчитывается по формуле

$$e_r(i) = U_r(i) - U(i)$$

Из графиков на рис. 4.2. видно, что ошибки округления являются периодическими непрерывными функциями времени. Форма сигнала ошибки в пределах периода имеет вид детерминированной последовательности пилообразных импульсов с размахом равным шагу квантования, независимо от амплитуды ЗС  $A$ . Длительность, форма и частота повторения импульсов меняются в пределах периода, так как они зависят от крутизны синусоидального входного сигнала в функции времени. Период повторения ошибок округления быстро уменьшается с увеличением  $A$ . Из приведенных графиков следует, что спектр гармоник ошибок квантования ограничен номером гармоники, у которой амплитуда максимальная. Эта гармоника грубо определяется числом импульсов ошибок

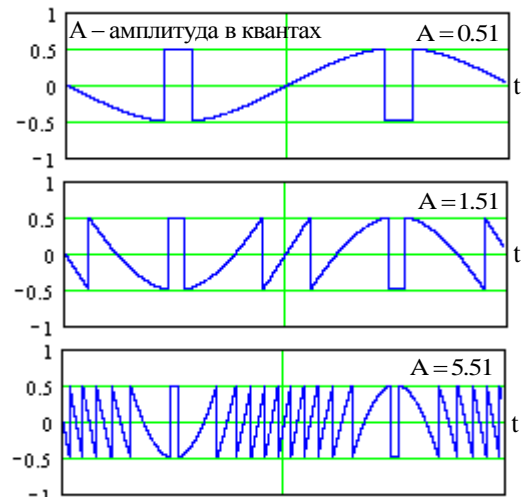


Рис. 4.2. Функции ошибок округления

Из приведенных графиков следует, что спектр гармоник ошибок квантования ограничен номером гармоники, у которой амплитуда максимальная. Эта гармоника грубо определяется числом импульсов ошибок

квантования за период ЗС. Чем больше значение  $A$ , тем выше номер этой гармоники  $\nu(A_{\nu_{\max}})$ . При  $A > 2$  этот номер рассчитывается по приближенной формуле

$$\nu(A_{\nu_{\max}}) \approx 2\pi A$$

Суммарная мощность ошибок округления  $P_r$  зависит только от шага квантования  $Q$

$$P_r = Q^2 / 12.$$

Функция ошибок округления периодическая и симметричная относительно оси времени, поэтому спектр ошибок округления дискретный в виде *нечетных гармоник сигнала ошибок округления*. Это позволяет производить расчет спектра гармоник с помощью БПФ по формуле

$$S(\nu) = 2 \cdot \text{fft}(e_r(i)) \cdot m^{-0.5}$$

где  $m = 2^x \cdot n$ ,  $n$  - число рассчитываемых гармоник спектра, представляемое числом  $2^y$ ,  $m$  - число точек отсчета на период ЗС  $i = 0, 1 \dots m-1$ . Чем больше  $m$ , тем большее число гармоник может быть вычислено. Чем больше значение переменной  $x$ , тем выше точность вычисления гармоник. Для перехода к частотному спектру достаточно изменить масштаб по оси абсцисс, заменив номера гармоник  $\nu$  частотой  $F(\nu) = \nu \cdot F$ .

На графиках (рис.4.3.) по оси ординат определяется амплитуда гармоник ошибок округления в квантах, а по оси абсцисс - номера гармоник. Из приведенных графиков видно, что в спектре ошибок округления имеется составляющая первой гармоники.

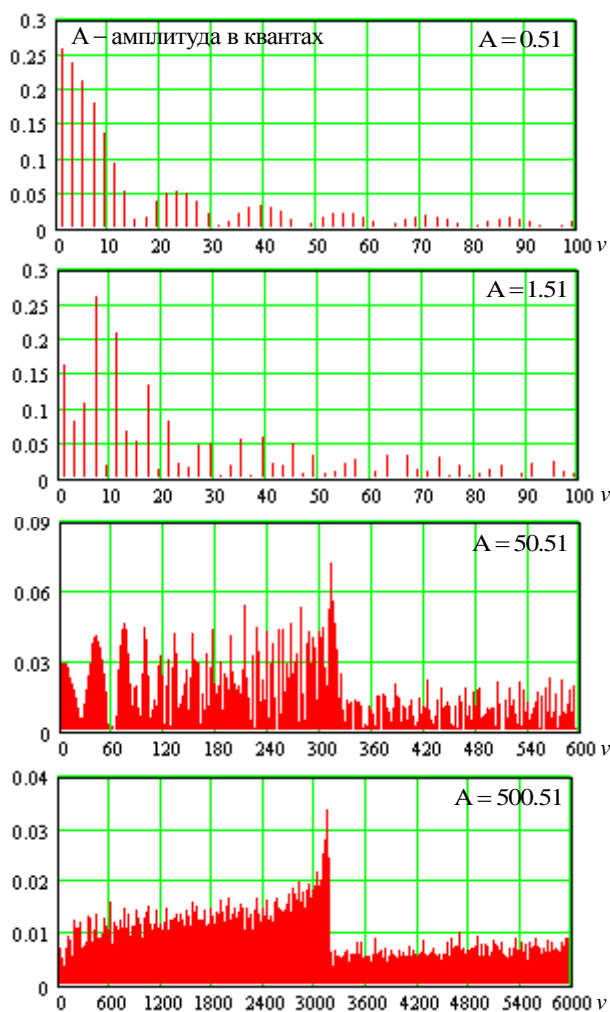


Рис.4.3. Спектры гармоник ошибок округления

При минимально возможной амплитуде ЗС  $A = 0,51$  максимум спектра совпадает с первой гармоникой ошибок округления и спектр быстро сходится. При дальнейшем увеличении  $A$  максимальные значения составляющих спектра уменьшаются и смещаются все к более высоким гармоникам в соответствии с приведенным выше равенством. Амплитуды гармоник выше номера  $\nu(A_{\nu_{\max}})$  резко уменьшаются в 2...4 раза и образуют равномерный спектр. Приведенные графики справедливы в равной мере как при 8, так и 16-разрядном квантовании, когда амплитуды сигналов и ошибок выражаются в квантах. При использовании логарифмического масштаба по оси ординат вид графиков будет совершенной иной.

## 4.2. Спектр ошибок квантования

*Частотный спектр ошибок квантования* формируется в процессе дискретизации квантованного ЗС. При выполнении этой операции осуществляется *балансная модуляция*, при которой дискретизируются как ЗС, так и ошибки округления, преобразуясь в ошибки квантования, как это показано на рис.4.4. Мощности ошибок квантования и округления одинаковые, но спектры совершенно различные.

В спектр АИМ сигнала прежде всего входят нечетные гармоники спектра ошибок округления  $v$ , попадающие в полосу частот до частоты Найквиста  $F_N$ . Их число может быть несколько тысяч и спектр простирается до 3...5 мегагерц.

В результате модуляции при каждой гармонике частоты дискретизации образуются верхняя и нижняя боковые полосы. Верхние боковые полосы являются *прямым*, а нижние - *зеркальным отображением спектра ошибок округления*. Мощности компонент спектра во всех полосах равны мощности ошибок округления.

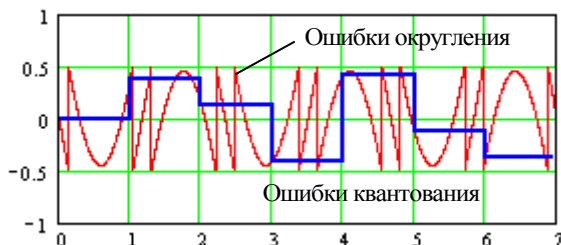


Рис. 4.4. Дискретизация ошибок округления

Необычность формирования спектра ошибок квантования заключается в том, что ширина спектра модулирующего сигнала во много раз превышает частоту дискретизации. Поэтому составляющие спектра нижних боковых полос попадают в область отрицательных значений (рис.4.5). Физически это означает, что происходит отражение спектров нижних боковых полос от нулевой частоты, и отрицательные значения частот становятся положительными.

Гармоники ошибок округления, частота которых не превышает частоту Найквиста, образуют *прямой спектр* ошибок квантования. Все гармоники выше  $f_N$  попадают в звуковой диапазон в виде продуктов модуляции

$$f_{\Delta} = s \cdot f_s - vF.$$

Эти продукты имеют знак плюс, если  $(s \cdot f_s - f_N) \leq v \cdot F \leq s \cdot f_s$ , и знак минус, если  $s \cdot f_s \leq v \cdot F \leq (s \cdot f_s + f_N)$ .

Сказанное поясним на частном простом примере кратных частот, когда  $F=1$  кГц и  $f_s = 48$  кГц. В этом случае в звуковой диапазон попадают гармоники ошибок округления с 1 по 23, образующие *прямой спектр*. Гармоники с 25 по 47 попадают в этот же диапазон,

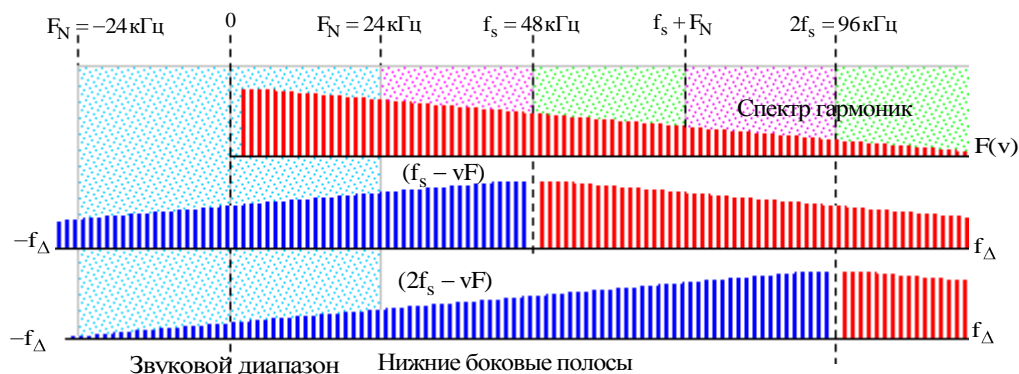


Рис.4.5. Формирование спектра ошибок квантования

как продукты модуляции нижней боковой полосы 1 порядка. Они создают частоты от 23 до 1 кГц с интервалом 1 кГц. Гармоники с 49 по 71 преобразуются в такие же частоты, но с отрицательным знаком. Аналогичные преобразования происходят с гармониками выше 71 в нижних боковых полосах 2 и более высокого порядков и т. д. Суммирование составляющих спектра происходит с учетом фазы каждой гармоники. Если частоты  $f_s$  и  $F$  не кратны, то при небольшом отклонении частоты ЗС от кратного значения каждая гармоника сигнала ошибок округления создает в звуковом диапазоне свою составляющую спектра.

Все вместе продукты модуляции в звуковом диапазоне частот формируют так называемый *зеркальный спектр*. Сумма спектральных компонент прямого и всех зеркальных спектров нижних боковых полос в области звукового диапазона является спектром ошибок квантования. В него всегда входят все гармоники ошибок округления в

виде продуктов модуляции. Так как зеркальные частоты могут совпадать по частоте между собой и с составляющими прямого спектра, то суммарная мощность спектральных составляющих может быть существенно меньше, чем мощность ошибок округления.

*Прямой спектр* определяется на основе комплексного спектра ошибок округления

$$S_{fr}(v) = 2 \text{fft}(e_r) \cdot m^{-0.5},$$

где  $v = 1, 2, \dots, v_{\max}$ ,  $v_{\max}$  – максимальный *нечетный номер гармоники ЗС* в полосе до частоты Найквиста,  $m$  – число точек отсчета за период ЗС. Этот номер рассчитывается с использованием Булевой алгебры:

$$v_{\max} = \begin{cases} \alpha, & \text{if mantissa}(\alpha) \neq 0 \\ (\alpha - 1), & \text{if mantissa}(\alpha) = 0 \end{cases}$$

где  $\alpha = \text{floor}\left(\frac{f_s}{2F}\right)$ ,  $\text{mantissa}(\alpha) = \frac{\alpha}{2} - \text{floor}\left(\frac{\alpha}{2}\right)$

Расчет составляющих *зеркального спектра* производится также на основе спектра ошибок округления

$$S_{fr}(v) = 2 \text{fft}(e_r) \cdot m^{-0.5}$$

где  $v = (v_{\max} + 2), (v_{\max} + 4) \dots (2n - 1)$ ,  $n$  – число рассчитываемых гармоник.

Частоты этих составляющих в звуковом диапазоне определяются равенством

$$F(v) = F(v) = |\gamma_s(v) \cdot f_s - F(v)|,$$

где

$$\gamma_s(v) = \text{round}\left(\frac{F(v)}{f_s}\right), \quad \gamma_s(v) - \text{это порядковый номер гармоники частоты}$$

дискретизации. Приведенные формулы позволяют рассчитывать частотный состав спектра ошибок квантования даже при отклонении ЗС от кратной частоты на доли герца.

К основным параметрам гармонического спектра относится нижняя граничная частота  $F_n(x)$  и интервал между составляющими  $dF$ . В общем случае, когда частоты  $f_s$  и  $F$  субкратны и отношение этих частот может быть представляется простой дробью  $y/x$ , период амплитудно-модулированного сигнала после дискретизации в  $x$  раз больше периода ЗС. Поэтому нижняя граничная частота спектра ошибок квантования определяется равенством

$$F_n(x) = \frac{F}{x}.$$

Посередине каждого интервала между кратными частотами субкратная частота имеет максимальный порядок  $x = 2$ . В обе стороны от этого значения порядок субкратности

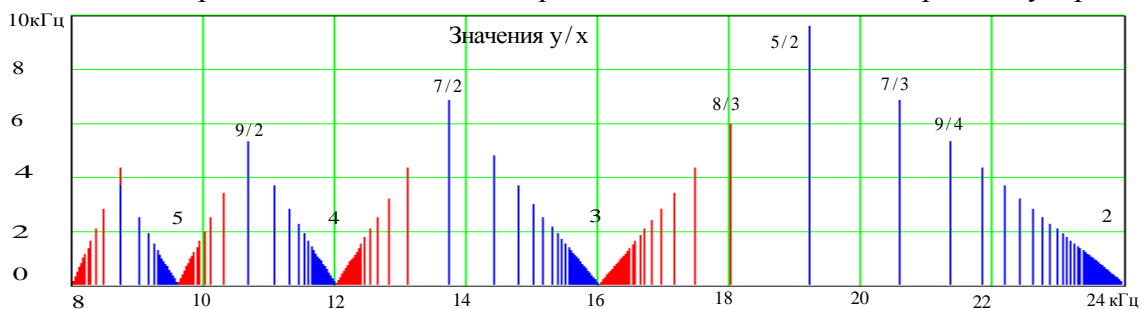


Рис.4.6. График изменения  $F_n(x)$  от частоты ЗС

увеличивается, стремясь к бесконечности вблизи кратных частот. Нижняя граница спектра ошибок квантования на этих частотах максимальна и она уменьшается в обе стороны от середины до нулевого значения вблизи кратных частотах (рис.4.6). Иногда эта частота является основным тоном звука ошибок квантования.



При отклонении частоты ЗС возникает низкочастотный звук, основной тон которого прямо пропорционален величине отклонения  $\Delta F$ . Высота основного тона этого звука определяется равенством

$$F_t = |\Delta F_{kr}| \cdot y_{kr},$$

где  $y_{kr}$  – значение числителя коэффициента  $k$  на кратной (субкратной) частоте,  $\Delta F_{kr}$  – отклонение от кратной (субкратной частоты). Из этого равенства следует, что чем больше значение  $y_{kr}$  (ниже частота ЗС), тем быстрее увеличивается высота тона отклонения. Такой же звук возникает и при отклонении от субкратных частот. Если  $y_{kr}$  число четное, то в спектре присутствуют как четные, так и нечетные гармоники  $F_t$ . Если же это число нечетное, то в спектре присутствуют только нечетные гармоники тона отклонения. Частота  $F_t$  всегда кратна или равна *нижней граничной частоте спектра*  $F_n(x)$ .

В спектре ошибок округления присутствуют только нечетные гармоники, так как функция ошибок округления всегда четная. Функция ошибок квантования может быть как четной, так и нечетной в зависимости от значения  $y$  коэффициента кратности. На [рис.4.7](#). приведены 4 временные диаграммы *функций ошибок квантования* с разными коэффициентами кратности.

Из графиков видно, что функции с четным значением  $y$  обладают симметрией 1 рода,

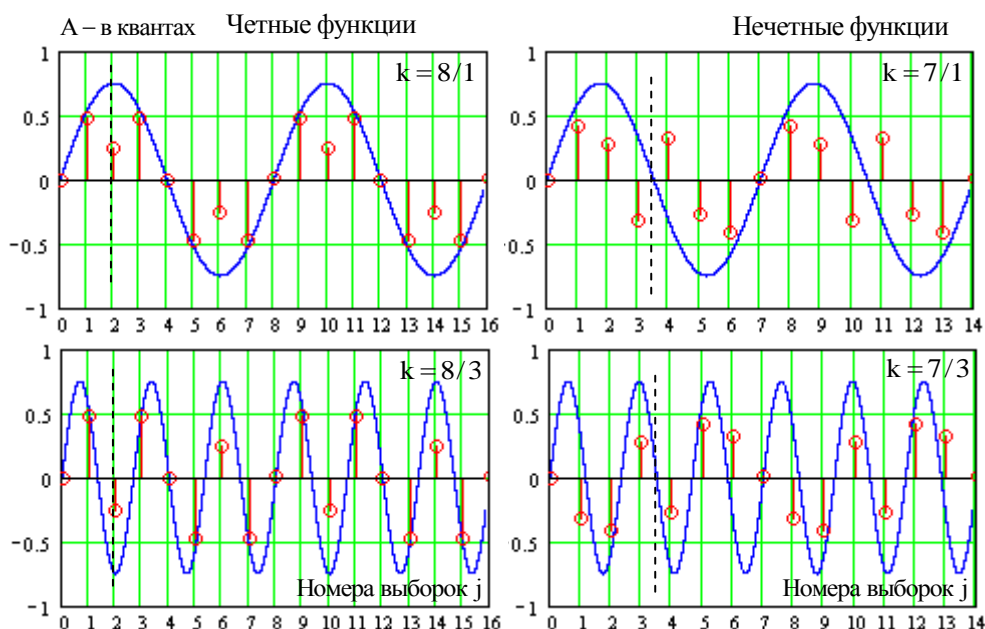


Рис.4.7. Функции ошибок квантования с разными коэффициентами кратности

а у функций с нечетным значением  $y$  – симметрия второго рода. Это значит, что если  $y$  число четное, то в спектре ошибок квантования кратных частот будут только нечетные гармоники частоты ЗС. Если же  $y$  число нечетное, то будут как четные, так и нечетные гармоники. Отличие для субкратных частот в том, что это будут гармоники не частоты ЗС, а граничной частоты спектра гармоник  $F_n(x)$ , а частота ЗС при этом будет одной из гармоник.

Спектр ошибок квантования всегда линейчатый, а частотный интервал между составляющими спектра  $dF$  для кратных и субкратных частот равен

$$dF = \begin{cases} F_n(x), & \text{если } y - \text{нечетное} \\ 2F_n(x), & \text{если } y - \text{четное} \end{cases}$$

Это иллюстрируется графиками спектров ошибок квантования при четных и нечетных значениях  $y$ , приведенными на [рис.4.8](#) и [4.9](#). На этих графиках красными квадратиками обозначены нечетные гармоники прямого спектра ошибок квантования, а синими кружочками – спектральные составляющие зеркального спектра.

Как видно, спектр ошибок квантования *на кратных частотах* состоит только из гармоник ЗС, и все составляющие зеркального спектра совпадают с гармониками прямого спектра. Если  $y$  четное – в спектре присутствуют только *нечетные гармоники*,

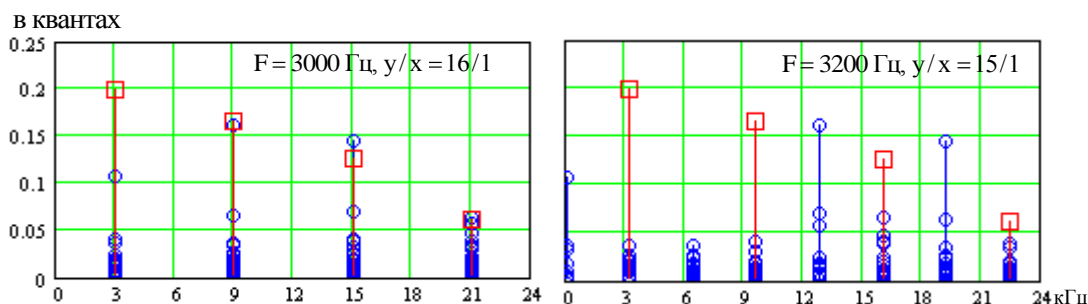


Рис.4.8. Спектры ошибок квантования на кратных частотах ЗС ( $A=3,49$ )

если нечетное, то в спектре *появляются четные гармоники* ЗС и, что очень важно, составляющая на нулевой частоте. Все они порождаются только зеркальным спектром.

На *субкратных частотах* в спектре возникают составляющие ниже частоты ЗС, и спектр становится гармоническим по отношению к нижней граничной частоте. При этом *нижняя граничная частота рассматривается как частота основного тона*, а ее гармоники - как *обертон*, которые создают своеобразную окраску звука и тембр их звучания зависит от характера огибающей спектра гармоник. В этом созвучии высшие гармоники могут значительно превышать уровень нижней граничной частоты. Слух определяет частоту основного тона независимо от того, есть четные гармоники или нет.

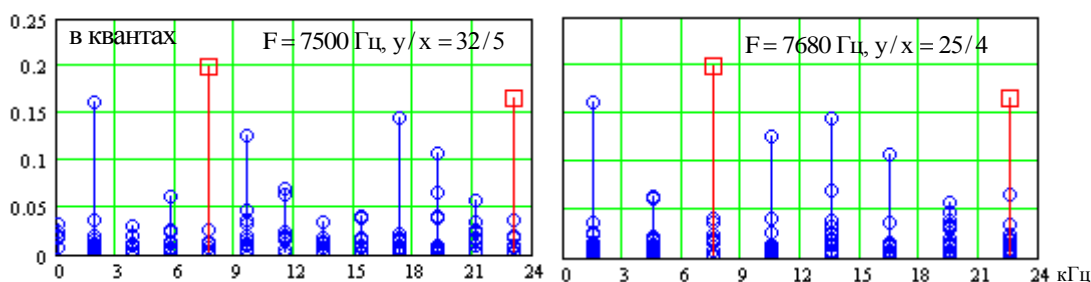


Рис.4.9. Спектры ошибок квантования на субкратных частотах ЗС ( $A=3,49$ )

Однако, это возможно, если число гармоник не менее 5-6. Если их меньше, то они воспринимаются как отдельные звуки.

Из приведенного анализа следует, что частотный состав спектра ошибок квантования однозначно определяется коэффициентом кратности, а составляющие этого спектра также однозначно определяются амплитудой ЗС в квантах. Таким образом, графики спектров на [рис.4.8](#) и [4.9](#) в равной мере справедливы как при 8, так и 16-разрядном квантовании, так как все составляющие выражены в квантах.

Эти же графики будут совершенно иными, если спектры, как это принято,

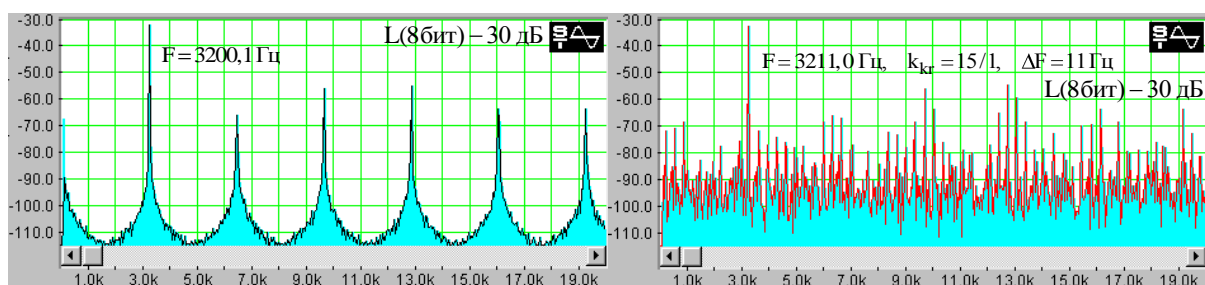


Рис.4.10. Спектры ошибок квантования при отклонении частоты ЗС от кратного значения

представлять в логарифмической шкале по оси ординат относительно нулевого электрического уровня (шкала FSdB). В этом случае спектры ошибок квантования при 8 и

16-разрядном квантовании будут одинаковы при существенно разных уровнях ЗС  $L(8bit)$  и  $L(16bit)$ . Связь этих уровней квантования описывается равенством

$$L(16 bit) = L(8 bit) - 48, dB.$$

Из этой формулы следует, что по отношению сигнал/шум и спектру ошибок квантования уровень минус 48 дБ 16-разрядного цифрового звука соответствует нулевому уровню 8 разрядного звука. Это значит, что ниже уровня минус 48 дБ 16-разрядный звук воспринимается как 8-разрядный, со свойственными ему проблемами с ошибками квантования. Однако, крайне неприятному 8-разрядному звуку с уровнем ЗС минус 30 дБ соответствует 16-разрядный звук с уровнем ЗС уже минус 78 дБ и, следовательно, ошибки квантования проявляются только на очень тихой музыке, близкой к порогу слышимости. Важно, что частотная характеристика слуха очень сильно зависит от уровня звукового давления, поэтому одинаковые спектры субъективно воспринимаются совершенно по разному и часто они вообще не слышны.

Таким образом, механизм формирования и изменения спектра ошибок квантования в зависимости от различных факторов не зависит от числа используемых разрядов. На этом основании измерение спектров ошибок квантования и прослушивание проводились только при 8-разрядном квантовании.

Особо важно как изменяется спектр и звук ошибок квантования, когда частота цифрового звука отклоняется от кратного значения на небольшую величину  $\pm\Delta F$  от 0 до 50 Гц. Если это происходит с аналоговым звуком, то на слух это заметить трудно. В случае цифрового звука около всех гармоник, *включая нулевую*, возникают боковые полосы спектральных компонент зеркального спектра, в которых преобладают гармоники тона отклонения. При малой величине отклонения они быстро затухают, с увеличением отклонения полосы начинают перекрываться и спектр становится более равномерным (рис.4.10). Составляющие спектра тона отклонения  $F_t$  максимальны вблизи всех гармоник частоты  $F_n(x)$  спектра.

Если, например,  
 $F = 3211 \text{ Гц}$ , то  $k_{kr} = 15/1$ ,  $k = f_s / F = 48000 / 3211$ ,  $x = 3211$ ,  $F_n(x) = F / x = 1 \text{ Гц}$ ,  
 $\Delta F = 11 \text{ Гц}$ ,  $F_t = y_{kr} \cdot \Delta F = 15 \cdot 11 = 165 \text{ Гц}$

и, следовательно, спектр состоит из нечетных гармоник  $F_n(x)$  с интервалом 2 Гц и нечетных гармоник тона отклонения с интервалом 330 Гц (рис.4.10). Причем амплитуды гармоник тона отклонения на 20...40 дБ больше других составляющих спектра, поэтому

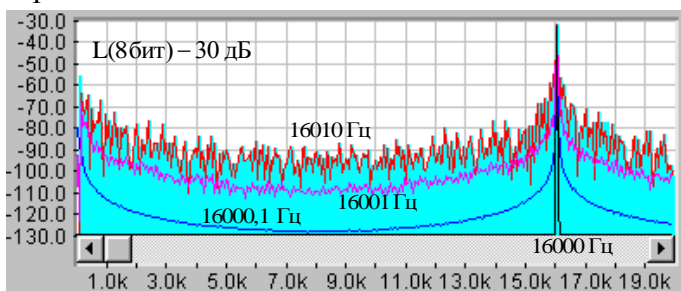


Рис.4.11. Спектры ошибок квантования при отклонении от кратной частоты 16 кГц

они хорошо видны на графике спектра и должны быть слышны. Таким образом, в приведенном примере, звук ошибок квантования порождается очень низкочастотными биениями нечетных гармоник нижней граничной частоты спектра и биениями гармоник тона отклонения. Низкочастотные биения на слух воспринимаются как рокот с большим числом неприятных обертонов. Звучание нечетных

гармоник тона отклонения ближе к созвучию. На слуховое восприятие этих биений сильное влияние оказывает частотная маскировка, которая на высоких и низких частотах ЗС проявляется по разному.

Зависимость формы огибающей спектра ошибок квантования от величины  $\Delta F$  иллюстрируется графиками на рис.4.11. При кратной частоте ЗС 16 кГц с уровнем минус 30 дБ в спектре все составляющие зеркального спектра имеют частоту 16 кГц, поэтому при 8-разрядном квантовании измеренное значение SNR составляет 146 дБ. При отклонении от кратного значения в спектре появляются все составляющие зеркального

спектра и значение SNR уменьшается до своего теоретического значения около 18 дБ, независимо от величины отклонения.

Как видно из графиков, пока отклонение мало, могут быть слышны, причем слабо, только очень низкие частоты и их гармоники. С увеличением  $\Delta F$  спектр выравнивается, и увеличиваются амплитуды составляющих спектра, особенно в области максимальной слышимости. Это связано, с увеличением частоты тона отклонения и интервалов между его гармониками. Так при  $\Delta F = 10$  Гц составляющие спектра в этом диапазоне возрастают на 30...40 дБ. Это значит, что с увеличением отклонения возрастает громкость звучания ошибок квантования.

Возникновение НЧ составляющих спектра при отклонении от кратных и субкратных частот вызывает ощущение не чистого, крайне неприятного звука, так как при этом искажается и временная структура сигнала ошибок квантования. При небольшом отклонении от кратных частот в сигнале на выходе ИКМ демодулятора всегда появляется огибающая по форме близкая к квантованной полусинусоиде с периодом равным  $1/F_t$  (рис.4.12), по форме напоминающая амплитудную модуляцию. Глубина этой «модуляции» достигает 100% при  $k = 2$  и она медленно уменьшается с увеличением значения  $k$ . Ее возникновение связано с биениями составляющих спектра вблизи кратной частоты. Эти биения создают амплитудную модуляцию сигнала ошибок квантования, которая хорошо заметна на слух.

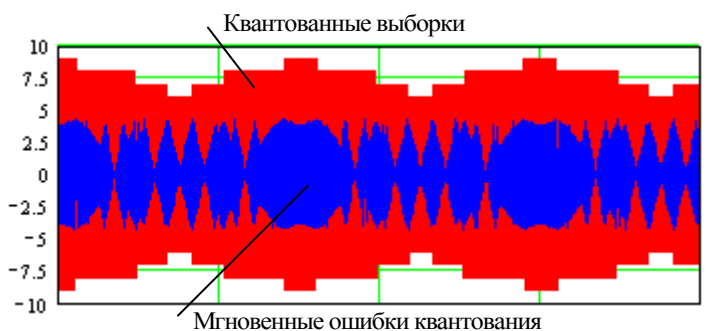


Рис.4.12. Огибающие сигналов квантованных выборок и ошибок квантования

На рис.4.12 приведены выходной сигнал ИКМ демодулятора и сигнал ошибок квантования, увеличенный в  $A$  раз при отклонении ЗС от частоты 12 кГц на 20 Гц. Как видно, периодичность огибающей АИМ сигнала передается на функции ошибок квантования и, следовательно, на ее спектр. Благодаря частотной фильтрации, амплитудные искажения огибающей синусоидального ЗС на выходе ФНЧ демодулятора устраняются во всем звуковом диапазоне частот, а зависимость огибающей сигнала ошибок квантования от коэффициента кратности остается.

Рассматривая спектры ошибок квантования необходимо хотя бы кратко остановиться на их зависимости от числа используемых разрядов и уровня ЗС. Если ограничиться увеличением числа разрядов от 8 и выше, то с достаточной точностью можно считать, что с каждым дополнительным разрядом все составляющие спектра ошибок квантования одинаково уменьшаются на 6 дБ. Это значит, что уже при 16 разрядах в большей части звукового диапазона они будут ниже абсолютного порога слышимости.

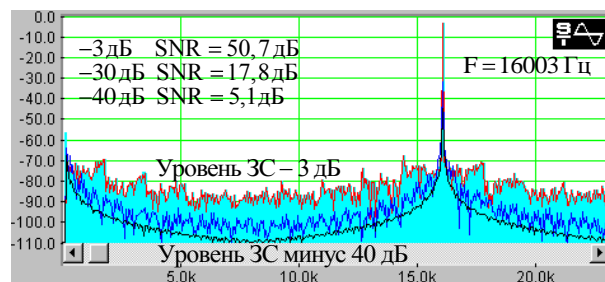


Рис.4.13. Спектры ошибок квантования при уровне ЗС минус 3, 30 и 40 дБ ( $q=8$ )

Зависимость спектра ошибок квантования от амплитуды ЗС на кратных и субкратных частотах проявляется в том, что одни составляющие спектра могут увеличиваться, а другие уменьшаться на несколько децибел или вообще исчезать. Так как составляющих много, это приводит к не очень заметному изменению громкости и тембра звучания ошибок квантования.

Эта зависимость выражена значительно сильнее, когда спектр ошибок квантования является почти сплошным и число составляющих спектра очень велико. Из графиков на [рис.4.13](#) создается иллюзия, что интегральная мощность шума квантования становится больше с увеличением уровня ЗС, что противоречит теории. На самом деле измеренные значения SNR мало отличаются от расчетных, определяемых равенством

$$SNR = 6,02 \cdot q + 1,76 - L,$$

где  $q$  - число разрядов,  $L$  - уровень ЗС в дБ.

Этот любопытный феномен объясняется тем, что с увеличением уровня ЗС расширяется спектр ошибок округления и число составляющих зеркального спектра в звуковом диапазоне увеличивается. Однако, они все когерентны уже имеющимся составляющим, поэтому происходит не энергетическое суммирование, а алгебраическое, с учетом фаз. В результате происходит усиление одних составляющих и ослабление других. Общая мощность ошибок квантования при этом не меняется. Так как увеличение уровня составляющих спектра достигает 25...30 дБ, то громкость звука ошибок квантования с увеличением уровня ЗС становится существенно больше.

Аналогичные преобразования спектра ошибок квантования происходят и при 16-разрядном кодировании, когда амплитуда ЗС составляет 1...10 квантов, что соответствует уровню ниже минус 70 дБ.

### 4.3. Звучание ошибок квантования

Характер звучания ошибок квантования можно уверенно оценить только при относительно грубом 8-разрядном квантовании, при котором абсолютный порог слышимости меньше влияет на результаты прослушивания. При этом ошибки квантования слышны почти во все звуковом диапазоне и вызываемые ими искажения замечают даже не подготовленные слушатели.

*На кратных частотах* звучание ошибок квантования в большинстве случаев вообще незаметно или проявляется только в изменении тембра. На высоких частотах 16 и 12

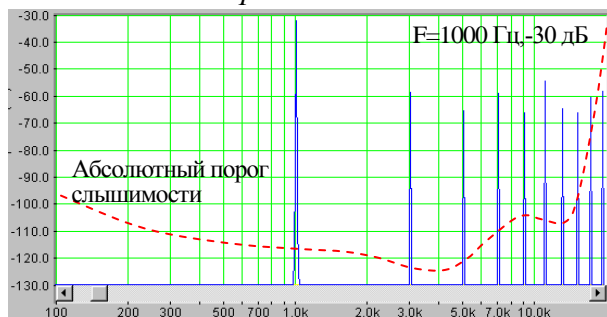


Рис.4.14. Спектр ошибок квантования

кГц, у которых в звуковом диапазоне нет никаких гармоник, естественно, ошибки квантования не слышны. В диапазоне от 2 до 9,6 кГц 2-3 гармоники попадают в область с высоким абсолютным порогом слышимости, поэтому ошибки квантования тоже не слышны. В этом диапазоне на частотах 2 и 3,2 кГц дополнительно имеет место еще частотная маскировка ближайших гармоник.

На ЗС с частотой 1 кГц почти все гармоники находятся выше абсолютного порога слышимости ([рис.4.14](#)). При уровне ЗС минус 30 дБ они слышны в виде свистящего звука, но с увеличением уровня звука начинает действовать эффект частотной маскировки. На частотах 500, 100 и 50 Гц большая часть гармоник попадает в область максимальной слышимости от 2 до 5 кГц. Они создают приятно звучащее созвучие, в котором обертона звучат громче, чем основной тон.

На ЗС с частотой 1 кГц почти все гармоники находятся выше абсолютного порога слышимости

*При небольшом отклонении частоты цифрового звука от кратного значения* всегда возникает низкочастотный тон отклонения, который особо заметен на высоких частотах при нечетных значениях  $u_{kr}$ . Это связано с возникновением боковых полос из гармоник тона отклонения около всех гармоник ЗС, включая нулевую. В качестве примера на [рис 4.15 А](#) приведены спектры ошибок квантования при отклонении от кратных частот ЗС 12 и 16 кГц на 3 Гц. Как видно, при частоте ЗС 16 кГц составляющие спектра вблизи нулевой частоты на 30...40 дБ выше, чем при частоте ЗС 12 кГц, поэтому громкость их

звучания выше. Из спектра в области низких частот, приведенного на рис.4.15B, следует, что слышны нечетные гармоники тона отклонения. При частоте ЗС 12 кГц в спектре вблизи нулевой частоты нет гармоник тона отклонения и звук ошибок квантования определяется биениями гармоник нижней граничной частоты спектра  $F_n(x)$ . Звук крайне неприятный.

Звук тона отклонения определяется и биениями его гармоник около частоты ЗС и его гармоник. Однако, низкочастотные биения слышны только при условии, если хорошо слышны сами частоты, которые создают биения. Это происходит только при частотах ЗС

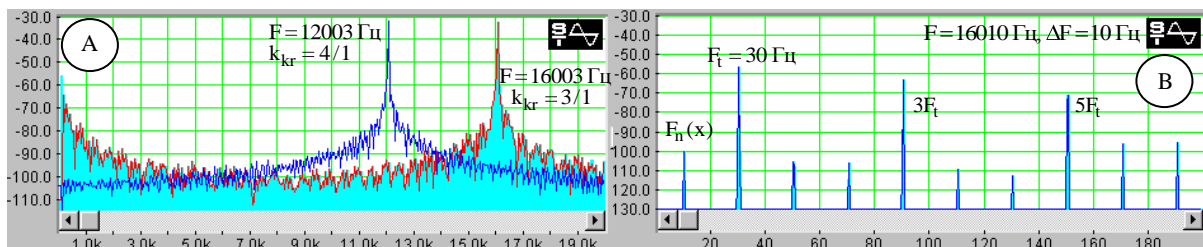


Рис.4.15. Спектры ошибок квантования при отклонении частоты ЗС от кратного значения

ниже 5...6 кГц.

Пока величина отклонения от 16 кГц ( $k=3/1$ )  $\Delta F$  не превышает 0,03...5 Гц слышны периодические сухие щелчки, выше – биения в виде глухого рокота и с 50 Гц – звонкое созвучие с основным тоном  $F_1$ , причем, чем больше отклонение, тем уровень громкости звука выше. Это связано с тем, что с увеличением  $\Delta F$  гармоники тона отклонения перемещаются в область максимальной слышимости.

Эффект меньшей громкости звучания ошибок квантования при небольших отклонениях от кратных частот с четным значением  $y$  имеет место и на всех более низких кратных частотах, но он не так сильно выражен. Это связано с тем, что звук тона отклонения порождается биениями гармоник около гармоник ЗС ниже частоты 5.6 кГц. Начиная с 3 кГц, звук отклонения становится все менее заметным на слух и на частотах меньше 1000 Гц из-за маскировки его практически не слышно.

На субкратных частотах в спектре ошибок квантования возникают составляющие ниже частоты ЗС, которые обычно воспринимаются на слух как тональный призыв. Это

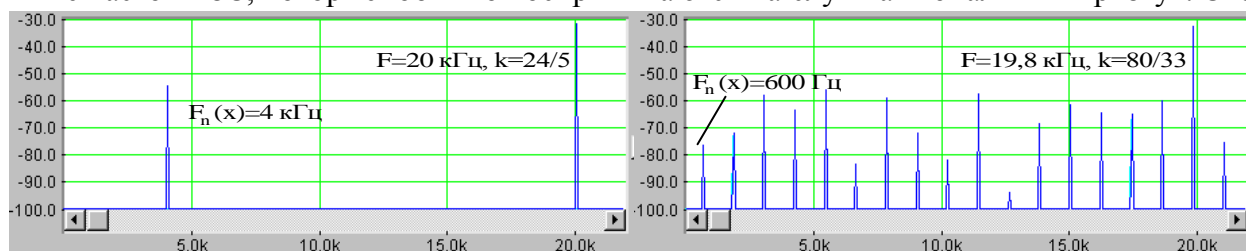


Рис.4.16. Спектры ошибок квантования на субкратных частотах

особо проявляется на высоких частотах с малым порядком субкратности, как, например, на 20 кГц (рис.4.16), когда вместо ЗС сигнала слышен только призыв с частотой 4 кГц. Объясняется это тем, что основной тон созвучия определяется на слух при числе нечетных гармоник не менее 5-6. Если их меньше, то они воспринимаются как отдельные частоты.

Если значение  $x$  достаточно велико, спектр ошибок квантования на высоких частотах простирается вниз до очень низких частот и тогда возникает призыв в виде созвучия с частотой основного тона  $F_n(x)$  (рис.4.16), гармоники этой частоты определяют тембр созвучия. На более низких частотах призывы маскируются ЗС. На субкратных частотах незаметны различия в звучании ошибок квантования при четном и нечетном значении  $y$ .

При небольшом отклонении частоты цифрового звука от любого субкратного значения происходит расширение спектра, резко увеличивается число его составляющих и всегда возникает низкочастотный звук, также как этот имеет место при отклонении ЗС

от кратного значения. При этом звучание ошибок квантования также зависит от значения у коэффициента кратности (четное он имеет значение или нечетное).

Главное отличие, оказывающее влияние на звучание ошибок квантования при отклонения ЗС от субкратных частот от аналогичного отклонения от кратных частот в том, что у этих сигналов исходно принципиально разные спектры ошибок квантования. При кратных частотах этот спектр простирается только вверх от частоты ЗС, где биения слышны плохо. При субкратных частотах он простирается от высоких частот вниз (рис.4.16), в области, где биения гармоник тона отклонения слышны хорошо. Кроме того, на субкратных частотах значения у достигают 100 и более, поэтому даже при очень небольшом отклонении от субкратной частоты возникает хорошо слышимый тон  $F_t$ .

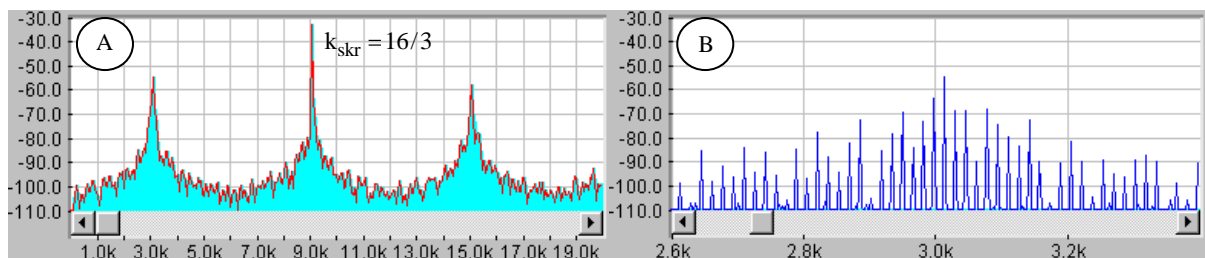


Рис.4.17. Спектр ошибок квантования при отклонении от субкратной частоты ЗС 9 кГц на 1 Гц  
А- полный спектр, В – спектр в области биений в диапазоне 2800...3200 Гц

Возникающие призвуки особенно заметны на высоких субкратных частотах, когда НЧ составляющие спектра попадают в область максимальной слышимости от 2 до 4 кГц (рис.4.17). Например, при частоте ЗС 9001 Гц низшая граница спектра равна 1 Гц, но уровень спектральных составляющих в этой области ниже минус 100 дБ. Зато в районе частоты 3 кГц возникает множество составляющих с интервалом равным частоте тона отклонения 16 Гц. Между ними возникают биения, которые очень хорошо слышны.

На субкратных частотах выше 5 кГц, на которых спектральные компоненты ошибок квантования попадают в область максимальной слышимости 2...4 кГц (рис.4.17), звучание ошибок квантования может приобретать очень своеобразный характер. Когда отклонению на  $\Delta F$  соответствует тон около 5 Гц, возникает звук похожий на звон колокольчиков, раздражающе действующий на слух. С увеличением отклонения звуковые биения становятся глухими и с 50 Гц постепенно переходят в созвучие.

При отклонении от кратных и субкратных частот ЗС ниже 4...5 кГц, при котором возникает тон отклонения с частотой 5 Гц, всегда звучание ошибок квантования похоже на переливы колокольчиков. Высокая слышимость биений с частотой 5 Гц связана с хорошо известным в психоакустике фактом, что частотная и амплитудная модуляция ЗС наиболее заметны при частоте модуляции 4..5 Гц.

Таким образом, при синусоидальных звуковых испытательных сигналах независимо от частоты и уровня ЗС ошибки квантования никогда не воспринимаются как шум. Даже удивительно, что при очень маленьких отклонениях ЗС от кратных и субкратных частот, когда частотный интервал между составляющими спектра ошибок квантования меньше 1 Гц во всем звуковом диапазоне, их звучание даже близко не напоминает шум и воспринимается как низкочастотный рокот. Если же в качестве испытательного сигнала использовать несколько некогерентных тональных звуков, то тогда биения составляющих спектра ошибок квантования воспринимаются как обычный белый шум.

На тестовых синусоидальных сигналах никаких особо неприятных для слуха звуков ошибки квантования не создают. На субкратных частотах из-за них возникают НЧ призвуки в виде тона или созвучия, либо звонкие биения. В связи с очень большим числом кратных и субкратных частот, особенно на высоких частотах, наиболее

характерным искажением является звук в виде рокота. На музыкальных сигналах с 8-разрядным квантованием этот рокот воспринимается как «грязное звучание» или дробовой шум. При 16-разрядном квантовании ни на тестовых синусоидальных сигналах, ни на музыке ошибки квантования практически не слышны.



## 5. Специальные аудио технологии

### 5.1. Технология Dithering

**Особенности цифрового звука.** Цифровой звук формируется с помощью квантователя с принципиально нелинейной передаточной функцией в виде лестницы (рис 5.1.). По этой причине порождаются детерминированные ошибки квантования звуковых сигналов (ЗС), спектр и характер звучания которых зависит как от коэффициента кратности частоты ЗС и частоты дискретизации, так и от вида самого ЗС. Спектр может быть гармоническим, субгармоническим и близким к белому шуму. На слух эти ошибки наиболее часто воспринимаются в виде неприятного ощущение грязного звука.

Округление цифровых данных при квантовании может производиться по двум алгоритмам: до ближайшего двоичного значения или к меньшей величине. Первый алгоритм принято называть *rounding*, второй - *truncating*. При аналого-цифровом преобразовании на основе импульсно-кодовой модуляции используется первый алгоритм округления, который реализуется путем смещения шкалы квантования (рис.5.1). В этом случае выходной сигнал квантователя симметричен относительно оси времени, и квантование осуществляется *с порогом, равным 0,5 шага квантования Q*.

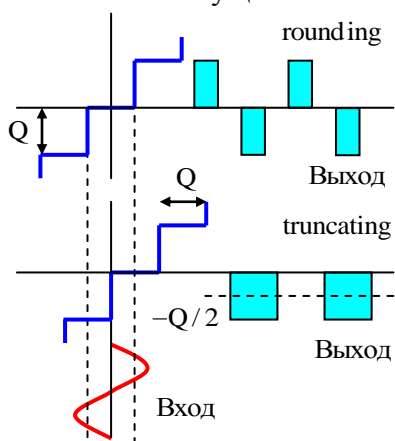


Рис. 5.1. Передаточные функции квантователя

Пока входной сигнал меньше этого порога выходной сигнал квантователя равен нулю, это значит, что квантование осуществляется с центральной отсечкой. Выше порога квантования выходной сигнал имеет вид последовательности импульсов со скважностью, зависящей от уровня ЗС.

Ситуации, когда входной сигнал может быть ниже порога квантования, возникают при реверберации звука и отражении звуковых колебаний, которые очень медленно затухают по амплитуде, но несут важную акустическую информацию об объеме помещения, где происходит звукозапись. При квантовании эта информация частично теряется и *пространственное восприятие звука сильно искажается*.

Величина ошибок квантования определяется только числом используемых двоичных разрядов  $q$ , и, следовательно, выбранным шагом квантования  $Q$ . При выбранном числе разрядов эффективное значение ошибок квантования не зависит от уровня ЗС и определяется равенством

$$E_{qef} = 0,289Q$$

Поэтому в цифровом звуке, в отличие от аналогового, чем меньше уровень ЗС тем относительные искажения квантования больше, а при уровне ЗС близком к 0,5 кванта они достигают 100%. Искажения, вызванные ошибками квантования, *уверенно заметны на слух* уже при 8-разрядном кодировании, хотя при этом величина искажений не превышает 0,5%. Это значит, что при 16-ти разрядном кодировании, используемом при записи CD, *реальный динамический диапазон* цифрового звука не превышает 48 дБ, а не 96 дБ, как это пишется в рекламе.

Указанные особенности цифрового звука в значительной мере устраняются или делаются менее заметными на слух при использовании технологии *Dithering*. Эта технология при аналого-цифровом преобразовании позволяет линеаризовать передаточную функцию квантователя, расширить реальный динамический диапазон ЗС,

преобразовать детерминированные ошибки квантования в случайные и при этом существенно понизить их заметность на слух.

При аналого-цифровом преобразовании ЗС технология *Dithering* заключается в добавлении к сигналу небольшого шума. 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.5.1). 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.

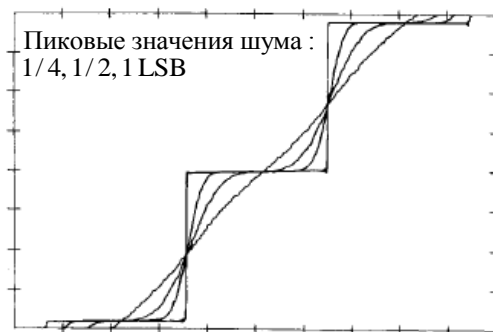


Рис.5.2. Линейризация передаточной функции квантователя

Технология *Dithering* решает три основные задачи. Первое, это *линейризация передаточной функции квантователя*. Под действием шума дизеринга в процессе квантования происходит хаотическое перескакивание между соседними уровнями и в результате передаточная функция спрямляется (рис.5.2.). Чем больше пиковые значения шума относительно шага квантования, тем линейность передаточной функции становится лучше.

Вторая задача технологии *Dithering* - это *декорреляция ошибок квантования*.

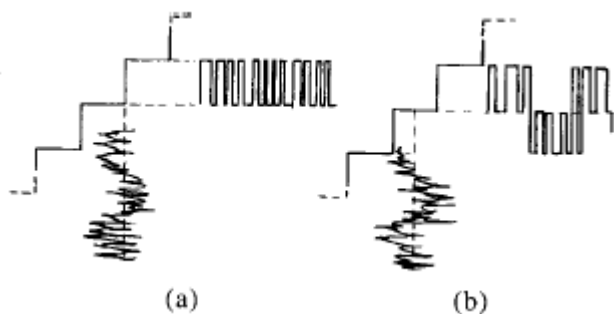


Рис.5.3. К вопросу расширения динамического диапазона АЦПс помощью технологии *Dithering*  
a) – *truncating* b) *rounding*

Благодаря шуму дизеринга происходит беспорядочное дрожание квантуемого сигнала относительно шкалы квантования. Поэтому ошибки квантования становятся случайными, корреляционные связи между ними уменьшаются или вообще пропадают. При этом происходит декорреляция и рандомизация детерминированных ошибок квантования. Чем больше пиковые значения шума относительно шага квантования, тем сильнее декорреляция ошибок квантования но сильнее ухудшается отношение

сигнал/шум. В результате гармонические и субгармонические искажения преобразуются в белый шум, который на слух действует менее раздражающе и субъективно не искажает музыкальные сигналы.

Третья задача технологии *Dithering* - это *расширение полного динамического диапазона АЦП*, в котором используется операции округления. В таких АЦП пока пиковое значение ЗС на входе не превышает половины шага квантования, выходной сигнал отсутствует, а динамический диапазон *DR* определяется числом разрядов *q*

$$DR = 6 \cdot q$$

Если, например,  $q=8$ , это значит, что ЗС с уровнем ниже минус 48 дБ через АЦП не передаются.

Расширение динамического диапазона за счет передачи через АЦП ЗС с амплитудой меньше 0,5 кванта при использовании технологии *Dithering* достигается необычным путем благодаря возникновению паразитной *модуляции шума, вводимого при использовании технологии Dithering звуковым сигналом.*

Это поясняется **рис.5.3**. При использовании технологии квантования *rounding* вводимый шум с амплитудой чуть больше 0,5 кванта на входе квантователя порождает на

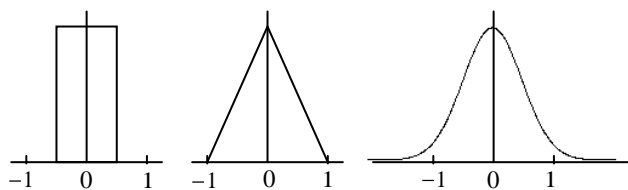


Рис. 5.4. Законы изменения плотности вероятности пиковых значений шума (PDF)

выходе стационарный поток двух полярных импульсов со случайной частотой и длительностью с амплитудой 1 квант. Благодаря нелинейности квантователя звуковой сигнал с уровнем меньше шума вызывает *модуляцию частоты повторения и длительности* случайной последовательности импульсов (временнй импульсная модуляция).

Естественно в спектре этой последовательности появляется составляющая ЗС, которая может быть выделена. В квантователе с технологией *truncating* происходит примерно тоже. Отличие только в том, что шум любого уровня на входе порождает стационарный поток на выходе одно полярных импульсов с амплитудой 1 квант.

*Аналоговый и цифровой шум*, используемые в технологии *Dithering* могут иметь различные *законы распределения плотности вероятности пиковых значений (PDF-Probability Density Function)* - такие как прямоугольный (*RPDF*), треугольный (*TPDF*) и Гаусса (*GPDF*) (**рис.5.4**). Пиковые значения шума могут быть от 0,5 младшего разряда до 10 и более разрядов.

Шум с гауссовым законом распределения возникает из-за тепловых шумов активных и пассивных элементов электронных схем. Поэтому он часто используется в операции «*dithering*» при аналого-цифровом преобразовании. Два других шума обычно реализуются в цифровом варианте, и применяется при реквантизации кодовых слов в процессе редактирования цифровых данных, при фильтрации и изменении уровня.

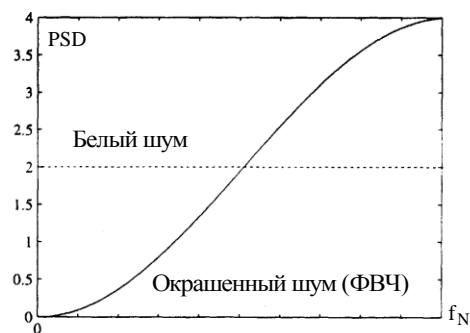


Рис.5.5. Графики спектральной плотности мощности шума

Таблица 5.1. Характеристики шума			
Вид функции PDF	Пиковые значения	Мощность шума	Уменьшение SNR, дБ
Прямоугольный-R	$\pm 0,5Q$	$Q^2/12$	-3
Прямоугольный-R1	$\pm Q$	$Q^2/3$	-6
Треугольный-T	$\pm Q$	$Q^2/6$	-4,77
Гаусса-G	$\pm 1,5Q$	$Q^2/4$	-6

отдельных случаях используется окрашенный шум с равномерным подъемом *PSD* в области высоких частот (**рис.5.5**). Такой окрашенный шум формируется из белого шума с использованием фильтра верхних частот (ФВЧ). В этом случае в названии закона распределения добавляется слово *shaped*. В звуковых редакторах разных фирм используются ФВЧ с различными передаточными функциями, поэтому значения *SNR* при одном и том же законе распределения шума могут существенно отличаться.

*Спектральная плотность мощности* вводимого шума (*PSD-Power Spectral Density*) обычно имеет равномерное распределение в диапазоне от нуля до частоты Найквиста  $f_N$  (белый шум). Лишь в

Надо отметить, что в звуковом редакторе WavLab в технологии *dithering* используется шум с равномерной спектральной плотностью в диапазоне от 20 до 22 кГц, поэтому шум квантования почти не слышен.

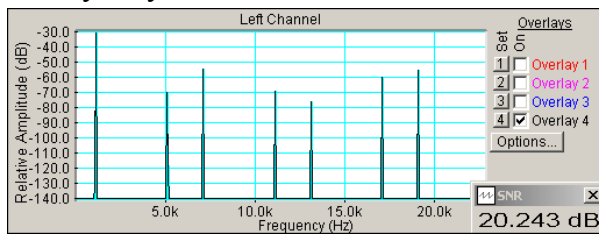


Рис.5.7. Спектр сигнала с реквантованием  $16 \Rightarrow 8$  бит

значения белого шума с широко известными функциями *PDF*, которые в технологии *Dithering* обеспечивают преобразование дискретного спектра ошибок квантования в белый шум. Как видно, это достигается при различных мощностях шума и

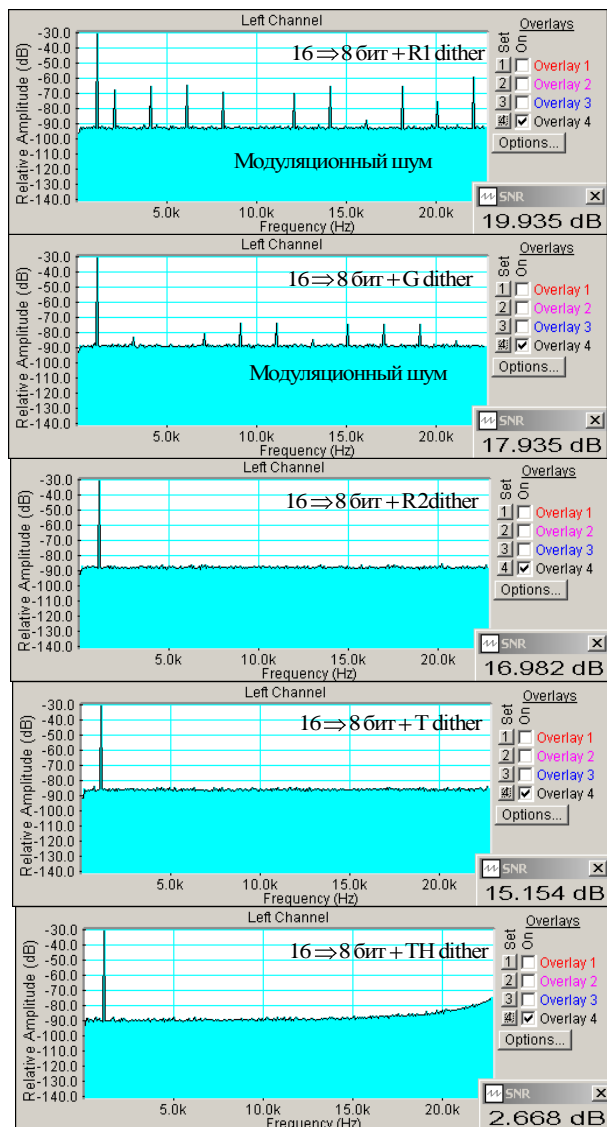


Рис.5.8. Спектры при реквантовании с различными законами распределения шума

сопровождается заметным ухудшением *SNR*. Если при реквантовании используется операция округления *rounding*, то для декорреляции ошибок квантования с помощью шума с прямоугольной функцией *PDF* достаточно, чтобы его пиковые значения достигали половины кванта. Если же применяется технология *truncating*, то для декорреляции ошибок требуется в 4 раза увеличить

Применение в технологии *dithering* шума для декорреляции ошибок квантования неизбежно сопровождается ухудшением объективных характеристик ИКМ тракта: шум на выходе квантователя увеличивается, а значение *SNR* уменьшается.

В табл.5.1. приведены пиковые значения белого шума с широко известными функциями *PDF*, которые в технологии *Dithering* обеспечивают преобразование дискретного спектра ошибок квантования в белый шум. Как видно, это достигается при различных мощностях шума и сопровождается заметным ухудшением *SNR*. Если при реквантовании используется операция округления *rounding*, то для декорреляции ошибок квантования с помощью шума с прямоугольной функцией *PDF* достаточно, чтобы его пиковые значения достигали половины кванта. Если же применяется технология *truncating*, то для декорреляции ошибок требуется в 4 раза увеличить мощность вводимого шума с тем, чтобы пиковые значения шума достигали 1 кванта. Шум с такими характеристиками используются в звуковых картах фирмы *Sound Forge* при выполнении операций реквантования звуковых фонограмм.

На рис.5.7 и 5.8. показано как видоизменяется спектр реквантованного сигнала и значение *SNR* в результате применения технологии *Dithering* с разными законами распределения пиковых значений шума. При реквантовании  $16 \Rightarrow 8$  бит без применения технологии *Dithering* (рис.5.7), спектр ошибок квантования остается дискретный, и в нем присутствуют некоторые четные и нечетные гармоники ЗС с уровнем минус 55...75 дБ.

При использовании технологии *Dithering*, как видно из рис.5.8, независимо от закона распределения шума дискретный спектр ошибок квантования преобразуется в белый или окрашенный шум со спектральной плотностью мощности минус 90 дБ. Чем меньше пиковые значения шума,

тем меньше значение *SNR*, но хуже декоррелируются ошибки квантования, а шум квантования модулируется гармониками ЗС

Когда реквантование осуществляется с использованием белого шума с прямоугольной функцией *PDF* –R1 потери *SNR* в этом случае практически отсутствуют.

Однако, при этом не достигается полная декорреляция ошибок, так как пиковые значения шума декорреляции не превышают 0,5 кванта. По этой причине в спектре остаются дискретные составляющие, которые создают модуляцию шума квантования ЗС и его гармониками. Это явление называется *модуляционным шумом*. Увеличение пиковых значений шума до одного кванта ( $R_2$ ) устраняет этот эффект, но при этом  $SNR$  уменьшается на 3 дБ.

Использование при реквантовании гауссового шума  $GPDF$  позволяет несколько уменьшить модуляционный шум, но значение  $SNR$  при этом ухудшается на 2 дБ. В случае применения при реквантовании шума с треугольной огибающей функции  $PDF$  ( $TPDF$ ), модуляционный шум полностью устраняется, но  $SNR$  уменьшается на 5 дБ. Наиболее сильно ухудшается  $SNR$  при использовании шума с треугольным законом распределения с повышением спектральной плотности шума в области высоких частот. В этом случае  $SNR$  равен 2,7 дБ, что на 17 дБ меньше исходного значения. Применение такого закона распределения позволяет снизить слуховую заметность шума квантования, несмотря на значительное ухудшение отношения сигнал/шум. Это объясняется тем, что высокочастотный шум слухом не воспринимается.

В настоящее время стандартом  $AES17$  принята технология декорреляции ошибок квантования, при которой используется случайный или псевдослучайный шум, у которого треугольная форма огибающей плотности вероятности ( $TPDF$  – *Triangle Probability Density Function*) мгновенных значений шума. Шум имеет размах от пика до пика  $2\text{ LSB}$  (*Least Significant Bit* – биты младших разрядов), и с равномерную спектральную 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, зависимо 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.

В соответствии с указанным стандартом при использовании технологии « $TPDF\ dither$ » значение  $SNR$  в функции числа разрядов  $q$  и уровня ЗС  $L_s$  рассчитывается по формуле

$$SNR(TPDF)_{\max} = 6,02q - 3,01, \text{ dBFS} . \quad SNR(TPDF) = SNR(TPDF)_{\max} + L_s, \text{ dBFS}$$

Если используется повышенная частота дискретизации  $f_s$ , то формула принимает вид

$$SNR(TPDF) = 10\lg\left(\frac{F_{\max}}{0,5f_s}\right) + 6,02q - 3,01 + L_s, \text{ dBFS} .$$

Уровень спектральной плотности мощности рассчитывается по формуле

$$L(PDF) = -SNR_{\max} - 10\lg(F_{\max}) .$$

При 8-разрядном квантовании и  $F_{\max} = 22\text{ кГц}$  с применением технологии « $TPDF\ dither$ »  $SNR = 15\text{ дБ}$ , и  $L(PDF) = -88,6\text{ дБ}$ , что соответствует графикам на [рис.5.8](#).

Для оценки слышимости ошибок и шума квантования необходимо учитывать особенности восприятия человеком тональных звуков и шума. Тональные звуки слышны только при условии, что их звуковое давление выше абсолютного порога слышимости.

При построении графика абсолютного порога слышимости в шкале  $FS$  следует учитывать:

- 1) 0 дБ в шкале  $FS$  всегда соответствует максимальный уровень звукового давления  $L_{\max} SPL$ .

2) 0 дБ в шкале *SPL* (порог слышимости) всегда соответствует уровню минус  $L_{\max} SPL$  в шкале FS.

Примем  $L_{\max} = 120dB SPL$ , что соответствует давлению, близкого к болевому порогу слуха человека. Из **рис.5.9** видно, при частоте 3С 1 кГц абсолютный порог слышимости

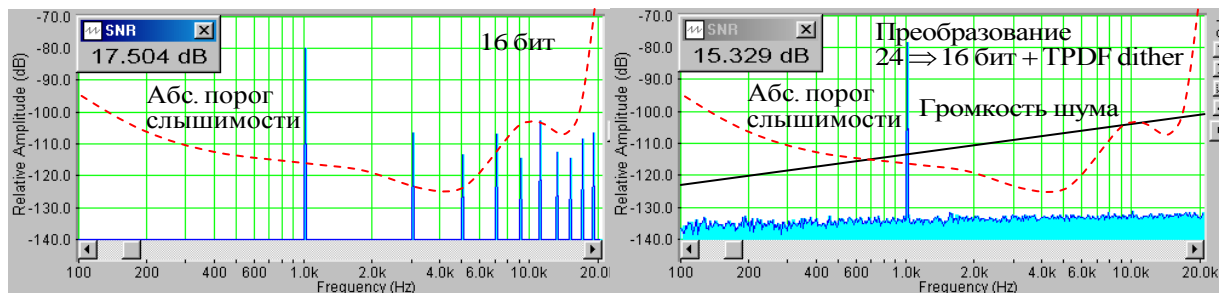


Рис.5.9. К вопросу слышимости ошибок и шума квантования

превышают только 3, 5 и 7 гармоники спектра ошибок квантования.

При решении вопроса слышимости шума квантования надо учитывать, что *органы слуха реагируют не на общую мощность шума, а на мощность шума в критических полосах слуха*. В качестве основных критериев слышимости шума принимаются его громкость и абсолютный порог слышимости тональных звуков (в шкале *SPL*). Шум слышен, если его уровень громкости выше абсолютного порога слышимости.

Расчет и построение графика громкости шума в функции частоты производится в следующем порядке. Сначала с помощью БПФ снимается график спектральной плотности мощности шума квантования при выбранном значении звукового давления. Затем производится его частотная коррекция, при которой учитывается, что громкость белого шума в критических полосах увеличивается с повышением частоты. Таблица корректирующих коэффициентов опубликованы, они меняются от 16,5 дБ на частоте 20 Гц до 33,9 дБ на частоте 20 кГц, на частоте 1 кГц этот коэффициент примерно равен 20 дБ. График частотной зависимости громкости белого шума достаточно точно представляется прямой линией с крутизной подъема 3 дБ/октаву (**рис.5.9**).

Как видно из этого рисунка, при звуковом давлении 120 *SPL* шум квантования слышен в диапазоне от 0,7 до 9 кГц и от 11 до 16 кГц. При увеличении звукового давления частотный диапазон слышимости шума квантования расширяется, а при его уменьшении сокращается. При звуковом давлении 100 дБ *SPL* шум квантования уже не слышен.

При увеличении числа разрядов реквантованного сигнала уровень шума квантования снижается на 6 дБ с каждым разрядом, поэтому в случае преобразования  $24 \Rightarrow 20$  бит плюс *TPDF dither* при давлении 120 дБ *SPL* уровень шума квантования находится на всех частотах ниже абсолютного порога слышимости.

При расчете и построении графика громкости шума квантования очень важно учитывать все возможные погрешности спектрального анализа с помощью БПФ, связанные с тем, что ширина полосы сканирования не равна 1 Гц и, следовательно, снятый в программе *SpectraLab* спектр не является графиком спектральной плотности мощности.

Ширина полосы сканирования зависит от выбора оконной функции, числа выборок и частоты дискретизации. Коэффициент коррекции определяется равенством

Табл.5.2.

Window	Scaling Bandwidth
None (rectangular)	1.00000 bins
Hamming	1.36283 bins
Hann	1.50000 bins
Blackman-Harris-4	2.00435 bins
Gaussian	2.21535 bins
Rife-Vincent-4	2.31000 bins
Rife-Vincent-5	2.62653 bins
Equiripple	2.63191 bins
Flat-top	3.82211 bins

$$K_{ww} = 10 \cdot \log\left(\frac{N_s}{SB \cdot f_s}\right),$$

где  $N_s$  – число используемых выборок (*FFT Size*),  $SB$  (*Scaling bandwidth*) – коэффициент оконной функции (из табл.5.2.). Если, например, выбрана оконная функция *Blackman-Harris-4*,  $N_s = 65536$  и  $f_s = 48$  кГц, то  $K_{ww} = -1,67$  дБ. С уменьшением *FFT Size* погрешность измерения спектральной плотности шума квантования возрастает, коэффициент коррекции соответственно также увеличивается и может достигать до 20 дБ.

Слуховая заметность шума квантования при использовании технологии *Dithering* может быть уменьшена, если с помощью ФВЧ осуществляется подъем спектра

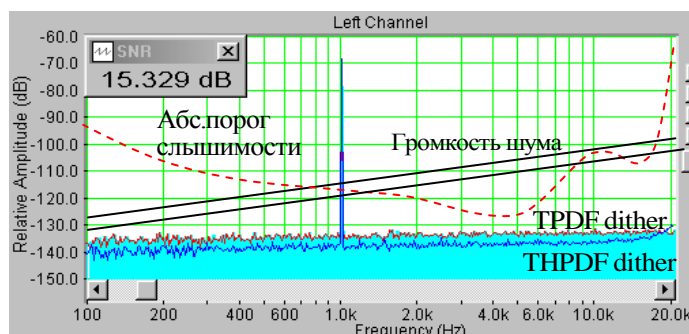


Рис.5.10. Громкость шума при реквантовании 24 ⇒ 16 бит при использовании TPDF и THPDF dither

вводимого белого шума в области частот выше 16...20 кГц до частоты Найквиста (рис.5.10). При использовании в этом случае шума с треугольным законом распределения (технология *THPDF dither*) уровень громкости шума квантования в звуковом диапазоне понижается 4...5 фон по сравнению с технологией *TPDF dither*. Такой алгоритм реквантования используется в звуковых картах фирмы Sound Forge под названием *High Pass Triangular Dither*.

**Заключение.** В области аудиотехники считается, что за последние 10 лет разработка и теоретическое обоснование применения технологии «*TPDF dither*» при аналого-цифровом и цифро-аналоговом преобразовании были наиболее значимыми практическими результатами по улучшению качества звучания цифровых записей.

Широкое применение технологии *Dithering* ограничивается тем, что после этой операции нельзя производить цифровое микширование, цифровую фильтрацию и цифровое регулирование уровня, иначе могут возникнуть серьезные искажения звука. Эта технология может быть только завершающей операцией. Поэтому при аналого-цифровом преобразовании она применяется только при записи мастер-дисков и мастер-лент, причем не всегда. Более широко ее применяют в процессе реквантования после окончания редактирования цифровых звукозаписей (микширование, изменение уровня, фильтрация...). Это связано с тем, что при выполнении этих операций вычисления часто производятся с 32-разрядными словами, а при записи используются только 16 разрядов.

## 5.2. Технология Noise Shaping

**Основы теории.** 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  $SNR$ .

В настоящем разделе рассматривается применение такой технологии только при импульсно-кодовой модуляции в процессе реквантования цифровых музыкальных фонограмм, 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, хотя значение  $SNR$  ухудшается.

Реализация технологии  $NS$  осуществляется путем использования в реквантователе частотно-зависимой отрицательной обратной связи по ошибке квантования. На рис.5.11 приведена простейшая схема реквантователя  $NS$  с элементом временной задержки  $\tau = 1/f_s$  в цепи отрицательной обратной связи на один такт. Этот элемент и сумматор

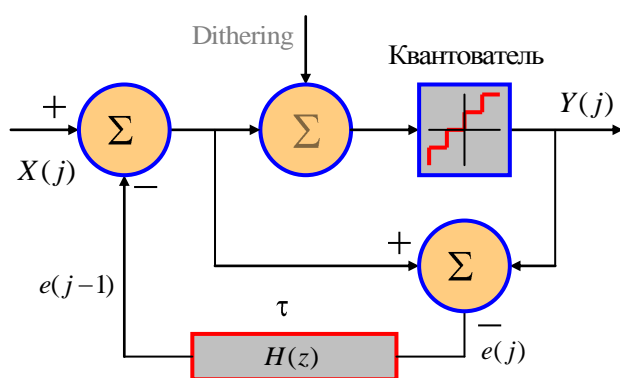


Рис.5.11. Схема реквантователя  $NS$  первого порядка

выполняют функцию фильтра высоких частот первого порядка для ошибок квантования. Реквантование может осуществляться с использованием операций rounding или truncating.

На входе и выходе схемы сигналы  $X(j)$  и  $Y(j)$  в цифровой форме, работа всех узлов тактируется с частотой дискретизации, а операция *Dithering* выполняется в цифровом виде. Применение этой операции линеаризует передаточную функцию квантователя и преобразует ошибки квантования в белый шум.

В принципе технология  $NS$  работает и без этой операции, только качество звука при этом значительно ухудшается. Это обстоятельство позволяет при выводе расчетных формул исключать операцию *Dithering* и рассматривать шум от внешнего источника как аддитивную составляющую.

В функции дискретного времени  $j$  (номер выборки) сигнал на выходе реквантователя можно представить в виде равенства

$$Y(j) = X(j) + e(j) - e(j-1),$$

из которого следует, что мгновенная ошибка реквантования определяется как

$$e(j)_{ns} = e(j) - e(j-1).$$

Эта формула отражает основную суть технологии  $NS$ . Очевидно, что ошибки двух последовательных выборок на низких частотах почти не отличаются и разностная ошибка стремится к нулю. На высоких частотах скорость изменения ЗС большая и эти ошибки могут отличаться очень значительно, они могут иметь и разную полярность, поэтому суммарная ошибка квантования сильно возрастает.

В приведенной схеме отрицательная обратная связь стремится уравнивать выходной сигнал с входным сигналом по ошибкам квантования. Поэтому чем больше число разрядов входного сигнала реквантователя, тем меньше уровень шума на выходе и больше динамический диапазон. При 16-разрядном выходном сигнале входной сигнал может иметь от 20 до 32 разрядов.



В цифровой технике для определения передаточных функций используются не преобразования Лапласа, а  $z$ -преобразования, поэтому в приведенной схеме передаточная функция цепи обратной связи определяется равенством

$$H(z) = z^{-1}$$

и выходной сигнал имеет вид

$$Y(z) = X(z) + (1 - z^{-1}) \cdot e(z).$$

Из этой формулы следует, что модули коэффициентов передачи по сигналу и ошибке соответственно равны:

$$|T_x(z)| = 1, \quad |T_e(z)| = 2 \left| \sin\left(\frac{\theta}{2}\right) \right|,$$

где  $z = e^{i\theta}$ ,  $\theta = 2\pi\eta$ ,  $\eta = \frac{f}{f_s}$ ,  $i = \sqrt{-1}$ ,  $i$  – мнимая единица,  $0 < \theta < \pi$ ,  $f$  – текущая частота,

$\eta$  – нормализованная частота,  $f_s$  – частота дискретизации. Таким образом, как следует из приведенной формулы, передаточная функция по сигналу не зависит от частоты, а передаточная функция для ошибки квантования имеет такой же вид как у фильтра верхних частот.

Наиболее важной характеристикой технологии *NS* является форма огибающей спектральной плотности мощности шума квантования (*PSD*) на выходе реквантизатора

$$S(e)_{ns1} = S(e) \cdot \left( 4 \left| \sin^2\left(\frac{\theta}{2}\right) \right| \right).$$

В этой формуле  $S(e) = const$ , это *PSD* шума квантования реквантователя (белый шум),

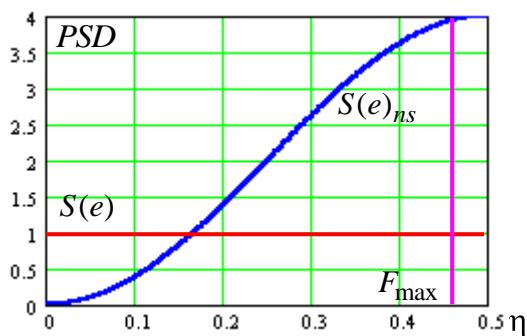


Рис.5.12. Спектральные плотности шума квантования

определяемая числом разрядов реквантователя без учета влияния отрицательной обратной связи

$$S(e) = \frac{Q^2}{12 \cdot f_s},$$

где  $Q$  – шаг квантования.

На рис.5.12. приведены графики этих спектров. Из них видно, что в реквантователе 1 порядка уменьшение спектральной плотности  $S(e)_{ns}$  происходит только на частотах ниже  $\eta = 0,16$ , что при  $f_s = 48$  кГц соответствует частоте ЗС 7,68 кГц. Выше этой частоты спектральная плотность увеличивается и на

частоте Найквиста она больше  $S(e)$  в 4 раза. Мощности шума квантования на выходе реквантователя с применением и без применения технологии *Noise Shaping* рассчитываются по формулам

$$P_{ns} = S(e) \cdot \left[ 2\theta_1 \cdot \left( 1 - \frac{\sin\theta_1}{\theta_1} \right) \right], \quad P = S(e) \cdot \theta_1,$$

где  $\theta_1 = \frac{2\pi F_{max}}{f_s}$ ,  $F_{max}$  – максимальная частота звукового диапазона.

Из этих равенств следует, что в диапазоне от 0 до частоты Найквиста ( $\eta = 0,5$ ) при использовании технологии *NS* мощность шума квантования удваивается. Без учета отрицательной обратной связи величина *SNR* на выходе реквантователя в звуковом диапазоне частот определяется известным равенством

$$SNR = 6,02q + 1,78 + 10 \cdot \log\left(\frac{f_s}{2F_{\max}}\right),$$

где  $q$ - число разрядов реквантователя. Отрицательная обратная связь приводит к ухудшению  $SNR$ , которое определяется величиной приращения

$$\Delta SNR_1 = -10 \cdot \log\left(2\left(1 - \frac{\sin(\theta_1)}{\theta_1}\right)\right)$$

Для реквантователя 1 порядка  $\Delta SNR_1 = -3$  дБ.

Для формирования огибающей спектральной плотности необходимой формы используются реквантователи более высокого порядка вплоть до 9...12. На [рис.5.13](#)

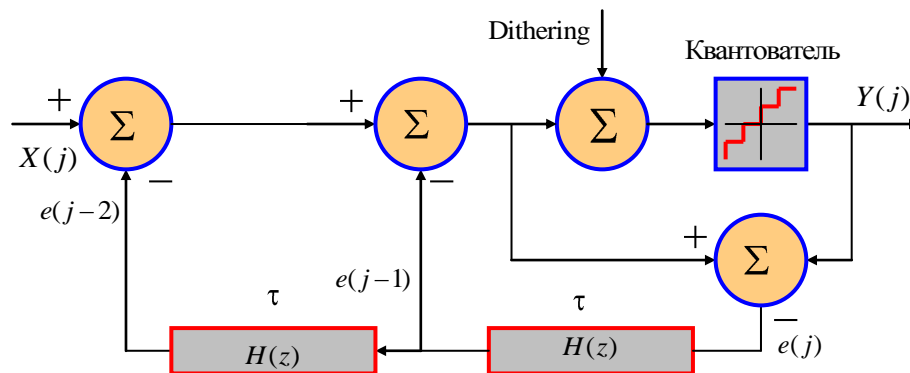


Рис.5.13. Схема реквантователя NS второго порядка

приведена схема реквантователя NS 2 порядка с двумя цепями отрицательной обратной связи. В этой схеме выходной сигнал реквантователя в форме  $z$  – преобразования представить в виде равенства

$$Y(z)_2 = z^{-1} \cdot X(z) + (1 - z^{-1})^2 \cdot e(z),$$

из которого не сложно определить модуль передаточной функции ошибки квантования

$$|T_e(z)_2| = \left[2 \left|\sin\left(\frac{\theta}{2}\right)\right|\right]^2.$$

Аналогичным образом можно показать, что для схемы с реквантователем NS  $\zeta$  порядка

$$|T_e(z)_\zeta| = \left[2 \left|\sin\left(\frac{\theta}{2}\right)\right|\right]^\zeta,$$

а спектральная плотность мощности шума квантования определяется равенством

$$S(e)_{ns\zeta} = S(e) \cdot \left(4 \left|\sin^2\left(\frac{\theta}{2}\right)\right|\right)^\zeta.$$

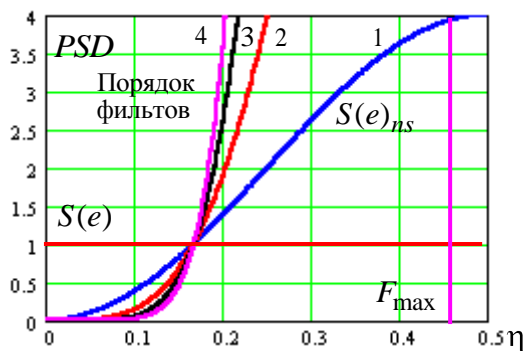


Рис.5.14. Графики спектральной плотности шума квантования

Из этих формул и графиков на [рис.5.14](#) следует, что с повышением порядка фильтра реквантователя резко уменьшается спектральная плотность шума квантования ниже  $\eta = 0,16$  и увеличивается скорость ее нарастания выше этой частоты. Максимальное значение спектральной плотности мощности на частоте Найквиста для фильтра  $\zeta$  порядка

$$S(e)_{ns\zeta_{\max}} = S(e) \cdot 4^\zeta,$$

а ухудшение  $SNR$  может быть рассчитано по формуле

$$\Delta SNR_{\zeta} = -10 \lg \left( \int_0^{\theta_1} \left( 4 \left| \sin^2 \left( \frac{\theta}{2} \right) \right| \right)^{\zeta} d\theta \right) .$$

Если для реквантователя  $NS$  1 порядка это ухудшение составляет минус 3 дБ, то для реквантователя 5 порядка - уже минус 24 дБ.

Применение фильтров более 3 порядка с последовательным включением элементов задержки приводит к нестабильности работы реквантователя, поэтому они реализуются в виде нескольких ступеней с применением местных обратных связей. Причем все связи имеют собственные весовые коэффициенты.

*В практическом приложении* технология  $NS$  всегда осуществляется одновременно с технологией *Dithering*, поэтому все преимущества последней, описанные выше, естественно сохраняются. Новым является добавление фильтра верхних частот. Это позволяет понизить спектральную плотность мощности шума реквантования в диапазоне от 0 до 14...16 кГц, за счет резкого увеличения мощности шума реквантования на более высоких частотах, до частоты Найквиста.

В результате в области средних частот  $SNR$  увеличивается и динамический диапазон расширяется. При этом общее значение  $SNR$  резко возрастает, но высокие частоты человек плохо воспринимает или вообще не слышит, поэтому громкость шума квантования понижается, а качество звучания музыки улучшается. Во всех современных звуковых картах для реквантования на основе технологии  $NS$  используются достаточно сложные фильтры высоких порядков, при построении которых учитываются особенности слуха человека в отношении слышимости шума.

Все расчеты по громкости шума технологии  $NS$  такие же, как это изложено в разделе, поэтому приведем только один пример, показывающий различие между применением только технологии *Dithering* и одновременным применением двух технологий. Для этих целей был использован звуковой редактор *Steinberg WaveLab*..

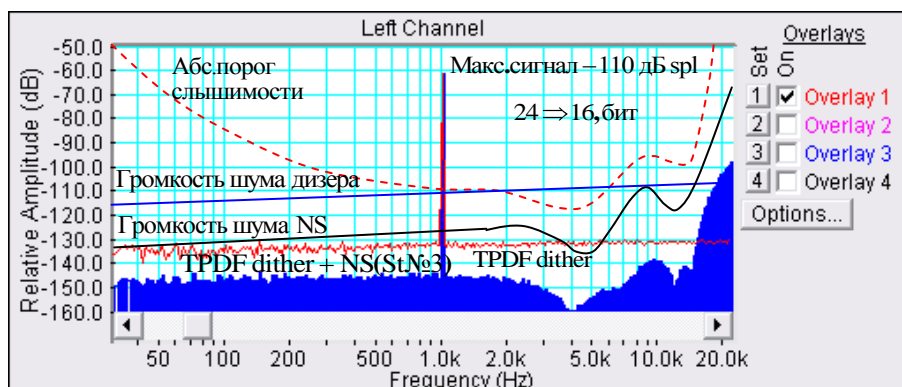


Рис.5.15. К вопросу слышимости шума дизера и шума технологии  $NS$

который позволяет получить лучшие результаты. Никакой информации об используемых в этом редакторе фильтрах нет, поэтому результат можно получить только методом проб и ошибок.

В этом редакторе производилось реквантование  $24 \Rightarrow 16$ , бит сигнала с частотой 1000 Гц и уровнем минус 60 дБ. Технология *Dithering* выполнялась с использованием шума типа 1, (предполагаем, что это *TPDF Dither*). Технология  $NS$  выполнялась с использованием шума типа 1 и  $NS$  типа 3.

На **рис.5.15.** приведены графики спектральной плотности мощности шума квантования при использовании технологии *TPDF Dither* (график красного цвета) и *TPDF Dither + NS* (график темно-синий). На этом же рисунке приведен график абсолютного порога слышимости.

В логарифмической шкале частот график частотной зависимости громкости шума реквантования при использовании только технологии достаточно точно представляется

прямой линией с крутизной подъема 3 дБ/октаву (рис.5.15.). На графике это синяя линия с подъемом в области высоких частот, она характеризует громкость шума реквантования при использовании только технологии *TPDF Dither*.

График громкости шума квантования при использовании технологии *NS* также рассчитывается по его спектральной плотности мощности с использованием корректирующих коэффициентов, о которых говорилось выше. Как видно, из рис.5.15., этот график лежит ниже графика абсолютного порога слышимости на 10...20 дБ во всем звуковом диапазоне и, поэтому шум реквантования не слышен. Он не будет слышен и при повышении  $L_{\max} SPL$  до 120 дБ.

### 5.3. Технология Super Bit Mapping (SBM)

Эта технология, предложенная фирмой *Sony*, предназначена для изготовления современных компакт дисков. В настоящее время в студийных условиях в процессе изготовления мастер - диска нет особых проблем произвести запись звука с 20 и более разрядами. Однако, стандартом *CD* предусмотрена возможность использования только 16 бит, поэтому имеется возможность произвести округление 20 разрядных слов оптимальным образом так, чтобы звучание при 16 разрядном коде было почти такое же, как при 20 разрядном.

Основой технологии *SBM* является применение реквантователя *NS*, который формирует огибающую спектра шума квантования в виде *F*- взвешенной кривой равной громкости (рис.5.16). Этим достигается распределение шума квантования по всему диапазону от нуля до частоты Найквиста. При этом спектральная плотность мощности

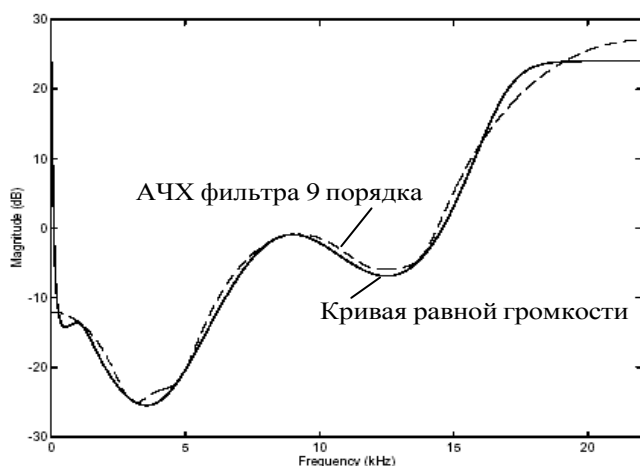


Рис.5.16. Кривая равной громкости и АЧХ фильтра 9-го порядка

шума понижается в области максимальной чувствительности слуха от 1 до 6 кГц и существенно увеличивается на участках от 0 до 60 Гц и выше 15 кГц, где абсолютный порог слышимости резко падает. Достигается это при использовании в реквантователе фильтра 9-го порядка.

В технологии *SBM* для уменьшения слышимости шума квантования используется еще эффект частотной маскировки шума квантования ЗС. С этой целью звуковой диапазон разбивается на полосы, близкие к критическим. В каждой полосе осуществляется

непрерывное Фурье-преобразование и рассчитывается кривая частотной маскировки. В соответствии с этими расчетами непрерывно меняется передаточная функция в цепи отрицательной обратной связи квантователя  $H(z)$ .

При больших уровнях ЗС, когда шум квантования маскируется ЗС, преобразование спектра шума квантования почти не используется. Когда же уровень ЗС мал, эффект маскировки не работает и тогда в действие вступает преобразователь спектра. Технология *SBM* наиболее эффективна при воспроизведении тихой музыки, когда ошибки квантования наиболее заметны.

На рис.5.17. приведена эквивалентная схема такого реквантователя. В этой схеме на вход сумматора подаются 20- разрядные выборки  $U(j)_{in}$  с частотой дискретизации  $f_s$ . В переквантователе с этой же частотой производится грубое округление, при котором младшие 4 разряда отбрасываются, и на выходе формируется 16-разрядный сигнал

$U(j)_{out}$ . Путем сравнения входа и выхода квантователя формируется сигнал ошибки округления  $e(j)$ , который подается на входной сумматор. Отрицательная обратная связь стремится сделать как можно меньшей разницу между входным и выходным сигналами и этим приближает 16-разрядный сигнал к 20-разрядному. В цепи отрицательной обратной связи включен фильтр 9 порядка с передаточной функцией по частоте  $H(f)$ , приведенной на рис.5.16., которая определяет форму огибающей спектра шума квантования. Для расширения динамического диапазона во время квантования осуществляется операция *Dithering*.

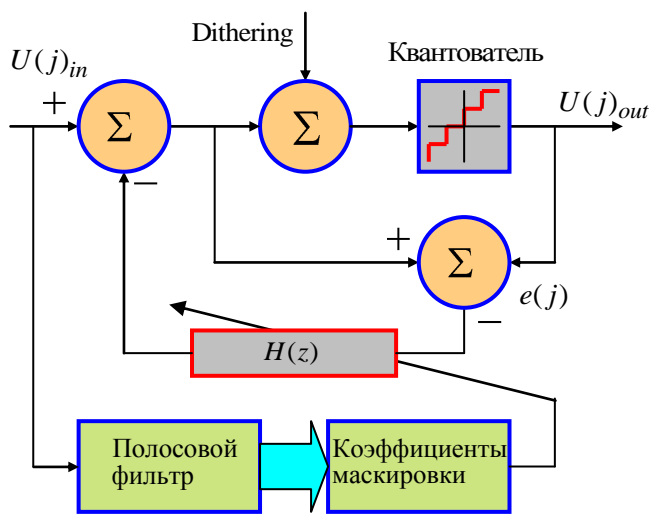


Рис.5.17. Реквантователь NS с психоакустической обратной связью

Применение технологии *SBM* отражает одно из направлений развития современной цифровой звукотехники, в соответствии с которым считается, что для обеспечения высококачественного звучания вполне достаточно 16 разрядов и частоты дискретизации 44,1 кГц. Очевидная недостаточность такого числа разрядов вполне может быть компенсирована использованием

специальных технологий, обеспечивающих отношение сигнал/шум такое же, как и при 20-разрядном кодировании, когда шумы квантования находятся ниже порога слышимости.

Именно поэтому в настоящее время разработаны и широко применяются технологии подобные *SBM*. Из них наиболее известны: *Apogee UV22*, *Gambit*, *Harmonia Mundi* и *Sonic Solution Turbo Bit Mapping*. Эти технологии отличаются тем, что их применение удорожает только студийную аппаратуру, тогда как запись музыки может воспроизводиться простейшими плеерами.

Другим направлением развития является увеличение числа разрядов до 20 и даже до 24 с повышением частоты дискретизации до 96 или 192 кГц. В этом случае никакие технологии *Dithering*, *Noise Shaping*, *SBM* не требуются и проблемы решаются применением носителей с высокой плотностью записи, таких как *DVD-Audio* или *Blue Ray Disc*.

Другим направлением развития является увеличение числа разрядов до 20 и даже до 24 с повышением частоты дискретизации до 96 или 192 кГц. В этом случае никакие технологии *Dithering*, *Noise Shaping*, *SBM* не требуются и проблемы решаются применением носителей с высокой плотностью записи, таких как *DVD-Audio* или *Blue Ray Disc*.

#### 5.4. Технология *Oversampling*

*Передискретизация* - это способ уменьшения шума квантования путем повышения частоты выборок в несколько раз по сравнению с базовой частотой из стандартного ряда, 44,1 и 48 кГц. Коэффициент передискретизации  $K_{os}$  показывает во сколько раз повышается частота дискретизации. Этот коэффициент определяется через функцию  $2^x$

$$f_{ks} = f_s \cdot K_{os}, \quad K_{os} = 2^x, \quad 2 \leq x \leq 10$$

Передискретизация может быть аналоговой и цифровой.

*Аналоговая передискретизация* отличается только тем, что выборки делаются с повышенной частотой дискретизации  $f_{sk}$  и применяется она только в АЦП. Эти идеи используются в системе *DVD-Audio*, где частота дискретизации может быть 96 и 192 кГц, в системе *Super Audio CD* частота дискретизации еще намного выше и равна 2,88224 МГц.

Когда рассматриваются вопросы передискретизации, всегда предполагается, что в тракте используется технология «*Dithering*», осуществляющая декорреляцию ошибок квантования, при которой детерминированные ошибки квантования преобразуются в шум

квантования с равномерной спектральной плотностью от 0 до частоты Найквиста  $f_N$  (рис.5.18).

При равномерном законе распределения плотность вероятности шума квантования определяется равенством

$$P_D(e) = 1/Q,$$

и его эффективное значение рассчитывается по формуле

$$\bar{e}(z_q = Q/2) = \sqrt{\int_{-Q/2}^{Q/2} e^2 \cdot P_D(e) \cdot d(e)} = \frac{Q}{2\sqrt{3}},$$

Модуль спектральной плотности мощности шума квантования  $S_D(e)$  зависит от мощности шума квантования и частоты Найквиста  $f_N$

$$|S_D(e)| = \frac{Q^2}{12 \cdot f_N}.$$

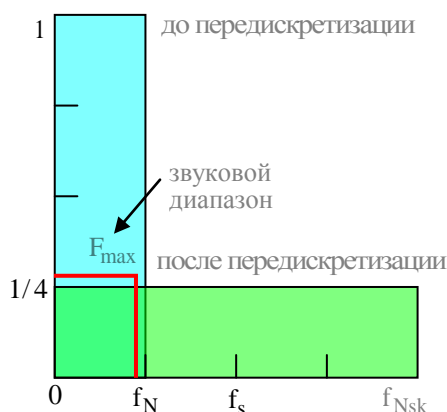


Рис.5.18. Спектр шума квантования

Значение SNR в цифровых звуковых трактах рассчитывается как отношение максимального эффективного синусоидального напряжения на выходе ФНЧ ИКМ демодулятора  $\bar{A}_{max}$  к эффективному значению напряжения шума квантования  $\bar{e}$

$$SNR = 20 \cdot \lg \left( \frac{\bar{A}_{max}}{\bar{e}} \right), \text{ где } \bar{A}_{max} = \frac{Q \cdot 2^{(q-1)}}{\sqrt{2}},$$

поэтому

$$SNR = 6,02 \cdot q + 1,76; \quad q > 1, \text{ дБ.}$$

Повышение частоты дискретизации приводит к увеличению частоты Найквиста, и, как следствие, к расширению полосы частот шума квантования до  $f_{Nsk}$  и уменьшению модуля спектральной плотности мощности, определяемого равенством

$$|S_D(e)| = \frac{Q^2}{12 \cdot f_{Nsk}}.$$

Поэтому расчетное соотношение для SNR преобразуется к виду

$$SNR = 6,02 \cdot q + 1,74 + 10 \cdot \lg \left( \frac{f_s \cdot K_{os}}{2F_{max}} \right),$$

дБ,

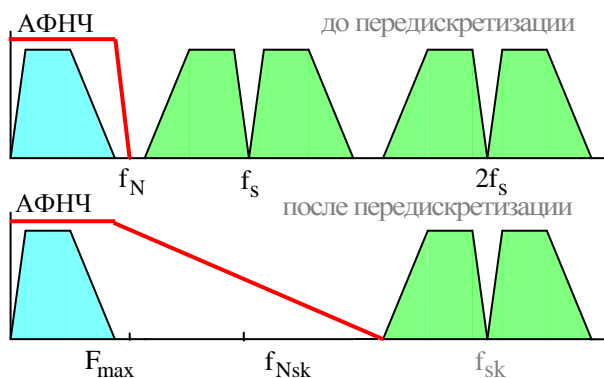


Рис.5.19. Спектры АИМ при передискретизации

где  $F_{max}$  - максимальная частота звукового диапазона.

Из приведенной формулы следует, что SNR увеличивается на 3 дБ при каждом удвоении частоты дискретизации. Это объясняется тем, что при увеличении частоты дискретизации спектр шума квантования расширяется и во столько же раз его спектральная плотность мощности

уменьшается.

Относительно небольшое увеличение SNR сопровождается двукратным увеличением скорости цифрового потока и необходимостью двойного увеличения плотности записи. Поэтому при достаточно высокой частоте передискретизации возникают серьезные

технические проблемы в ее реализации. По этой причине в ИКМ трактах при аналоговой передискретизации в модуляторе коэффициент  $K_{os}$  равен всего 2 или 4.

Использование аналоговой передискретизации позволяет в ИКМ трактах значительно упростить антиэлайзинговый аналоговый ФНЧ на входе модулятора, который предназначен для исключения возможности перекрытия спектров полезного сигнала и продуктов модуляции. Реализация такого фильтра в обычных ИКМ трактах весьма сложна, так как у него должна быть линейная АЧХ в рабочем диапазоне и крутой спад вблизи частоты Найквиста с затуханием не менее 90 дБ. При использовании передискретизации требования к крутизне спада этого АФНЧ существенно уменьшаются, так как существенно увеличивается частота Найквиста (рис.5.19), и в тоже время обеспечивается отсутствие комбинационных частот с участием ЗС.

Аналоговая передискретизация позволяет также значительно снизить требования к ФНЧ ИКМ демодулятора. Если, например, частота дискретизации равна 48 кГц, то в обычном тракте необходимо подавить частоты нижней боковой полосы 1 порядка выше 24 кГц. Это может быть выполнено лишь весьма сложным ФНЧ 7...11 порядка. При 2-х

кратной передискретизации граница нижней боковой полосы повышается до  $1,5 f_{Nsk} = 72$  кГц (рис.5.19) и проблем с созданием ФНЧ не возникает.

При высокой частоте дискретизации целесообразно в ИКМ трактах частоту среза ФНЧ на выходе демодулятора увеличивать до 25...40 кГц. При этом SNR становится меньше, но зато расширяется полоса звуковых частот. Звучание становится более прозрачным, за счет уменьшения частотно-фазовых искажений на верхней границе звукового диапазона.

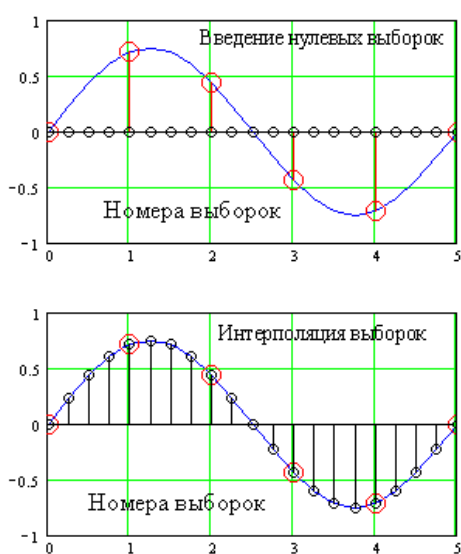
В некоторых публикациях по этому поводу авторы полагают, что можно использовать на выходе демодулятора ИКМ простые ФНЧ, как, например, фильтр Чебышева 3 порядка. Крутизна среза у них мала, но они обеспечивают хорошее подавление высокочастотных составляющих спектра модуляции. При этом забывается, что ФНЧ предназначен еще для накопления и

интерполяции выборок при реконструкции ЗС, поэтому его постоянная времени не может быть малой, иначе возникнут значительные нелинейные искажения.

**Цифровая передискретизация** прежде всего используется в ЦАП на основе сигма-дельта модуляции с коэффициентом  $K_{os}$  до 128 и даже 1024 крат, так как при этом отсутствуют проблемы, связанные со скоростью цифрового потока и плотностью записи. Надо обратить внимание на то, что передискретизация увеличивает коэффициент корреляции между последовательными отсчетами. Это обеспечивает большую точность работы интерполяционных фильтров. По этой причине операция декорреляции ошибок квантования производится после фильтрации.

Существуют *два способа цифровой передискретизации* (рис.5.20). В первом варианте между уже существующими цифровыми выборками вводятся дополнительные, рассчитанные путем интерполяции. Другой способ получения значений промежуточных значений выборок состоит во вставке нулевых выборок, после чего вся последовательность подвергается цифровой фильтрации.

Цифровые фильтры с предварительным расчетом интерполированных выборок имеют хорошие технические характеристики, но для них требуется значительная вычислительная мощность и они дороги. В CD проигрывателях с такими фильтрами присутствует надпись



Цифровая 4-х кратная передискретизация

Рис.5.20. Цифровая передискретизация

«18 bit 20×Oversampling». Это значит, что используется 20-ти кратная передискретизация, при которой  $SNR$  соответствует 18 разрядному кодированию.

Цифровые фильтры на основе введения нулевых выборок значительно проще и дешевле. Однако, при фильтрации в них возникают специфические помехи, которые могут достигать величины сигнала. Эти помехи большей частью фильтруются, но некоторая часть попадает в звуковой диапазон.

Базовые цифровые фильтры выпускаются с коэффициентом передискретизации 2, 4 и 8 крат. При необходимости они могут включаться последовательно с максимальным общим коэффициентом передискретизации до 1024. Число разрядов на выходах фильтров может быть от 17 до 28, поэтому требуется их усечение.

При использовании цифровой передискретизации в ЦАПх ИКМ трактов полностью справедливы приведенные выше формулы для  $SNR$ . Из этих формул следует, что передискретизация позволяет использовать ЦАПы с меньшим числом разрядов без ухудшения  $SNR$ . При этом каждое уменьшение кодового слова на один разряд (6 дБ) может компенсироваться двукратным повышением частоты дискретизации. Например, для того чтобы вместо 16-ти разрядного ЦАПа использовать более дешевый 10-ти разрядный ЦАП, частоту дискретизации следует увеличить в 64 раза. При этом потребуется ЦАП с быстродействием во столько же раз большим, что обойдется еще дороже. Поэтому такой вариант использования передискретизации не нашел широкого практического применения.



## 6. Сигма-дельта модуляция

Сигма-дельта модуляция предназначена для аналого-цифрового и цифро-аналогового преобразований звуковых сигналов. В отличие от импульсно-кодовой модуляции она позволяет использовать при этих операциях достаточно грубые преобразователи с числом разрядов вплоть до одного, обеспечивая при этом отношение сигнал шум до 120...140 дБ, что необходимо для профессиональной записи звука. Технология производства АЦП и ЦАП на основе сигма-дельта модуляции значительно проще и дешевле, поэтому такие преобразователи широко используются в современных цифровых магнитофонах, в оптической звукозаписи и звуковых картах компьютеров.

В отличие от ИКМ АЦП и ЦАП на основе сигма-дельта модуляции работают на частоте дискретизации в 4 и более раз выше стандартного значения, соответствующего требованиям теоремы В.П. Котельникова. В них используются грубые квантователи с числом разрядов от 1 до 6 с частотно-зависимой отрицательной обратной связью. Квантование уже квантованного цифрового сигнала в ЦАП на основе сигма-дельта модуляции называют реквантизацией.

### 6.1. Технология Noise Shaping

Основой сигма – дельта модуляции является так называемая технология Noise Shaping, при которой под действием частотно-зависимой отрицательной обратной связи изменяется вид огибающей спектральной плотности мощности шума квантования таким образом, что в звуковом диапазоне она существенно уменьшается, за счет увеличения ее за пределами этого диапазона. Технология Noise Shaping включает в себя операции Dithering, с помощью которой осуществляется декорреляция ошибок квантования и

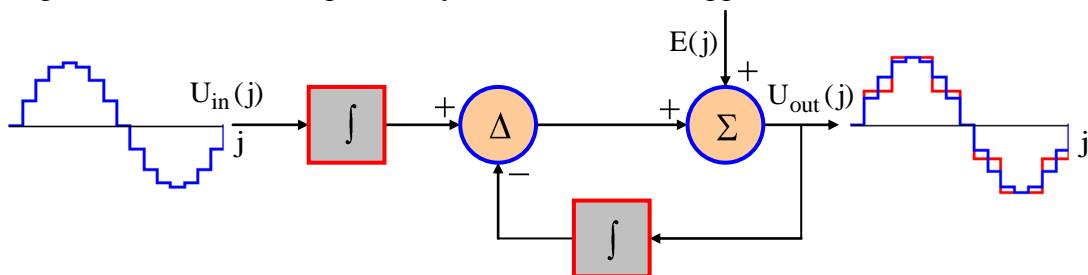


Рис.6.1. Квантование (реквантование) с отрицательной обратной связью

преобразование их в шум с равномерной спектральной плотностью, и аналоговую или цифровую передискретизации.

Реализация технологии Noise Shaping поясняется рис.6.1., в котором квантователь (реквантователь), представлен сумматором, на один вход которого подается квантуемый сигнал, а на второй - ошибки или шум квантования, соответствующий числу используемых двоичных разрядов. Для преобразования спектра ошибок квантования без изменения спектра квантуемого сигнала в предложенной модели используется отрицательная обратная связь, два интегратора и вычислитель разности между входным сигналом и сигналом обратной связи. Чтобы эта модель была справедлива как при квантовании аналоговых так и реквантовании цифровых сигналов, входной и выходной сигналы представляются в аналоговой форме в виде последовательности выборок с частотой дискретизации. В качестве этих выборок может быть сигнал на выходе устройства выборки и хранения после амплитудно-импульсной модуляции или уже квантованный цифровой сигнал, но в аналоговой форме представления.

Благодаря применению двух одинаковых интеграторов, один на входе и другой в цепи обратной связи, обеспечивается частотная независимость передачи квантуемого сигнала с входа на выход, так как разностный сигнал на входе сумматора в этом случае на всех частотах равен нулю. Сигнал ошибки квантования проходит только через один

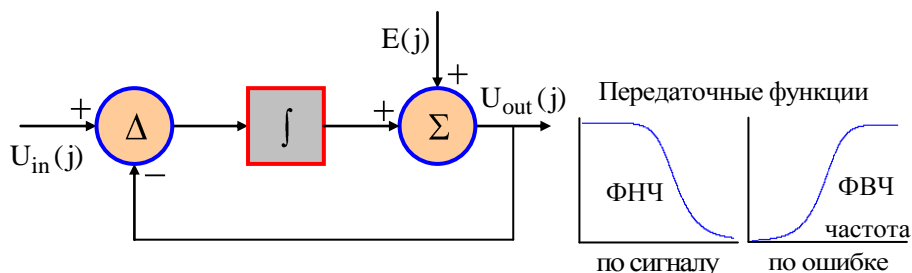


Рис.6.2. Эквивалентная схема формирователя спектра

интегратор, поэтому возникает частотная зависимость его коэффициента передачи. Поэтому приведенную схему называют формирователем спектра шума квантования.

Путем линейных преобразований схема на рис.6.2 может быть приведена к виду с одним интегратором в прямой ветви. В этой схеме интегратор имеет передаточную функцию

$$T(s)_{\text{int}} = \frac{1}{s},$$

где  $s = j2\pi F\tau_i$  - оператор преобразования Лапласа,  $\tau_i$  - постоянная времени интегратора.

Передаточная функция для сигнала

$$T(s)_s = \frac{Y(s)}{X(s)} = \frac{1}{1+s}, \text{ (когда } N(s) = 0 \text{)}$$

такая же как у фильтра нижних частот (ФНЧ) первого порядка. Передаточная функция для ошибки квантования

$$T(s)_n = \frac{Y(s)}{N(s)} = \frac{s}{1+s}, \text{ (когда } X(s) = 0 \text{)}$$

имеет такой же вид как у фильтра верхних частот (ФВЧ) первого порядка. Таким образом, приведенной схеме в звуковом диапазоне частот квантуемый сигнал передается без изменений, а спектральная плотность ошибок квантования понижается на низких частотах и повышается на высоких. Отрицательная обратная связь стремится уравнивать выходной сигнал с входным.

## 6.2. Математические модели сигма-дельта модулятора

В цифровой схемотехнике работа всех узлов тактируется и для их анализа используются не преобразования Лапласа, а  $z$ -преобразования, поэтому эквивалентная схема формирователя спектра шума квантования с интегратором 1 порядка несколько изменяется (рис.6.3). Для ее работы необходимо, чтобы сигнал обратной связи был сдвинут по времени на один такт по отношению к входному. Предполагается также, что ошибки квантования декоррелированы и спектр их равномерен от 0 до частоты Найквиста, а на входе действует сигнал после передискретизации с коэффициентом передискретизации  $K_{os} = 2^x$ ,  $x = 0, 1, \dots, 10$ .

В приведенной схеме используется дискретно-временной интегратор 1 порядка с собственной петлей обратной связи. У него единичный коэффициент усиления и он осуществляет внутреннюю задержку на один такт:  $\tau = 1/f_{sk}$ . В схеме интегратора функция  $z^{-1}$  является  $z$ -оператором временной задержки, сумматор - выполняет функции аналогового накопителя в дискретной форме. Передаточная функция такого интегратора в форме  $z$ -преобразования имеет вид

$$H(z) = \frac{z^{-1}}{1+z^{-1}},$$

где  $z = e^{i\theta}$ ,  $\theta = \frac{2\pi f}{f_{sk}}$ ,  $i = \sqrt{-1}$ ,  $i$  – мнимая единица,  $0 < \theta < \pi$ ,  $f$  – текущая частота,  $\theta$  – нормализованная частота,  $f_s$  – частота используется другая модель

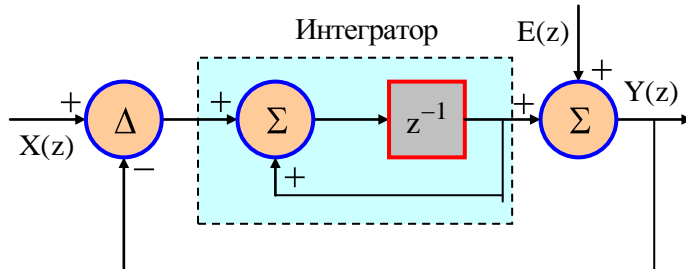


Рис.6.3. Модель  $\Sigma\Delta$  модулятора 1 порядка

В технической литературе схему формирователя спектра ошибок квантования, часто называют сигма-дельта модулятором ( $\Sigma\Delta$  или  $\Sigma\Delta$ ). При этом уточняют, что он аналоговый, если производится квантование, или цифровой, если производится реквантование.

Вместо схемы на рис.6.3. часто используется другая модель  $\Sigma\Delta$  1 порядка с задержкой в цепи обратной связи (рис.6.4.). В этих схемах сигнал на выходе модулятора в функции дискретного времени можно представить в виде равенства

$$Y(j)_{out1} = X(j) + e(j) - e(j-1),$$

из которого следует, что мгновенная ошибка квантования определяется как

$$e(j)_{SDM} = e(j) - e(j-1).$$

Эта формула отражает основную суть  $\Sigma\Delta$ -модуляции. Очевидно, что ошибки двух последовательных выборок на низких частотах почти не отличаются и разностная ошибка стремится к нулю. На высоких частотах скорость изменения ЗС большая и эти ошибки могут отличаться очень значительно, они могут иметь и разную полярность, поэтому суммарная ошибка квантования сильно возрастает.

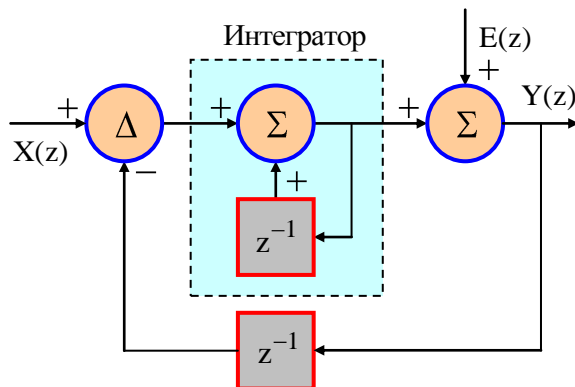


Рис.6.4. Модель  $\Sigma\Delta$  модулятора 1 порядка

В приведенной схеме выходной сигнал в форме  $z$ -преобразования имеет вид

$$Y(z)_{out1} = z^{-1} \cdot X(z)_{in} + (1 - z^{-1}) \cdot e(z).$$

Из этой формулы следует, что модули коэффициентов передачи по сигналу и ошибке

соответственно равны :

$$|T_x(z)| = |z^{-1}| = 1, \quad |T_e(z)_1| = |1 - z^{-1}| = 2 \left| \sin\left(\frac{\theta}{2}\right) \right|$$

Это значит, что при использовании интегратора 1 порядка квантуемый сигнал проходит на выход модулятора без частотных искажений и только задерживается на один такт, а ошибка квантования является нелинейной функцией частоты.

На рис.6.5 приведена модель  $\Sigma\Delta$ -модулятора второго порядка с последовательным включением двух интеграторов 1 порядка. Сигнал на выходе этого модулятора в функции дискретного времени можно представить в виде

$$Y(j)_{out2} = X(j)_{in} + e(j) - 2e(j-1) + e(j-2).$$

Из этой формулы следует, что суммарная ошибка квантования определяется алгебраической суммой 4 значений ошибок квантования, сдвинутых по времени.

Выходной сигнал в форме  $z$ -преобразования

$$Y(z)_{out2} = z^{-1} \cdot X(z)_{in} + (1 - z^{-1})^2 \cdot e(z).$$

Поэтому модуль передаточной функции ошибки квантования равен

$$|T_e(z)_2| = |(1 - z^{-1})^2| = \left[ 2 \left| \sin\left(\frac{\theta}{2}\right) \right| \right]^2$$

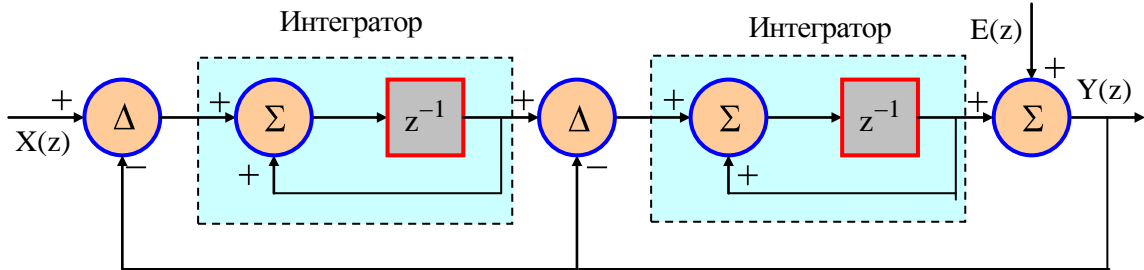


Рис. 6.5. Модель  $\Sigma\Delta$  модулятора 2 порядка

Аналогичным образом можно показать, что для схемы с интегратором  $\zeta$  порядка передаточная функция по ошибке определяется равенством

$$|T_e(z)_\zeta| = [2|\sin(\theta/2)|]^\zeta.$$

Наиболее важной характеристикой технологии Noise Shaping является форма огибающей спектральной плотности мощности шума квантования (PSD). На выходе модулятора  $\zeta$  порядка она может быть рассчитана по формуле

$$S_D(e)_\zeta = S_D(e) \cdot |T_e(z)_\zeta|^2 = S_D(e) \cdot \left( 4 \left| \sin^2\left(\frac{\theta}{2}\right) \right| \right)^\zeta,$$

где  $S_D(e)$  - PSD шума квантования квантователя (реквантователя), определяемая числом используемых разрядов и передискретизацией без учета влияния отрицательной обратной связи, это характеристика шума, подаваемого в эквивалентной схеме на сумматор.

Графики на рис.6.6 иллюстрируют форму огибающей спектра ошибок квантования на выходе модулятора с интеграторами 1 и 2 порядков при  $K_{os} = 4$ . Для сравнения приведен также график спектральной плотности  $S_D(e)$ , характеризующей шум непосредственно квантователя. Как видно, в звуковом диапазоне SPD уменьшается и тем сильнее, чем выше порядок интегратора.

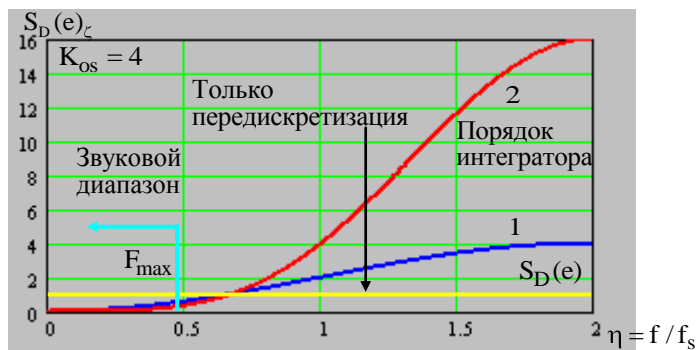


Рис.6.6. Графики огибающей спектральной плотности мощности шума квантования SDM 1 и 2 порядков

в звуковом диапазоне и увеличиваться на частоте Найквиста. С увеличением  $K_{os}$  форма графиков остается неизменной, а меняется масштаб по оси частот, так как увеличивается частота Найквиста. При этом звуковой диапазон, ограниченный относительной частотой  $\eta = 0,5$ , на графике будет сужаться обратно пропорционально значению  $K_{os}$  и спектральная плотность шума будет в нем быстро уменьшаться.

Основной характеристикой цифрового тракта звукозаписи является отношение сигнал\шум квантования (SNR) и в общем случае, когда используются все технологии уменьшения шума квантования, он определяется равенством.

$$SNR_{qk} = (6,02 \cdot q + 1,76) + 10 \cdot \lg \left( \frac{K_{os} f_s}{2F_{max}} \right) - 10 \lg \left( \int_0^{\theta_1} \left( 4 \left| \sin^2 \left( \frac{\theta}{2} \right) \right| \right)^\zeta d\theta \right), q > 1,$$

где

$$\theta_1 = \frac{2\pi F_{max}}{K_{os} f_s},$$

$F_{max}$  – максимальная частота звукового диапазона.

Первый член этого равенства определяет зависимость SNR от числа разрядов, используемых при квантовании (реквантовании). Формула действительна только при  $q > 1$ . Если  $q = 1$  (1 бит SDM), то первый член равенства становится отрицательным и равен – 8,45 дБ. Второй член равенства определяет зависимость SNR от передискретизации, он увеличивается на 3 дБ при каждом удвоении частоты дискретизации. Третий член равенства определяет зависимость SNR от порядка SDM. При каждом удвоении частоты дискретизации он увеличивается на 6 дБ для SDM 1 порядка, на 12 дБ для SDM 2 порядка и на 18 дБ для SDM 3 порядка и так далее.

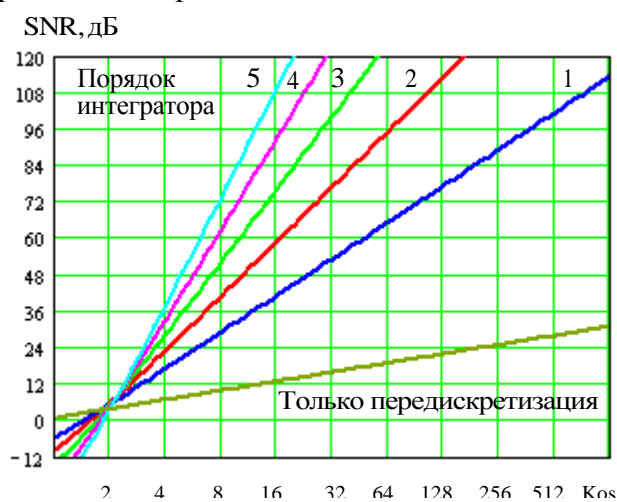


Рис.6.7. Графики SNR для SDM 1-5

дискретизации он увеличивается на 6 дБ для SDM 1 порядка, на 12 дБ для SDM 2 порядка и на 18 дБ для SDM 3 порядка и так далее.

Интеграл в формуле табличный, но аналитические выражения него очень громоздки. Известно много приближенных значений этого интеграла, но более рационально производить расчеты на ПК без использования эмпирических формул, так как они дают большую погрешность.

Графики на рис.6.7 иллюстрируют зависимость SNR от порядка SDM и коэффициента  $K_{os}$  при  $q = 2$ . Из них следует, что при отсутствии

передискретизации и  $K_{os} < 2$  технология Noise Shaping только ухудшает SNR и тем в большей степени, чем выше порядок интегратора. Если  $K_{os} > 2$ , то чем выше порядок модулятора, тем при меньшем значении  $K_{os}$  достигается необходимое для цифрового тракта значение SNR = 120 дБ, даже при использовании весьма грубого квантования ( $q = 2$ ).

### 6.3. Одноразрядный АЦП на основе сигма-дельта модуляции

В последние годы основной интерес вызывают одноразрядные АЦП на основе  $\Sigma\Delta$ -модуляции (1-бит SDM). Они очень успешно конкурируют с АЦП на основе ИКМ, особенно в цифровых магнитофонах, так как отличаются значительно более простой технологией изготовления, существенно дешевле, а по качественным характеристикам примерно равноценны.

Наиболее важным вопросом построения АЦП на основе сигма-дельта модуляции является выбор частоты дискретизации. Чем эта частота больше, тем при меньшем значении порядка интегратора достигается необходимое значение SNR 120...140 дБ. Однако, с повышением частоты дискретизации увеличивается и скорость цифрового потока, что приводит к уменьшению времени записи на дисковом носителе. Поэтому приходится искать компромисс между значениями частоты дискретизации и порядком

интегратора. В настоящее время в качестве такого компромиссного решения принят коэффициент передискретизации равным 64. Из графиков на рис.6.7. видно, что при таком значении  $K_{os}$  отношение сигнал шум 120 дБ достигается только при использовании

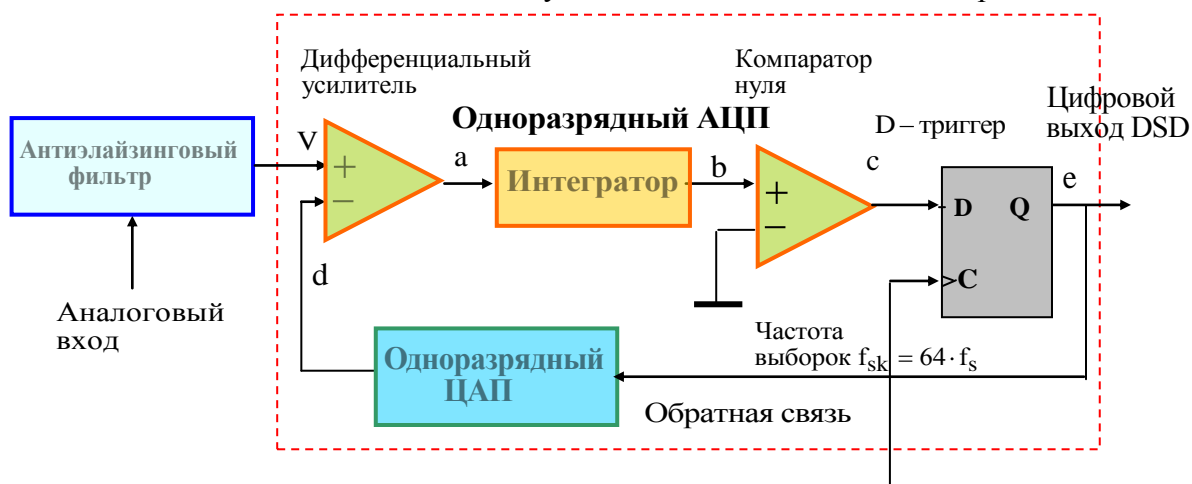


Рис.6.8. Функциональная схема сигма-дельта модулятора первого порядка

SDM не менее 3 порядка.

На рис.6.8. приведена упрощенная схема 1-бит SDM первого порядка. В этой схеме используется аналоговая передискретизация, поэтому частота дискретизации выбирается в  $2^x$  выше одной из стандартных частот 44,1 или 48 кГц. При  $K_{os} = 64$  и  $f_s = 48$  кГц  $f_{sk} = 3,072$  МГц. Аналоговый сигнал подается на вход модулятора через антиэлайзинговый фильтр, в качестве которого может использоваться простейший RC-интегратор, так как требуется подавлять частоту Найквиста (1,5 МГц) и выше.

В этой схеме нет классического дискретизатора, выполняющего функцию амплитудно-импульсной модуляции, отсутствует устройство выборки и хранения, но работа всех узлов тактируется частотой дискретизации. В качестве одноразрядного квантователя используется компаратор, который формирует выходной сигнал положительной полярности только при условии, что входное напряжение выше нуля. В приведенной схеме задержка на один такт осуществляется D- триггером, на счетный вход которого подается сигнал с частотой дискретизации  $f_{sk}$ .

Этот же триггер выполняет функцию дискретизатора. В таком варианте исполнения при аналого-цифровом преобразовании сначала производится квантование, а потом дискретизация. Одноразрядный ЦАП преобразует однополярный выходной сигнал D-триггера в двух полярный.

В соответствии с приведенным алгоритмом в начале каждого такта дифференциальный усилитель вырабатывает на своем выходе разностный сигнал “a” между входным напряжением “V” и выходным напряжением одноразрядного ЦАП. Интегратор добавляет напряжение “a” к своему выходному напряжению, сформированном в предыдущем такте. Это новое напряжение “b” подается на вход компаратора нуля. На выходе компаратора формируется логическая 1, если  $b > 0$  и логический 0, если  $b < 0$ . Временная диаграмма, поясняющая работу всех узлов АЦП при напряжении на входе +0,6 В приведена на рис. 6.9.

Выходной сигнал компаратора, называемый DSD (Direct Stream Digital), однополярный в виде непрерывной последовательности логических 1 и 0 без разделения их на символы и блоки. По цепи обратной связи этот цифровой поток формирует выходной сигнал одноразрядного ЦАПа. Этот сигнал сдвинут по времени на один такт. Если на выходе компаратора логическая единица формируется сигнал импульсный сигнал положительной полярности, а если 0, то отрицательной. Амплитуды сигналов на выходе

Алгоритм работы :

$$a_{j+1} = V_{j+1} - d_{j+1}$$

$$b_{j+1} = b_j + a_{j+1}$$

$$c_{j+1} = \begin{cases} 1; & b_{j+1} > 0 \\ 0; & b_{j+1} \leq 0 \end{cases}$$

$$d_{j+1} = \begin{cases} 1; & c_j > 0 \\ -1; & c_j = 0 \end{cases}$$



Одноразрядные SDM очень чувствительны к перегрузкам, поэтому максимальный размах сигнала его на входе на должен превышать  $0.84Q$ . При превышении этого уровня модулятор переходит в режим самовозбуждения. Ошибки квантования в одноразрядных SDM сильно коррелированы с сигналом, поэтому без операции «dithering» их использовать нельзя. В тоже время добавление шума с размахом  $0.8Q$  приводит к значительному уменьшению SNR и существенному уменьшению допустимого уровня входного сигнала по размаху до значения  $0.5Q$ .

По этим причинам в таких модуляторах невозможно использовать шум с треугольным законом распределения, как это делается в ИКМ трактах, и поэтому применяют шум с *прямоугольным законом распределения* в пределах  $0.5Q$ . Такой шум лишь улучшает линейной передаточной функции, но происходит это за счет **уменьшения SNR на 12 дБ** и ухудшения стабильности работы модулятора.

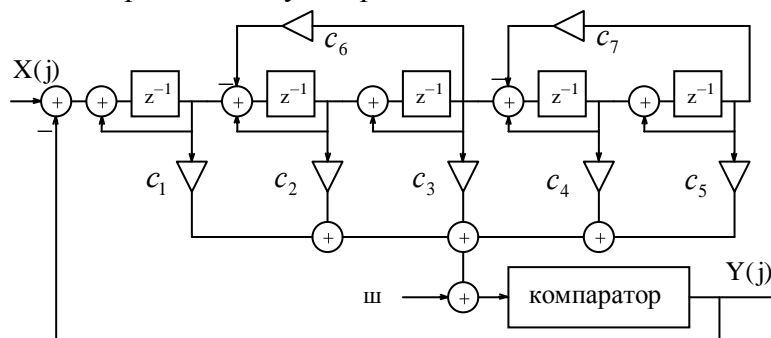


Рис.6.11. Схема одноразрядного SDM 5 порядка

Операция декорреляции ошибок квантования в 1-бит SDM не преобразует их в аддитивные, поэтому особого положительного эффекта она не приносит. Если декорреляция ошибок не используется, то в звуковом диапазоне уровень шума уменьшается до минус 130 дБ, добавление шума с размахом  $0.01Q$  снижает уровень шума до минус 140 дБ.

Модуляторы выше 2-3 порядков с последовательным включением интеграторов работают неустойчиво, поэтому они практического применения не нашли. Структурные схемы современных модуляторов более высоких порядков достаточно сложны. В них применяется каскадное включение интеграторов 1 порядка, охваченных местными обратными связями. В качестве примера на [рис.6.11](#) приведена блок-схема 1-бит АЦП на основе SDM 5-порядка, который используется при записи звука по стандарту Super Audio CD.

В этой схеме выходы всех интеграторов суммируются со своими весовыми коэффициентами  $c_i$  и подаются на вход компаратора. Для каждой цепи обратной связи также подбирается весовой коэффициент  $c_i$ , чтобы получить огибающую спектра шума квантования необходимой формы. Шум, осуществляющий декорреляцию ошибок квантования, вводится непосредственно у входа компаратора.

Для этой схемы на [рис.6.12](#) приведена огибающая спектра шума квантования. В области низких частот у нее подъем с крутизной 12 дБ/ декаду. Это делает зависимым уровня шума от частоты в звуковом диапазоне частот. В худшем случае, на частоте около 5 КГц, уровень шума повышается до минус 90 дБ. За пределами звукового диапазона уровень шума повышается линейно с крутизной 90 дБ/декаду и на 100 кГц достигает нулевого уровня. При этом спектральные компоненты шума соизмеримы с максимальными компонентами ЗС в области этих частот.

Повышение порядка интегратора существенно улучшает SNR в области низких частот. Для сравнения на [рис.6.12](#) приведен график спектральной плотности шума квантования для 1-бит SDM 7 порядка. В таком модуляторе уровень шума на частотах ниже 100 Гц понижается до минус 195 дБ. В пределах звукового диапазона уровень



шума не превышает минус 140 дБ. В практике уже есть разработки таких модуляторов до 12 порядка.

В связи с тем, что коды коррекции ошибок и канальная модуляция рассчитаны на работу с организацией цифровых данных в виде байт и блоков, одноразрядные АЦП на основе SDM высоких порядков сами по себе используются только в аппаратуре записи звука по стандарту Super Audio CD. В большинстве случаев после такого АЦП

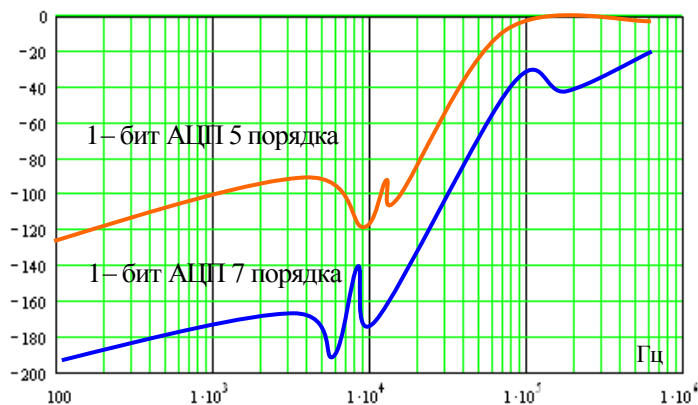


Рис.6.12. Частотная зависимость уровня шума квантования в 1-бит АЦП на основе SDM 5 и 7 порядков

включается демультиплексор, в котором частота дискретизации понижается до стандартного значения ИКМ тракта и формируются выборки с числом разрядов от 16 до 24. В качестве демультиплексора обычно используется прореживающий цифровой фильтр. Считается, что демультиплексор входит в состав АЦП, и поэтому его называют по числу выходных разрядов, например, 24-разрядным. Число выходных разрядов определяется по достигнутому значению SNR. Если  $SNR = 120$  дБ, это 20-разрядный АЦП, если 140 дБ – 24 разрядный.

#### 6.4. Одноразрядный ЦАП на основе сигма-дельта модуляции

В общем объеме производства аудиоаппаратуры более 90% приходится на долю звуковоспроизводящей техники. Поэтому очень важное значение имеет технология производства и стоимость цифро-аналоговых преобразователей. Так, например, стандартом на DVD-audio диски предусмотрено при записи использование 24-разрядной

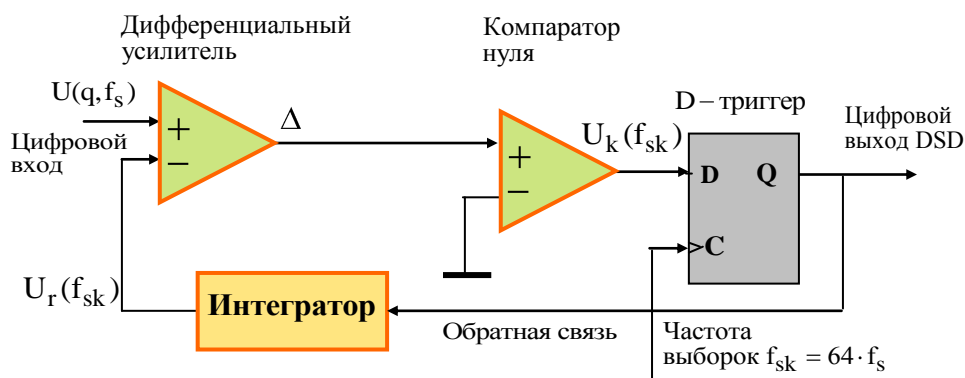


Рис.6.13. Функциональная схема 1 бит ЦАП на основе  $\Sigma\Delta$ -модулятора 1 порядка

ИКМ. Такой профессиональный АЦП стоит около 10000 долларов, тогда как для бытовых проигрывателей стоимость ЦАП не может превышать 5-10 долларов. Поэтому большинство ЦАП, используемых в проигрывателях CD и DVD-A дисков, а также в

звуковых картах компьютеров строятся на основе использования сигма-дельта модуляции и, в частности, одноразрядных ЦАП.

В цифровых SDM, используемых для цифро-аналогового преобразования, частота дискретизации известна априори и выбор коэффициента передискретизации не связан с процессом записи, как это имеет место в аналоговых SDM. Поэтому значение частоты дискретизации  $f_{sk}$ , на которой работает модулятор, может быть сколь угодно большим и ограничивается только быстродействием элементов схемотехники. Это значит, что необходимое значение SNR может достигаться при значительно меньшем порядке интегратора модулятора, чем в АЦП.

Хотя модели аналогового и цифрового модуляторов 1 порядка совершенно одинаковы, при практической реализации их функциональные схемы существенно различны. В первом - интегратор находится в цепи прямой связи, а во втором – в цепи обратной связи. Функции рекуантователя при цифро-аналоговом преобразовании выполняет компаратор нуля. Он преобразует  $q$ -разрядные выборки в непрерывный двоичный поток.

В приведенной на рис.6.13 схеме цифрового SDM 1 порядка на один из входов дифференциального усилителя подается цифровой  $q$  разрядный ИКМ сигнал  $U(q, f_s)$  с частотой выборки  $f_s$ . Однако, в соответствии с ранее принятыми условиями анализа, он

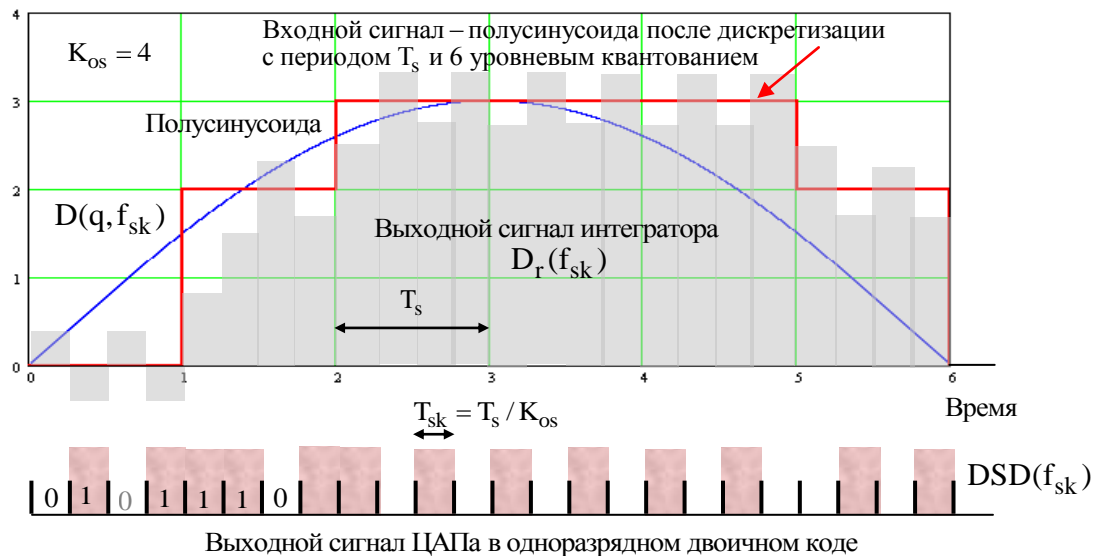


Рис.6.14. Временная диаграмма работы одноразрядного ЦАП

представляется в виде последовательности выборок ЗС, формирующих ступенчатый сигнал (рис.6.14). В дифференциальном усилителе производится сравнение его с опорным импульсным сигналом  $D_T(f_{sk})$  с высокой частотой дискретизации  $f_{sk} = K_{os} f_s$ . Этот сигнал формируется интегратором, на вход которого подается сигнал D-триггера, сдвинутый по времени относительно входного на 1 такт. Отрицательная обратная связь стремится уравнивать среднее значение выходного сигнала с входным.

Если опорный сигнал с интегратора меньше входного на выходе компаратора формируется логическая 1 и опорный сигнал увеличивается. Этот процесс повторяется в течение периода  $T_s$ , пока опорный сигнал не станет больше входного. Тогда выходной сигнал дифференциального усилителя становится меньше нуля ( $\Delta < 0$ ), на выходах компаратора и D-триггера формируется логический 0, поэтому опорный сигнал начинает уменьшаться. Выходная последовательность 1 0 1 0 ... будет продолжаться до тех пор пока на вход не поступит новая выборка входного сигнала. Это поясняется приведенным рисунком.

В приведенной схеме на выходе формируется однополярный сигнал DSD, такой же как в АЦП, у которого среднее значение меняется по времени по такому же закону, как и

на входе модулятора. Фактически все операции при цифро-аналоговом преобразовании осуществляются в цифровом виде, и на выходе может быть сформирована любая двоичная последовательность.

Обычно, выходной сигнал одноразрядного ЦАП подается на время импульсный преобразователь с цифровой модуляцией параметров импульсов. В настоящее время используются три вида ВИМ: PDM – Pulse Density Modulation (модуляция импульсов по плотности), PWM – Pulse Width Modulation (модуляция импульсов по ширине- широтно-импульсная модуляция) и PLM – Pulse Length Modulation (модуляция импульсов по длине). Все эти виды модуляции позволяют выделять ЗС из этого потока с помощью простейшего ФНЧ 3 порядка.

На основе одного SDM 1 или 2 порядка невозможно обеспечить необходимое значение SNR 120...140 дБ, поэтому используется их каскадное включение. На рис.6.15. приведена функциональная схема одноразрядного ЦАП фирмы Philips на основе

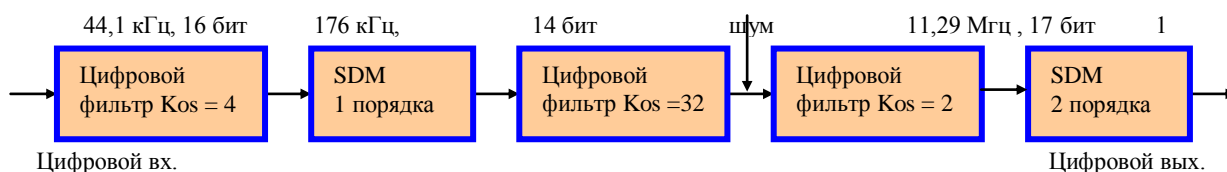


Рис.6.15. Блок-схема ЦАП фирмы Philips на основе технологии Noise Shaping

технологии «Noise Shaping». Входные цифровые 16-ти разрядные выборки с частотой дискретизации 44.1 кГц подаются на вход цифрового фильтра передискретизации. В схеме используется нерекursивный с 4-кратной передискретизацией FIR (finite impulse response) интерполяционный фильтр с линейной фазовой характеристикой. На первом этапе модуляции в результате переквантования число разрядов в выборках понижается с 16 до 14 и используется SDM 1 порядка.

Таблица 6.1. Характеристики ЦАП на основе  $\Sigma\Delta$ -модуляции и ВИМ

Вид модуляции	Коэффициент передискретизации	Частота дискретизации, МГц	Порядок интегратора,	Число уровней квантования
PDM	256	11,29	1 и 2	2
PLM	1024	45,1584	2	2

Затем еще раз производится передискретизация с помощью двух ступеней ( $K_{os}=32$  и 2). Между этими ступенями в тракт вводится шумовой сигнал, осуществляющий операцию «Dithering» с уровнем шума равны минус 20 дБ. Она уменьшает нелинейность передаточной функции из-за ошибок квантования. Общий коэффициент передискретизации равен 256 и частота дискретизации увеличивается до 11,29 МГц.

Во второй ступени модуляции используется SDM 2 порядка и формируется 1-

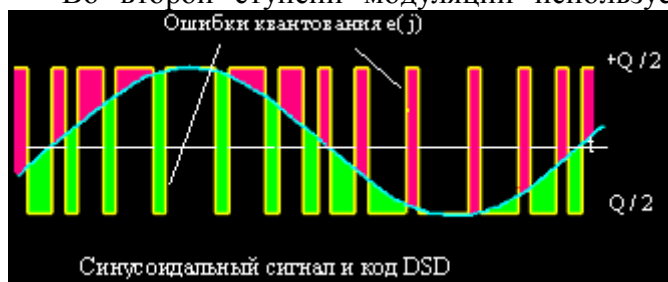


Рис.6.16. Ошибки квантования 1 бит ЦАП

разрядный цифровой поток. К выходу ЦАП подключается время-импульсный цифровой модулятор, преобразующий цифровые данные в последовательность импульсов, модулированных по плотности (PDM).

В табл.6.1. приведены некоторые данные по современным ЦАП фирм Philips, Matsushita и Sony с использованием технологии «Noise

Shaping» и ВИМ. Эти ЦАП обеспечивают SNR до 115...120 дБ, поэтому изготовители проигрывателей указывают в технических характеристиках эквивалентное число разрядов 18...20, иногда приводится и значение коэффициента передискретизации.

Одноразрядные ЦАП на основе SDM имеют целый ряд важных особенностей. Их несомненным преимуществом является высокая линейность во всем диапазоне изменения ЗС и отсутствие особых требований к точности изготовления элементов схемы.

В ИКМ трактах шум квантования возникает только при подаче ЗС, его мощность всегда равна  $Q^2/12$  и от уровня ЗС она не зависит. Если реквантователь является простым компаратором, шум квантования на его выходе существует независимо от ЗС (рис.6.16). У одноразрядных цифровых SDM сумма мощностей шума квантования и сигнала постоянна и равна  $Q^2/4$ .

При подаче ЗС мощность шума квантования уменьшается, но незначительно. В случае максимального допустимого ЗС эта мощность определяется равенством

$$P(e) = \frac{7Q^2}{32},$$

где  $Q$ - шаг квантования,  $e$  – мгновенное значение шума квантования. При этом мощность ЗС синусоидальной формы соответственно равна

$$P_{\text{sin}} = \frac{Q^2}{32}.$$

Таким образом, в лучшем случае SNR равен **минус 8,45 дБ**. Если уровень ЗС уменьшается до минус 60 дБ, то SNR становится равным минус 68,5 дБ! Это значит, что шум в 3000 раз больше ЗС. Чтобы в этом случае качественно воспроизводить звук приходится использовать очень высокую частоту дискретизации и SDM высоких порядков, что приводит к нестабильности работы из-за наличия отрицательных обратных связей.

Прямая зависимость шума квантования от уровня ЗС приводит к возникновению неустраняемого модуляционного шума. Для его устранения требуется использовать технологию Dithering, но при одноразрядной SDM это приводит к перегрузке реквантователя и возникновению самовозбуждения. Одноразрядные ЦАП крайне чувствительны к фазовому джиттеру частоты дискретизации, которая тактирует работу модулятора. Даже небольшой джиттер вызывает ошибки, равные кванту.

Еще одной проблемой построения одноразрядных ЦАП являются ФНЧ, реконструирующие ЗС. Выходной сигнал такого ЦАП представляет собой последовательность прямоугольных импульсов высокой частоты с очень высокой крутизной нарастания и спада. Для фильтрации ЗС требуются специальные конструкции ФНЧ, в которых подавляются ВЧ составляющие спектра, возникающие из-за емкостных связей. Из-за применения SDM высоких порядков в спектре ошибок квантования очень велики ВЧ составляющие за пределом звукового диапазона. Их также очень трудно фильтровать из-за наличия емкостных связей между элементами ФНЧ. По всем этим причинам одноразрядные ЦАП в высококачественных звуковых трактах не применяются.

## 6.5. Многоразрядные ЦАП

Много разрядные ЦАП предназначены для работы в высококачественных звуковых трактах, таких проигрыватели DVD-Audio и Super Audio CD или дорогие звуковые карты компьютеров с числом звуковых каналов 5 и больше. Обычно число разрядов таких ЦАП от 3 до 6, иногда может быть и до 14. Главной характеристикой ЦАП является число уровней выходного сигнала, которое выбирается всегда нечетным и может быть от 3 до 31.

На рис.6.17 приведена функциональная схема много разрядного ЦАП. Цифровой фильтр-интерполятор осуществляет повышение частоты дискретизации в 64 или 128 раз. Для обеспечения устойчивости работы порядок интегратора цифрового сигма-дельта модулятора 2 или 3. Если одно разрядные ЦАП являются линейными по принципу, то в много разрядных преобразователях эта линейность обеспечивается точностью технологии их изготовления.

Высококачественные много разрядные аналоговые ЦАП требуют специальной подгонки элементов для обеспечения линейности передаточной функции. В противном случае возникают ошибки преобразования, которые коррелированы со звуковым сигналом и поэтому очень заметны на слух. Для исключения таких требований в современных много разрядных ЦАП на основе SDM используется специальная

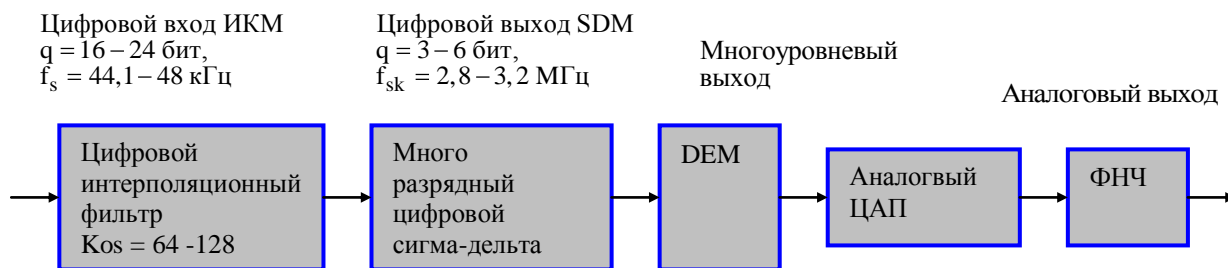


Рис.6.17. Функциональная схема много разрядного ЦАП на основе сигма-дельта модуляции технология обработки цифровых звуковых сигналов под названием DEM (Dynamic Element Matching), с помощью которой детерминированные ошибки ЦАП преобразуются в случайные и усредняются почти до нулевого значения. Существует множество вариантов реализации этой технологии.

Важнейший недостаток одно разрядных ЦАП это очень высокий уровень ошибок квантования компаратора, из-за которого SNR не превышает минус 8,45 дБ. В много

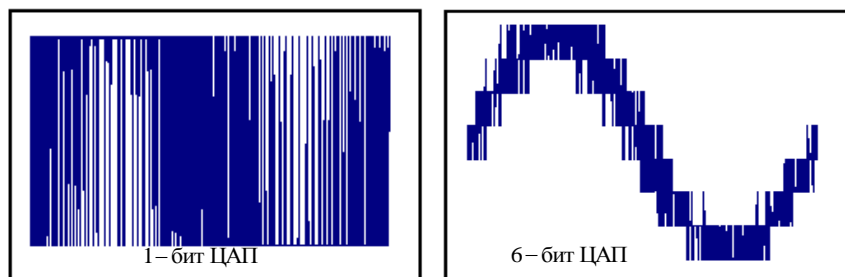


Рис.6.18. Влияние фазового джиттера на выходной сигнал ЦАП

разрядных ЦАП отношение сигнал-шум определяется по известной формуле

$$SNR = 6,02q + 1,76\text{дБ}$$

и уже при трех битах  $SNR = +20$  дБ, а при 6 битах -  $+38$  дБ. Различие огромное, главное, что сигнал выше уровня шума. Это существенно облегчает задачу фильтрации ошибок квантования как в звуковом диапазоне, так и за его пределами. Важно также, что форма сигнала на выходе модулятора приближается к форме воспроизводимого звукового сигнала и не носит чисто импульсного характера. Другое важное преимущество много разрядных ЦАП это значительно меньшая чувствительность к паразитной фазовой

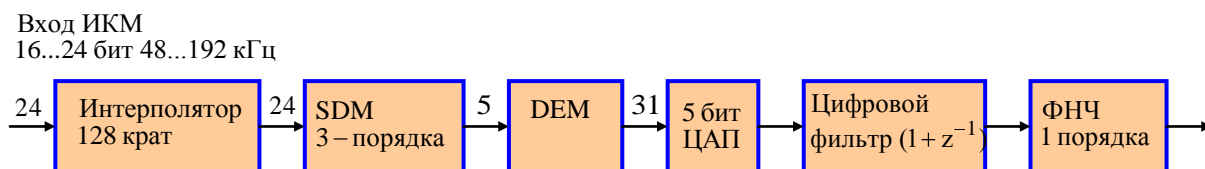
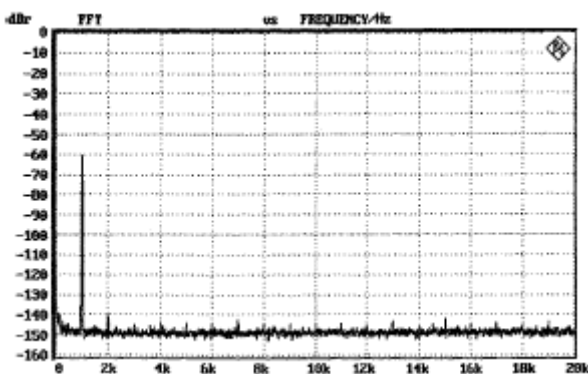


Рис.6.19. Функциональная схема ЦАП для проигрывателей DVD-Audio модуляции тактовой частоты модулятора (рис.6.18).

Уменьшение шума квантования и влияния фазового джиттера наиболее сильно проявляется когда используется нечетное число уровней квантования, так как при этом положительные и отрицательные значения цифрового сигнала становятся симметричными относительно нулевого значения.

В качестве примера на [рис.6.19](#) приведена функциональная схема много разрядного 31 уровневый ( $\pm 15$ ) ЦАП для высококачественных цифровых трактов, который работает с частотой дискретизации до 192 кГц с динамическим диапазоном 120 дБ. На [рис.6.20](#) представлен спектр выходного сигнала этого ЦАП при уровне ЗС минус 60 дБ, из которого видно, что во всем звуковом диапазоне уровень шума не превышает минус 140 дБ. Известно много квантования может быть  $\pm 3, 7$  и  $11$ . С увеличением числа разрядов качественные характеристики ЦАП улучшаются, но сильно усложняются устройство и алгоритм работы DEM.

Особое место занимают много разрядные ЦАП, не использующие технику DEM. Они используются в трактах среднего качества со значением SNR не выше 100 дБ, когда для снижения стоимости аналоговых ЦАП применяют 12-14 разрядные SDM 1 порядка, но



6.20. Спектр ЗС с уровнем минус 60 дБ на выходе ЦАП

хотят получить значение SNR такое же, как 16-ти разрядном ЦАП. В этом случае используются возможности сигма-дельта модуляции, при которой уменьшение разрядности на каждый бит может компенсироваться двукратным увеличением частоты дискретизации.

Функциональная схема такого ЦАП, используемая фирмой Филипс в CD плеерах, приведена на [рис.6.21](#). В этой схеме на вход цифрового фильтра интерполятора подается кодовая последовательность в формате 44,1/16. В интерполяторе осуществляется

повышение частоты дискретизации в 4 раза до 176,4 Гц. При этом число разрядов в выборках увеличивается до 17 бит. В преобразователе Noise Shaper 1 порядка осуществляется округление, при котором отбрасываются младшие 3 разряда, и в выходном сигнале остаются только старшие 14 разрядов. По цепи отрицательной обратной связи младшие разряды с задержкой на один такт подаются на сумматор. Формирователь шума стремится произвести округление таким образом, чтобы выходной

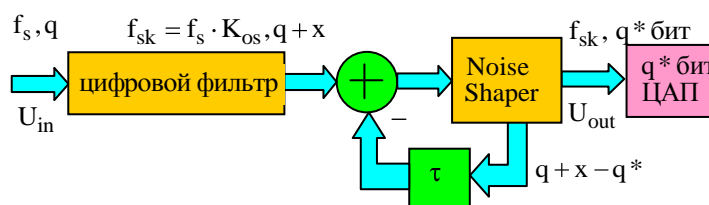


Рис.6.21 ЦАП на основе SDM 1 порядка

14 разрядный сигнал был как можно ближе к входному 17 разрядному сигналу. Значение SNR 14-разрядного ЦАП получается такое же, как у 16-ти разрядного около 100 дБ.

Физическая суть работы приведенной схемы поясняется [рис.6.22](#). Пусть на входе ЦАП действует дискретизированный и квантованный цифровой сигнал с числом уровней 9 в виде ступенчатого сигнала ([рис.6.22.a](#)) с частотой дискретизации  $f_s$ . Изменение уровня этого сигнала происходит в соответствии со шкалой квантования на 8 делений.

**Первая операция** – это усечение цифрового сигнала путем удаления *двух младших разрядов* ([рис.6.22.б](#)). При уменьшении числа разрядов на 2 бита шаг квантования увеличивается в 4 раза и квантование происходит по шкале с двумя уровнями. Округление при квантовании осуществляется к ближайшему делению шкалы на 2 уровня, поэтому в 1-й период дискретизации  $T_{s1}$  уровень квантованного сигнала больше уровня исходного и ошибка квантования имеет знак плюс.

Во 2-й период дискретизации  $T_{s2}$  уровень квантованного сигнала меньше входного уровня и ошибка имеет знак минус. В третий период дискретизации  $T_{s3}$  ошибка квантования очень большая со знаком плюс. Все это свидетельствует об увеличении шума квантования при уменьшении числа разрядов, как и следовало, ожидать.

**Вторая операция** – это 4-х кратное увеличение частоты дискретизации без изменения числа разрядов. Квантованный сигнал после выполнения этой операции приведен на рис.6.22.в).

**Третья операция** - компенсация ошибок реквантования. Система с обратной связью стремится, чтобы среднее значение ошибки за 4 такта (это исходный период дискретизации) равнялось нулю. Если в 1 такте ошибка положительная, то в следующий такт она делается отрицательной. Если сумма первых двух ошибок отрицательная, то следующая ошибка снова положительная. Если сумма трех ошибок отрицательная,

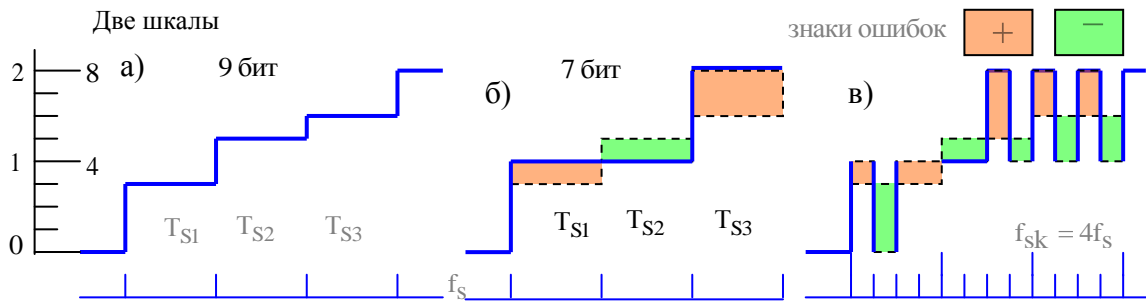


Рис.6.22. К вопросу уменьшения шума квантования при использовании технологии в много разрядном ЦАП с технологией Noise Shaping 1 порядка

следующая ошибка положительная.

В результате среднее значение ошибок квантования за период дискретизации всегда равно нулю. Таким образом, при использовании отрицательной обратной связи *результат каждого квантования зависит от результата предыдущего*. Очевидно, что такая компенсация ошибок может осуществляться достаточно точно только на низких частотах. С повышением частоты ЗС компенсация будет ухудшаться, и ошибки квантования будут увеличиваться.

## 7. Кодирование звуковых сигналов

### 7.1. Корректирующее кодирование

#### Основные понятия и определения

*Корректирующий код* предназначен для обнаружения и исправления кодовых ошибок. Принцип такого кодирования основан на создании избыточности в коде. Под этим понимается введение дополнительных проверочных бит в последовательность информационных данных. Поэтому кодирование – это процесс вычисления проверочных бит, а декодирование – процесс обнаружения ошибок и их исправление. Чем больше избыточность корректирующего кода, тем выше его обнаруживающая и исправляющая способности. При этом, естественно, возрастает скорость цифрового потока и уменьшается длительность записи на носителе. В цифровой звукозаписи для обнаружения и исправления кодовых ошибок используется только блочное кодирование

*Блочное кодирование* – это кодирование без памяти, потому что кодирование и декодирование таких кодов осуществляется на основе данных только одного блока. Блочный код может быть *двоичным и недвоичным*. В общем случае основой блочного кода является *символ*. Длина символа равна  $t$  бит и рассматривается как единое целое. В частном случае, когда  $t = 1$ , код становится двоичным. С символами оперируют также как с битами.

При блочном кодировании звука последовательность цифровых выборок левого и правого звуковых каналов разбивается на символы длиной 4, 8 или 16 бит. Эти символы образуют *информационное слово* из  $k$  символов. При кодировании в блок добавляется  $r$  *проверочных символов* длиной  $t$ , которые формируют *проверочное слово*. В результате кодирования формируется *кодированное слово* длиной  $n = k + r$  символов, и такой код обозначается, как  $(n, k)$ . Введение проверочных символов создает избыточность кода, определяемая равенством

$$R = \frac{n - k}{n},$$

где  $r = n - k$ .

Если проверочные символы располагаются в конце блока, тогда код называется *симметричным*. Номера всех символов в блоке  $i$  отсчитываются справа налево от нуля, этот номер определяет позицию символов в блоке и называется *локатором*. При выполнении математических операций символы и слова могут быть представлены векторной форме. В процессе корректирующего кодирования и декодирования вектора можно умножать, делить и суммировать по правилам многочленов из векторной алгебры.

Важнейшим понятием блочного кодирования является *поле кода*, под которым понимается все *конечное множество* кодовых комбинаций бит в символе, для которых определены правила выполнения всех математических операций (сложение, умножение.....). Такие конечные поля называются полями Галуа и обозначаются как  $GF(2^m)$ . Число элементов поля  $\sigma$  называется его *порядком* и определяется как

$$\sigma = 2^m.$$

*Двоичное поле* (в этом поле  $t = 1$ ,  $\sigma = 2$ ). Элемент поля (символ) может принимать только два значения 0 и 1. Математические правила:  $0 + 0 = 0$ ,  $1 + 1 = 0$ ,  $0 + 1 = 1 + 0 = 1$ ,  $0 \times 1 = 1 \times 0 = 0$ ,  $0 \times 0 = 0$  и  $1 \times 1 = 1$ . Информационные и кодовые слова блока в двоичном поле могут быть представлены в двоичной форме или в виде степенного полинома. В процессе корректирующего кодирования вектора можно умножать, делить и суммировать по правилам многочленов из векторной алгебры. Кодовое слово длиной  $n$  обычно



записывается как сумма информационного и проверочного слов, представляемых в виде полиномов степени не выше  $n-1$

$$a(x) = k(x) + r(x),$$

где

$$k(x) = a_1x^{n-1} + a_2x^{n-2} + \dots + a_i x^{n-i} + \dots + a_k x^{n-k}, \quad r(x) = a_{k+1}x^{r-1} + a_{k+2}x^{r-2} + \dots + a_{n-1}x + a_n,$$

$i = 0, 1, 2, \dots, n-1$ ,  $a_i$  -весовой коэффициент, который может принимать значения 0 или 1. Эти слова могут быть представлены и в двоичной форме.

Пример двоичного представления кодового слова длиной  $n=8$ :

7 6 5 4 3 2 1 0 – номера позиций кодового слова блока  
( $i = 0, 1, 2, \dots, n-1$ ),

$b = (1 \ 1 \ 0 \ 0 \ 1 \ 0 \ 1 \ 0)$  – кодовое слово в двоичной форме.

Это же слово в виде полинома степени  $n-1$ :

$$b(x) = x^7 + x^6 + x^3 + x.$$

Таблица 7.1. Поле $GF(2^4)$ при $F(x) = x^4 + x + 1$ степени $m$					
Двоичный код	Полином	Степень	Двоичный код	Полином	Степень
0000		0	1011	$x^3 + x + 1$	$\alpha^7$
0001	1	$\alpha^0$	0101	$x^2 + 1$	$\alpha^8$
0010	$x$	$\alpha$	1010	$x^3 + x$	$\alpha^9$
0100	$x^2$	$\alpha^2$	0111	$x^2 + x + 1$	$\alpha^{10}$
1000	$x^3$	$\alpha^3$	1110	$x^3 + x^2 + x$	$\alpha^{11}$
0011	$x + 1$	$\alpha^4$	1111	$x^3 + x^2 + x + 1$	$\alpha^{12}$
0110	$x^2 + x$	$\alpha^5$	1101	$x^3 + x^2 + 1$	$\alpha^{13}$
1100	$x^3 + x^2$	$\alpha^6$	1001	$x^3$	$\alpha^{14}$

**Недвоичное поле** (например, поле  $GF(2^4)$  -  $m = 4$ ,  $\sigma = 16$ ). В этом поле символ может принимать 16 кодовых значений, включая нулевое. Символы недвоичного поля могут быть представлены в двоичной форме, в виде многочлена и в показательной форме. Это связано с тем, что некоторые математические операции проще выполнять в двоичном виде, например, операции сложения. Умножение и деление удобнее производить в показательной или степенной формах. Переход из одной формы в другую производится с помощью таблицы.

Поле недвоичного кода образуется с помощью *примитивного элемента*  $\alpha$ , который является одним из корней *порождающего многочлена* степени  $m$ , например,  $F(x) = x^4 + x + 1$ .

Единичный элемент поля в показательной форме записывается в виде  $\alpha^0 = 1$ , все остальные элементы поля выражаются как степени примитивного элемента:  $\alpha^1 = 1\alpha^0$ ,  $\alpha^2 = \alpha^1 \alpha$ , ..... При построении поля операции умножения в показательной форме производятся по модулю  $\alpha$  в степени  $2^m - 1$ , поэтому справедливы такие равенства:  $\alpha^6 \alpha^9 = \alpha^{15} = \alpha^0$ . В **табл.7.1.** приведены результаты расчетов элементов поля в показательной форме.

Расчет элементов поля можно также производить, используя представление примитивного элемента в виде степенного полинома  $x$ . В этом случае операции умножения производятся по модулю порождающего многочлена  $F(x)$ . При этом результатом умножения является остаток от деления произведения многочленов на порождающий полином. В качестве примера приведем расчета элемента поля  $\alpha^4$ :

$$\alpha^4 = x^3 \times x \equiv r(x) \pmod{F(x)},$$

где  $r(x)$  – остаток от деления многочленов. Он вычисляется следующим образом:

$$\frac{x^4}{x^4 + x + 1} \left| \begin{array}{l} x^4 + x + 1 \\ \hline 1 \end{array} \right.$$

$r(x) = x + 1$  (остаток),

поэтому  $\alpha^4 = x + 1$ . Результаты этих расчетов элементов поля в полиномиальном виде приведены в **табл.7.1**. Операции сложения всегда выполняются в двоичном виде. Таким же образом производится расчет полей и более высокого порядка с другими порождающими полиномами.

Исправляющая способность блоковых корректирующих кодов определяется *минимальным кодовым расстоянием*  $d_{\min}$ . Под этим термином понимается минимальное число позиций, по которым *отличаются кодовые слова*. Например, для корректирующего кода (6,3) минимальное кодовое расстояние определяется по двум кодовым словам  $b1 = (1\ 0\ 1\ 1\ 1\ 0)$  и  $b2 = (1\ 1\ 1\ 0\ 0\ 0)$ . Эти слова отличаются в 3 позициях (1,2,4 справа на лево) и, следовательно, кодовое расстояние кода равно 3.

Для блоковых кодов фундаментальным является неравенство

$$d_{\min} \leq 1 + r,$$

которое определяет возможности кода обнаруживать и исправлять ошибки. Блоковые коды, для которых это неравенство превращается в равенство, называются *разделимыми кодами с максимальным расстоянием* (МДР)- это, например, коды Рида Соломона (RS). Для них гарантированное *число исправляемых ошибок*  $t$ , локаторы которых неизвестны, вдвое меньше числа проверочных символов

$$t = r/2.$$

**Обнаруживающая способность** МДР кода вдвое выше исправляющей способности. Число обнаруживаемых ошибок в блоке  $t^*$  определяется формулой

$$t^* = r.$$

Это значит, что при 4 проверочных символах в блоке обнаруживаются до 4 ошибок, а исправляются только 2.

Часто используются такие стратегии декодирования, при которых локаторы символов, которые *только возможно ошибочны*, известны. Такие ошибки называются *стираниями*  $\bar{t}$ . В этом случае кодовое расстояние для МДР блоковых кодов определяется другим равенством

$$d_{\min} = 2t + \bar{t},$$

поэтому исправляющая способность блокового кода значительно выше. При этом число исправляемых ошибок и стираний определяется равенством

$$2t + \bar{t} = r.$$

Из этой формулы следует, что при 4 проверочных символах в блоке данных может быть исправлено только 2 ошибки или 4 стирания. Возможно исправление одной ошибки и двух стираний.

### Кодовые ошибки

Под *кодowymi ошибками* могут пониматься ошибочные биты, символы, слова и блоки. Очевидно, что если ошибочен хотя бы один бит в символе, ошибочным является и символ, а также кодовое слово и блок, в который он входит. Если ошибки случайны и независимы, для их характеристики используется *понятие вероятности*. Этим термином обозначается отношение числа ошибочных бит, символов, кодовых слов или блоков к их общему числу за время измерения. Вероятности этих событий:  $P_b$ ,  $P_w$ ,  $P_s$ , и  $P_{blk}$ . Для случайных независимых ошибок эти вероятности связаны равенствами:

$$P_s = m \cdot P_b, \quad P_w = q \cdot P_b, \quad P_{blk} = n \cdot P_s$$

На практике условие независимости ошибок большей частью не выполняется и тогда ошибки характеризуются их относительной скоростью (rate): BER(bit error rate),

SER(symbol error rate), BLER (blok erroe rate) и WER (word error rate) Эти термины определяют отношение частоты ошибочных бит, символов, слов и блоков к частоте их повторения ( $F_b, F_w, F_s$  и  $F_{blk}$ ):

$$BER = \frac{F_{ошб}}{F_b}, \quad SER = \frac{F_{ошс}}{F_s}, \quad BLER = \frac{F_{ошбл}}{F_{blk}}$$

Так, для системы «Компакт Диск»  $F_{blk} = 7350$  Гц и по стандарту допустима скорость ошибочных блоков  $BLER = 0,02$ , поэтому максимальная частота повторения ошибочных блоков не должна превышать 147 бл/с.

Кодовые ошибки могут быть *одиночными*, когда они возникают случайным образом, или следовать *в виде групп*. Если ошибки в группе идут без перерыва – это называется *выпадение*, если же имеются разрывы- то это *пакетные ошибки*.

Основной характеристикой кодовых ошибок носителя записи является зависимость скорости ошибочных бит BER от SNR в канальном коде при воспроизведении записи (рис.7.1). Для построения системы перемежения кода важна зависимость вероятности кодовых ошибок от их длины. В качестве дополнительных характеристик используются такие характеристики как средняя продолжительность выпадения или пакета ошибок и среднее время между ошибками и выпадениями. Такие данные требуются при расчете систем перемежения кодовых данных для исправления выпадений и пакетов ошибок.

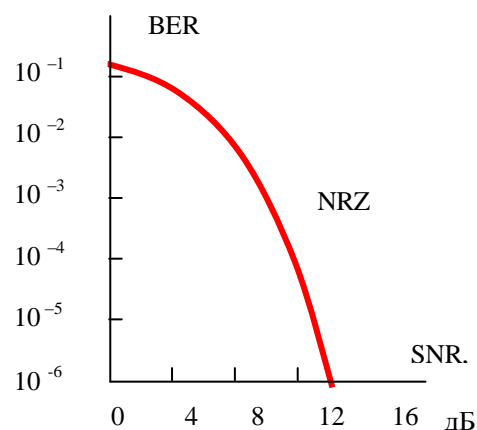


Рис.7.1. Зависимость вероятности ошибочных бит от SNR

### Обнаружение кодовых ошибок

Обнаружение ошибок при цифровой звукозаписи осуществляется с помощью специальных кодов, которые называются EDC (Error Detection Code). Для этих целей наиболее часто используется код CRC (Cyclic Redundancy Code). Это *двоичный циклический блочный код с избыточностью* длины  $n$ , обозначаемый как  $(n, k)$ .

Информационные  $k$  двоичных символов CRC кода образуют *алфавит информационных данных*  $A(k)$ . Число их кодовых комбинаций в этом алфавите составляет

$$N(k) = \sigma^k.$$

За счет избыточности кода объем возможных кодовых слов возрастает, и *полный алфавит* всех кодовых комбинаций расширяется до значения

$$N(n) = \sigma^n.$$

Из этого алфавита выбираются только  $\sigma^k$  кодовых слов, которые включают в основной алфавит кода, а остальные кодовые слова считаются *запрещенными* -  $N^*(n)$ . Если переданное кодовое слово из-за ошибки переходит в запрещенное, то такая ошибка обнаруживаются всегда. Если кодовое слово из-за ошибки переходит в другое разрешенное слово, то такая ошибка вообще не обнаруживается. Чем больше число проверочных символов, тем больше число запрещенных кодовых слов и меньше вероятность не обнаружения ошибочного кодового слова  $P_{cw}$

$$P_{cw} = \frac{N(k)}{N^*(k)} = 2^{-r}.$$

Как видно из этой формулы, *вероятность не обнаружения ошибок* в коде CRC зависит только от числа проверочных символов и не зависит от длины блока, что очень

важно. Это значит, что код можно использовать для обнаружения ошибок в блоках очень большой длины.

Код CRC с вероятностью 1 обнаруживает до 3 ошибочных бит и пакеты ошибок длиной меньше и равной  $r$  бит. Вероятность не обнаружения пакетных ошибок длиной  $r+1$  определяется равенством

$$P_{cw} = 2^{-(r-1)}$$

Обычно,  $r = 16$  или  $32$  бита. При этом вероятность не обнаружения ошибочного кодового слова находится в пределах от  $1,5 \cdot 10^{-5}$  до  $2,3 \cdot 10^{-10}$

Информационный алфавит кода CRC задается *порождающим многочленом*  $F(x)$ , по модулю которого происходит умножение двоичных векторов. Сам алфавит формируется путем циклического сдвига символов на один шаг. Поэтому все информационные слова делятся без остатка на порождающий полином. Циклическому сдвигу на один шаг соответствует умножение на  $x$ : например: 110011, 100111, 001111, 011111.....

Для обнаружения ошибок в код CRC вводится *проверочное слово* длиной  $r = n - k$ . Чтобы освободить место в блоке для  $r$  проверочных символов, информационные символы умножаются на  $x^r$

$$F(x) \cdot x^r / G(x) = F(x) \cdot x^r + r(x).$$

Вычисление проверочного слова производится путем деления информационного слова длиной  $k$  на порождающий полином. Остаток от деления является проверочным словом

$$r(x) = R \ k(x) / F(x) ,$$

он вводится в кодовое слово блока, поэтому оно всегда без остатка делится на порождающий многочлен. В стандартах по цифровой звукозаписи эта математическая операция записывается в другом виде

$$r(x) = \sum_{i=0}^{r-1} a_i x^i = I(x) \cdot x^r \bmod F(x), \quad I(x) = \sum_{i=r}^n a_i x^i.$$

При декодировании кодовые слова также делятся на порождающий многочлен. Если остаток от деления равен нулю - в блоке ошибок нет.

### Код Рида - Соломона

*Код Рида - Соломона* (RS) предназначен как для обнаружения, так и для исправления ошибок и широко используется при цифровой записи звука на магнитной ленте, оптических и магнитных дисках.

Это линейный блочный двоичный циклический код в поле Галуа  $GF(2^m)$ . Он относится к группе МДР, поэтому при выбранных значениях  $n$  и  $k$  у него наибольшее кодовое расстояние и лучшая исправляющая способность. Максимальная длина кода определяется равенством

$$n = 2^m - 1.$$

Часто используются укороченные коды RS, у которых отбрасываются старшие разряды. Важно, что кодовое расстояние укороченного кода такое же, как у исходного. Главное, код RS имеет эффективные алгебраические методы декодирования. Поле кода задается производящим многочленом  $F(x)$  и примитивным элементом.

*Кодирование кода RS производится с помощью проверочной матрицы H.* Она состоит из  $n$  столбцов и  $r$  строк. Номер столбца  $i$  (справа на лево),  $0 \leq i \leq n-1$ , номер строки  $j$  (сверху вниз),  $0 \leq j \leq r-1$ . На **рис.7.2.** приведен алгоритм кодирования кода RS (28,24) в поле Галуа  $2^8$ , запатентованный фирмой Sony для системы CD.

В этом алгоритме входное информационное слово длиной  $k$  преобразуется в кодовое, в которое добавляются  $r$  неизвестных проверочных символов. Эти символы необходимо

определить. Для этого с помощью проверочной матрицы составляется и решается система из  $r$  уравнений относительно неизвестных  $x_0, x_1, \dots, x_{r-1}$ . При этой операции используется транспонирование – преобразование вектор-строка в вектор-столбец. В результате

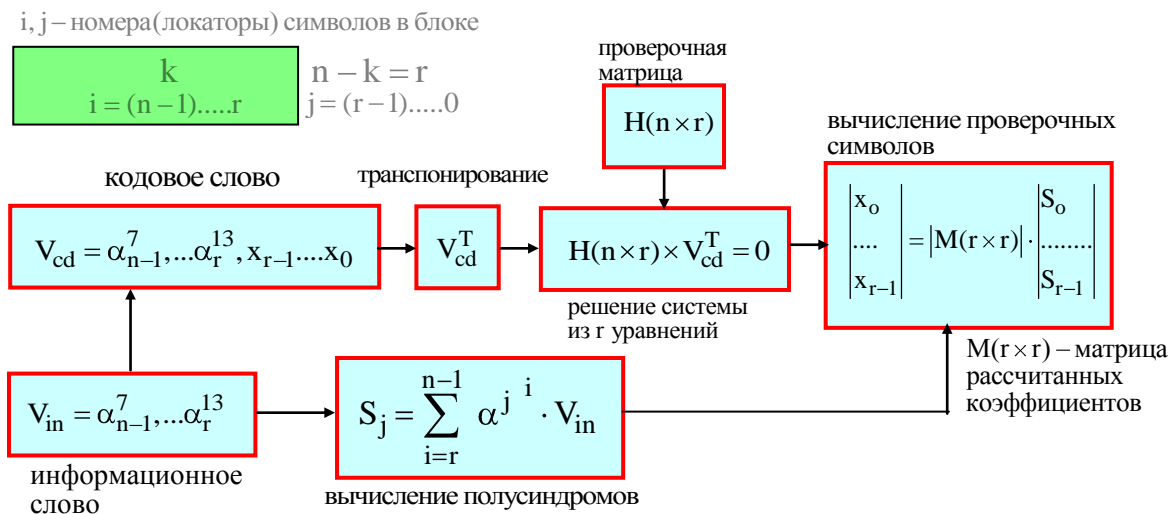


Рис.7.2.. Алгоритм кодирования кода Рида Соломона

решения проверочные символы определяются в матричной форме с помощью матрицы постоянных коэффициентов  $M(r \times r)$  и полусиндромов  $S_j$ . Полусиндромом называется произведение информационного вектора на строку проверочной матрицы.

**Декодирование кода RS** включает в себя следующие операции:

- расчет синдромов ошибок,
- обнаружение ошибок,
- идентификация ошибок (определение числа ошибок в блоке),
- определение локаторов ошибок,
- исправление ошибок.

**Синдромом ошибки** называется произведение кодового вектора на вектор  $j$  строки проверочной матрицы длиной  $n$ . Он рассчитывается по формуле

$$S(j) = \sum_{i=0}^{n-1} (\alpha^j)^i \cdot V_{cd}(i),$$

Число вычисляемых полусиндромов и синдромов ошибок равно числу проверочных символов.

**Обнаружение ошибок.** Если все синдромы ошибок равны нулю, это значит, что ошибок нет или они не обнаружены. Если известна вероятность ошибочных символов  $P_s$ , вероятность не обнаружения ошибок в этом случае определяется приближенным равенством

$$P_s = 7,3 \cdot 10^{-6} \cdot P_s^5$$

Обычно,  $P_s = 10^{-4} \dots 10^{-5}$ , поэтому  $P_s = 7,3 \cdot 10^{-26} \dots 7,3 \cdot 10^{-31}$ . Как видно, вероятность не обнаружения ошибок в коде RS ничтожно мала.

**Идентификация ошибок.** При этой операции определяется сколько именно ошибок в блоке. В зависимости от числа ошибок меняется дальнейшая стратегия декодирования. Одиночные и двойные ошибки исправляются, а если их больше производится отказ от декодирования и осуществляется операция стирания. Идентификация производится на основе использования синдромов ошибок по расчетным формулам, хранящимся в памяти декодера.

**Определение локаторов и значений ошибок.** Расчеты локаторов и исправление ошибок выполняются по алгебраическим формулам с использованием синдромов ошибок. Это

возможно лишь в частных случаях 1 или 2 ошибок, когда решаемая система уравнений не выше 4 порядка. Следует отметить, что при декодировании с исправлением 2 ошибок число неизвестных –4, это два локатора и две ошибки. Чтобы иметь возможность исправлять больше ошибок, необходимо увеличивать число проверочных символов. Уже при 3 ошибках необходимо решать систему из 6 уравнений. Это приводит очень сильному усложнению аппаратуры декодирования и резко снижается быстродействие декодера.

Код определяет лишь потенциальную исправляющую способность декодера. Реальная

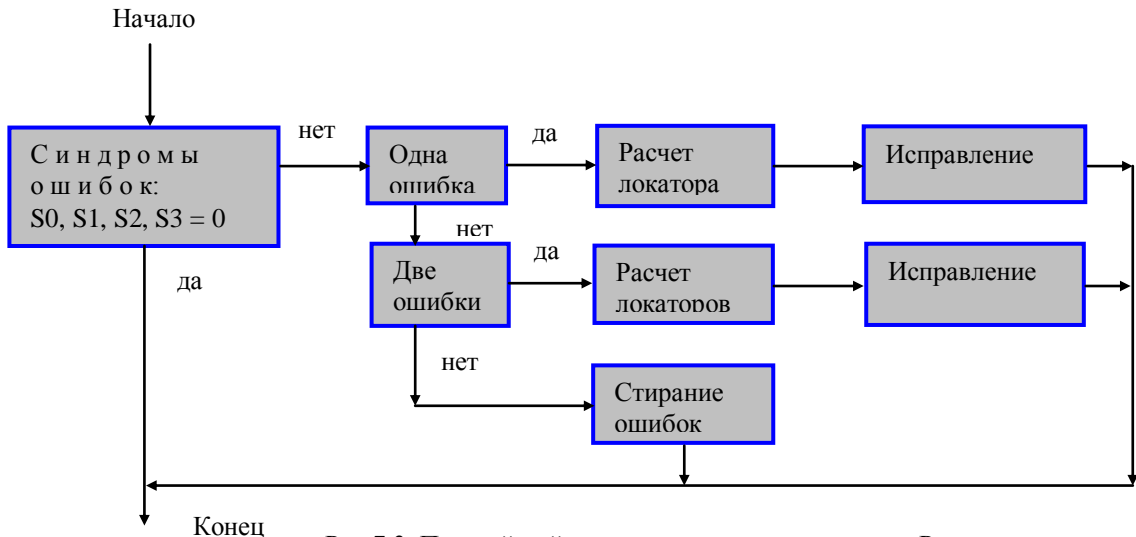


Рис.7.3. Простейший алгоритм декодирования кода Рида

исправляющая способность зависит от используемой стратегии декодирования. При одном и том же корректирующем коде стратегий декодирования может быть много. На рис.7.3. приведен простейший алгоритм декодирования с исправлением одной и двух ошибок. В этой схеме предполагается, что при кодировании используется код RS с 4 проверочными символами.

Декодирование начинается с вычисления синдромов ошибок  $S_j$ . Для приведенной схемы вычисляется 4 синдрома.

**Обнаружение ошибок в блоке** осуществляется путем проверки синдромов ошибок. Если

$$S_0 = S_1 = S_2 = S_3 = 0,$$

то ошибок в блоке нет, и на этом операция декодирования заканчивается. Если хотя бы один синдром не равен нулю, это значит, что блок ошибочен. Тогда вычисляются 3 вспомогательных коэффициента:

$$A = S_1^2 + S_0 S_2, \quad B = S_1 S_2 + S_0 S_3, \quad C = S_2^2 + S_1 S_3.$$

**Определение числа ошибок в блоке** (идентификация ошибок). Если

$$S_0 \cup S_1 \cup S_2 \cup S_3 \neq 0, \quad A = B = C = 0,$$

то это значит, что в блоке только одна ошибка. Если

$$S_0 \cup S_1 \cup S_2 \cup S_3 \neq 0, \quad A, B, C \neq 0,$$

то в блоке две ошибки.

**Определение локаторов ошибок.** В случае одиночной ошибки локатор  $i$  определяется с помощью синдромов ошибок из равенства

$$\alpha^i = \frac{S_1}{S_0}.$$

Если в блоке две ошибки, то локаторы  $i$  и  $j$  определяются значительно более сложным путем из двух основных равенств:

$$\alpha^i = \frac{\sigma_1}{X}, \quad \alpha^j = \frac{\sigma_2}{Y},$$

где

$$\alpha^i = \frac{\sigma_1}{X}, \quad \alpha^j = \frac{\sigma_2}{Y}, .$$

При одной ошибке в блоке исправление ошибки производится в соответствии с равенством

$$\Delta\alpha_i = S_0 .$$

Если в блоке две ошибки и известны их локаторы  $i$  и  $j$ , то они рассчитываются по формулам:

$$\Delta\alpha_j = \frac{\alpha^i S_0 + S_1}{\alpha^i + \alpha^j}, \quad \Delta\alpha_i = \Delta\alpha_j + S_0 .$$

При двухступенчатом декодировании в ситуации, когда вероятность ошибочных символов на входе декодера  $P_s = 10^{-4}$  вероятность неправильного исправления одиночных ошибок не превышает  $9,4 \cdot 10^{-19}$ , двойных –  $4,8 \cdot 10^{-12}$ , а вероятность отказа от декодирования может достигать  $4,65 \cdot 10^{-10}$ . Важно еще отметить, что высокая исправляющая способность кода RS реализуется лишь в относительно «хороших» каналах, когда вероятность ошибочных символов не превышает  $10^{-4}$ . В «плохих» каналах, когда вероятность ошибочных символов больше  $10^{-3}$  его исправляющая способность резко падает и использовать его нельзя.

### Перемежение

Даже самые мощные коды не в состоянии исправить более 4...8 ошибок в блоке. В тоже время кодовые ошибки часто бывают в виде выпадений или длинных пакетов, когда подряд могут быть ошибочными сотни и тысячи символов подряд. *Перемежение* – это способ преобразования выпадений и пакетов ошибок в 2-х...4-х кратные ошибки в блоке, которые могут быть исправлены кодом RS.

Такое преобразование осуществляется путем перестановки отдельных символов или целых групп символов в кодовой последовательности после корректирующего кодирования и восстановления исходного порядка перед декодированием. Подобная операция называется *перемежение – деперемежение* символов.

Перемежение обычно осуществляется на уровне символов с перестановкой их из одного блока в другой. Иногда оно осуществляется на уровне кодовых слов длиной  $q$ , формируемых при ИКМ. При этом осуществляется перестановка слов левого и правого звуковых каналов в пределах одного блока. Для исправления пакетов ошибок основной является межблочная перестановка всех символов блока.

Различают *короткое и длинное перемежения*. Первое предназначено для исправления выпадений малой длины (несколько символов), когда их вероятность достаточно высока. Второе предназначено для исправления длинных выпадений превышающих несколько блоков, при условии, что их вероятность невелика. Реализуются эти перемежения по разному.

На **рис.7.4.** приведен пример *короткого межблокового перемежения* символов с интервалом в 1 блок и глубиной перемежения 2 блока для кода с исправлением  $t = 2$  ошибок в блоке. В этой схеме производится перемещение в следующий блок только четных символов. Если в одном блоке будет подряд 4 ошибочных символа, которые

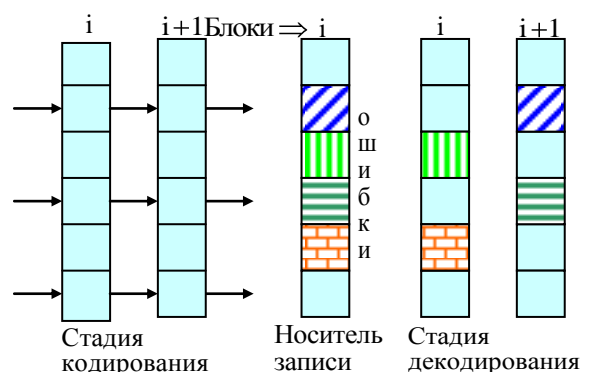


Рис.7.4. Короткое межблоковое перемежение-деперемежение символов

исправить невозможно, то после депережежения ошибочны будут подряд 2 блока, но в них будет только по две ошибки, которые могут быть исправлены.

Минимальное расстояние в блоках, на которое смещаются символы исходного блока, называется *интервалом перемежения* ( $I_p$ ). Другой важной характеристикой является *глубина перемежения*, которая определяется числом блоков, в которых размещаются символы после перемежения. Глубина перемежения – это зона (в блоках или символах), в которой размещаются все символы одного блока после их перестановки. Эта зона тем больше, чем длиннее интервал  $I_p$  и больше символов в блоке. Глубина перемежения определяет объем памяти, необходимой для обеспечения работы системы перестановки символов.

Для увеличения длины исправляемых выпадений, вплоть до нескольких блоков подряд, используется так называемое *сверточное перемежение*. Оно отличается тем, что все символы исходного блока перемещаются в разные блоки и интервал перемежения у каждого символа свой. Интервал перемещаемого 1 символа называется основным и обозначается  $I_{п1}$ .

При сверточном перемежении данные после корректирующего кодирования поступают в ОЗУ в виде прямоугольной матрицы. Число строк матрицы равно длине кодового вектора  $N$ , а число столбцов определяется произведением длины кодового вектора на основной интервал перемежения минус один. Это есть и глубина перемежения. При сверточном перемежении длина исправляемых выпадений в блоках равна  $I_{п1} \times t = 4$  блока.

В этой матрице запись данных блоков производится по столбцам в параллельном коде, а считывание символов осуществляется с главной диагонали матрицы в



Рис.7.5. Сверточное межблоковое перемежение символов

последовательном коде с заданным основным интервалом перемежения. Главная диагональ проходит с левого угла матрицы в правый. Эта диагональ начинается с 1 символа в первом блоке и заканчивается последним символом последнего блока матрицы. В примере на рис.7.5. интервал перемежения равен 2 блокам, а глубина перемежения – 15 блокам. Символы, которые считываются с главной диагонали матрицы, подчеркнуты. Ввод и считывание осуществляются непрерывно со сдвигом столбцов в матрице слева направо. Проверочные символы, которые обычно расположены в конце блока, тоже участвуют в перемежении.

При сверточном перемежении выпадение из нескольких блоков подряд преобразуется в длинный пакет ошибочных блоков с 1 или 2 ошибочными символами. Это позволяет полностью исправлять выпадения длиной 2  $I_p$  блоков подряд. На рис.7.6. показано, как выпадение длиной 4 блока с 32 ошибочными символами преобразуется в пакет из 18 ошибочных блоков подряд. В каждом из этих блоков по 1 или 2 ошибочных символа, которые могут быть исправлены декодером.



Следует особо подчеркнуть, что перемежение эффективно работает только при условии, что средний интервал между выпадениями существенно больше *глубины перемежения*. Если это условие не выполняется перемежение только ухудшит работу декодера и применять его нельзя.

При *внутриблочном перемежении* обычно производится несистематическая

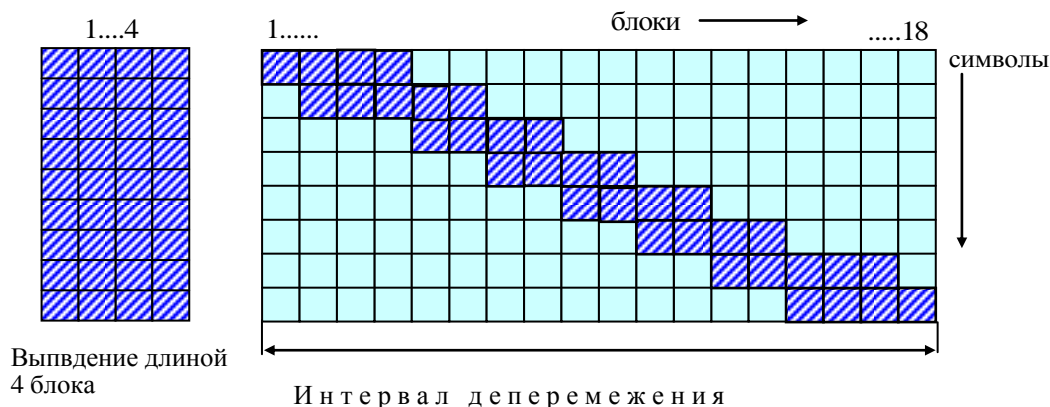


Рис.7.6. Преобразование выпадения из 4 ошибочных блоков в пакет из 18 ошибочных блоков при депережении

перестановка символов, младших и старших разрядов символов, а также осуществляется перестановка слов левого и правого звуковых каналов звуковых каналов. Этот вид перемежения предназначен для улучшения маскировки неисправленных ошибок.

### Системы кодирования с перемежением

Для обеспечения защиты от коротких и длинных выпадений без существенного повышения исправляющей способности кода в системах цифровой записи звука используются не просто коды, а системы кодирования, включающие в себя несколько ступеней кодирования и многократное перемежение символов в этих ступенях.

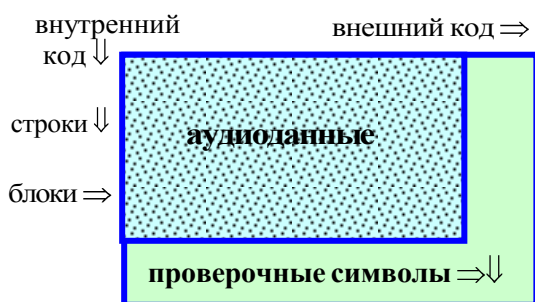


Рис.7.7. Product code (PC) (код-произведение)

Из таких систем чаще всего используются три: код – произведение с внутриблочным и межблочным перемежениями, каскадное кодирование с межблочным перемежением и комбинация этих двух систем. В системе код – произведение данные, подлежащие кодированию, собираются в ОЗУ в виде прямоугольной матрицы с организацией блоков по строкам и столбцам. Кодирование осуществляется по строкам и столбцам, а иногда еще и по диагонали

матрицы. Первое по времени кодирование называется внешним, второе – внутренним (рис.7.7). После кодирования производится перемежение. Во внутреннем и внешнем кодах чаще всего используются коды RS, но могут быть применены и другие коды.

В коде – произведение минимальное число исправляемых ошибок значительно больше, чем у отдельных кодов. Если один код может исправить  $t_1$  ошибочных символов, а второй –  $t_2$ , то код – произведение может исправить  $t$  ошибок.

$$t = 2t_1 \cdot t_2 + t_1 + t_2$$

Это значит, что при 4 проверочных символах по строкам и столбцам исправляются до 12 ошибок в блоке.

В системах с каскадным кодированием производится кодирование одним кодом, чаще всего кодом RS, несколько раз подряд (рис.7.8.). После внешнего кодирования осуществляется длинное перемежение, после внутреннего кодирования – короткое перемежение. Во второй ступени внутренним кодом кодируются и проверочные символы предыдущих ступеней. Это позволяет реализовать высокую избыточность кода без увеличения числа проверочных символов в ступенях выше 2... 4.

При декодировании каскадного кода все операции осуществляются в обратном

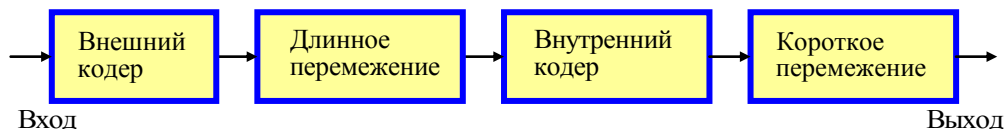


Рис.7.8. Каскадное кодирование

порядке. Сначала производится короткое перемежение. Главной задачей декодирования внутреннего кода является обнаружение ошибочных блоков и исправление коротких выпадений. Возможно также исправление стираний, введенных при канальном декодировании. Поэтому его исправляющая способность по ошибкам обычно невысока. Далее производится длинное перемежение для обеспечения возможности исправления длинных пакетов ошибок во 2 ступени декодирования..

Декодер внешнего кода 2 ступени прежде всего должен обнаруживать ошибки, не обнаруженные в 1 ступени декодирования, и исправлять длинные пакеты выпадений. Поэтому исправляющая способность декодера 2 ступени должна быть как можно выше. Иногда ему вводятся и функции исправления стираний, введенных в 1 ступени декодирования. При невозможности декодирования внешний декодер сам вводит стирания, которые далее могут быть исправлены с помощью маскировки.

### Маскировка ошибочных слов

При сколь угодно высокой исправляющей способности декодера всегда возможны кодовые ошибки, которые нельзя исправить. Они приводят к ошибкам в кодовых словах звуковых сигналов. При записи звука, в отличие от записи цифровых данных в компьютерах, в этих случаях можно *маскировать ошибки*, используя интерполяцию ошибочных слов. Этот процесс иллюстрируется рис.7.9., на котором номерами обозначены различные способы интерполяции.

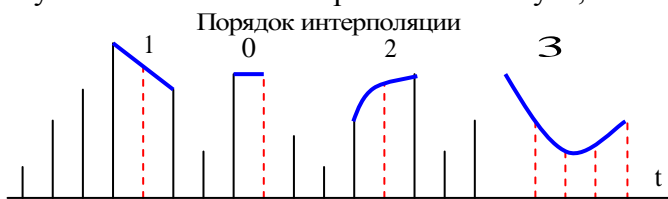


Рис.7.9. Интерполяция ошибочных выборок

Интерполяция может быть в виде мьютинга (4), при котором ошибочное слово заменяется нулевым значением с плавным спаданием и нарастанием звукового сигнала в течение 50 ... 60 слов. Она используется при длинных выпадениях. Возможна интерполяция нулевого порядка (2), при этом ошибочное слово заменяется предыдущим значением (удержание). Более часто используется интерполяция первого порядка (1), когда ошибочная выборка заменяется средним значением из двух ближайших выборок. Иногда используется полиномиальная интерполяция (3) второго и более высоких порядков, при которой отсутствующая выборка рассчитывается по нескольким значениям ближайших выборок. Поэтому она позволяет успешно интерполировать несколько неисправных слов подряд (рис.7.9).

Уровень шума интерполяции существенно зависит от порядка используемой интерполяции ошибок. На рис.7.10. приведены графики электрических уровней шума

интерполяции относительно нулевого (0,775 мВ) в зависимости от порядка интерполяции. Из них видно, что чем более высокого порядка интерполяция, тем меньше слышимость шума интерполяции. Этот шум всегда растет с повышением частоты интерполяции. На частотах интерполяции ниже 63 Гц уровень шума менее минус 100 дБ, даже при интерполяции 1-го порядка. Например, радиальная царапина на компакт диске имеет частоту интерполяции 3...10 Гц в зависимости от радиуса записи. В этой области частот даже при интерполяции 0 порядка уровень шума интерполяции ниже нулевого на 50 дБ. Интерполяция с помощью мьютинга слышна на всех частотах.

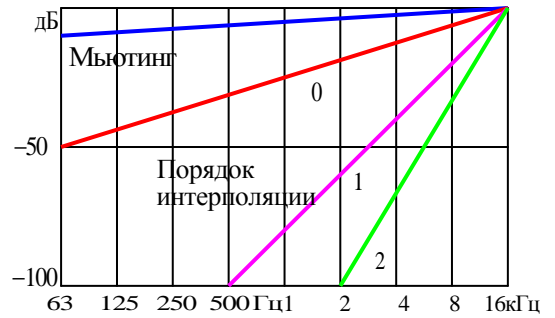


Рис.7.10. Графики уровня шума интерполяции

## 7.2. Канальная модуляция

При корректирующем кодировании число последовательных 1 и 0 двоичного кода ничем не ограничено, и спектр такого сигнала может простирается от 0 до тактовой частоты. Многие же цифровые тракты записи-воспроизведения не в состоянии пропустить без искажений такой широкий спектр частот. Для формирования сигнала записи канальная модуляция осуществляет преобразование спектра двоичной последовательности в соответствии с полосой пропускания сквозного тракта. Это достигается путем замены одних двоичных кодовых комбинаций другими, более оптимальными, поэтому такую модуляцию часто называют *канальным кодированием*, а сформированный при модуляции цифровой поток – *канальным кодом*.

Некоторые цифровые тракты не пропускают постоянную составляющую. В других трактах частотный диапазон от 0 до 50 кГц используется для передачи сигналов систем автоматического регулирования. Поэтому крайне желательно, чтобы в спектре канального кода было минимизировано значение постоянной составляющей и спектральных компонент в этом диапазоне частот.

После канальной модуляции сигнал записи представляет собой двухуровневую последовательность импульсов, в которой высокий уровень соответствует логической «1», а низкий - «0». В этой последовательности цифровая информация передается в виде дискретного изменения длительности импульсов и промежутков между ними. Эти

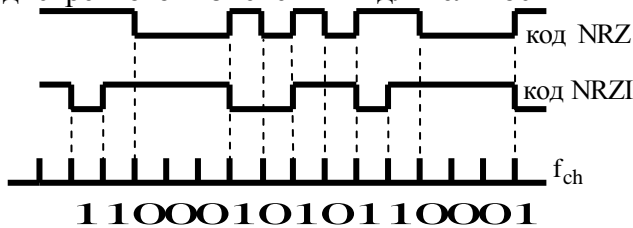


Рис.7.11. Канальные коды NRZ и NRZI

временные интервалы кратны периоду тактовой частоты и ограничены по длительности от минимального значения до максимального. На этом основании канальные коды часто называют RLL кодами (Run Length Limited Codes – коды с ограничением длительности временных интервалов). В качестве примера на **рис.7.11.** приведены временные диаграммы канального кода NRZ (Non Return to Zero), в котором обычно передается двоичная информация в цифровой схемотехнике.

Для обеспечения возможности демодуляции воспроизводимого цифрового сигнала канальный код должен обладать свойством самосинхронизации. Такое свойство означает, что при воспроизведении из сигнала канального кода может быть выделена тактовая частота, которая синхронизирует работу демодулятора. Без такой синхронизации демодуляция канального кода невозможна. Для выделения тактовой частоты необходимо, чтобы вероятность коротких интервалов случайной последовательности импульсов была максимальна. Код NRZ, например, таким свойством не обладает и поэтому для записи

данных не используется. Канальная модуляция позволяет также увеличивать плотность записи, повышать помехозащищенность и обнаруживать ошибки при декодировании.

В современной технике звукозаписи используются две основные группы канальных кодов: коды без избыточности и коды с избыточностью. В первой группе кодовые комбинации канального кода формируются с использованием некоторых логических правил. Например, в относящемся к этой группе коде NRZ, изменение уровня сигнала записи происходит только при смене кодовых комбинаций, когда после 1 следует 0 или наоборот. В коде NRZI (инверсный NRZ) изменение уровня сигнала происходит при каждой 1 (рис.7.11.). Канальные коды первой группы отличаются простотой технической реализации схем кодирования и декодирования, но большинство имеет значительную постоянную составляющую, что резко ограничивает возможности их применения в современной звукозаписи.

Канальные коды второй группы формируются путем организации бит данных в группы и добавления в эти группы избыточных символов, поэтому часто эти коды называются групповыми. В этих кодах в качестве вспомогательного часто используется код NRZI. Групповые коды могут быть фиксированной и переменной длины. Они отличаются тем, что одних кодах число символов в группе является величиной постоянной, а в других - оно может меняться.

*Групповое канальное кодирование фиксированной длины* заключается в том, что непрерывный поток двоичных данных разбивается на группы (символы) по « $m$ » бит. Затем их число в группе увеличивается до « $n$ » и такие символы и биты называются

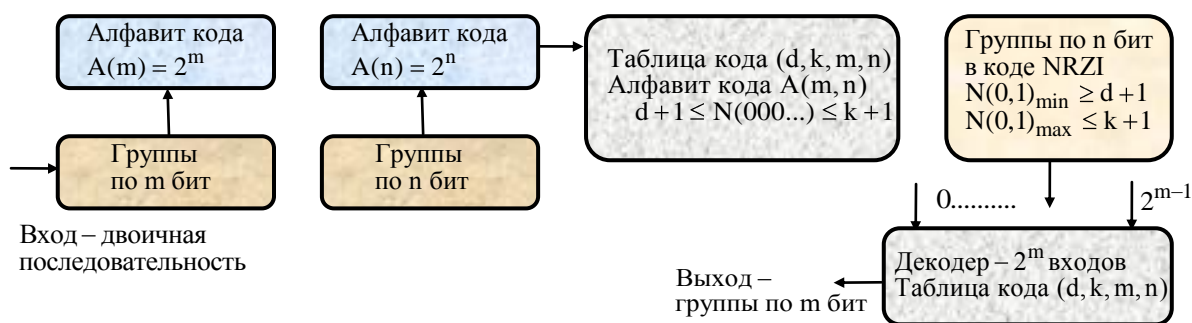


Рис.7.12. Групповое канальное кодирование

канальными. Таким образом, образуются два алфавита кода, число кодовых комбинаций в которых определяется равенствами:

$$A(m) = 2^m \text{ и } A(n) = 2^n.$$

Из общего алфавита  $A(n)$  с помощью компьютера выбираются только те кодовые комбинации, в которых между логическими «1» число последовательных нулей  $N(000\dots)$  не меньше  $d+1$  и не больше  $k+1$  (рис.7.12). Они образуют разрешенную часть алфавита  $A(m, n)$ , остальные кодовые комбинации составляют запрещенную часть алфавита  $A^*(n)$ . В цифровой аппаратуре записи звука обычно  $m = 8$  (1 байт), тогда  $A(m) = 256$ , значения  $n$  могут быть от 10 до 16. Поэтому возможно, что 256 разрешенных символов выбираются из 32768 кодовых комбинаций.

Параметры " $d$ " и " $k$ " определяют все основные характеристики канального кода. Групповой канальный код фиксированной длины обозначается как  $(d, k, m, n)$  или  $(d, k)$ , а еще чаще  $-m/n$ . Разрешенные кодовые комбинации сводятся в таблицу в специальном виде, называемом NRZI представление. Это значит, что двоичные табличные данные будут преобразовываться в последовательность временных интервалов переменной длительности с помощью этого кода. Поэтому в сигнале записи число последовательных единиц и нулей определяется неравенством

$$d + 1 \leq N(000\dots111\dots) \leq k + 1.$$

Таким образом, параметры  $d$  и  $k$  ограничивают число последовательных «0» в кодовой таблице и число последовательных «1» в коде записи.

Кодовые комбинации таблицы составляются так, чтобы длительности формируемых импульсов и интервалов между ними менялись от минимального до максимального значений  $c$

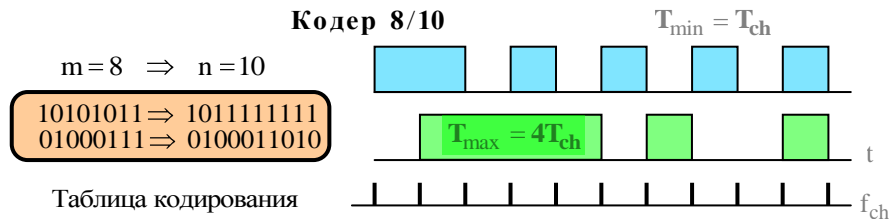


Рис. 7.13. Канальный код (0,3,8,10)

интервалом дискретности, который равен периоду тактовой частоты канального кода  $T_{ch}$ . Длительности этих интервалов определяются равенствами:

$$T_{\min} = T_{ch}(d + 1), \quad T_{\max} = T_{ch}(k + 1).$$

В качестве примера на рис. 7.13. приведем фрагмент таблицы для канального кода 8/10 (0,3,8,10), используемого в цифровых магнитофонах. Из таблицы двоичные данные считываются с тактовой канальной частотой в  $f_{ch}$  в коде NRZI. При этом каждой логической «1» в таблице соответствует изменение уровня сигнала вверх или вниз. В этом коде длительность временных интервалов :

$T_{\min} = T_{ch}$ ,  $T_{\max} = 4T_{ch}$ . Значения  $T_{\min}$ ,  $T_{\max}$  определяют верхнюю и нижнюю границы частотного спектра канального кода, в котором передается звуковая информация

$$f_{\min} = \frac{1}{2T_{\max}}, \quad f_{\max} = \frac{1}{2T_{\min}}.$$

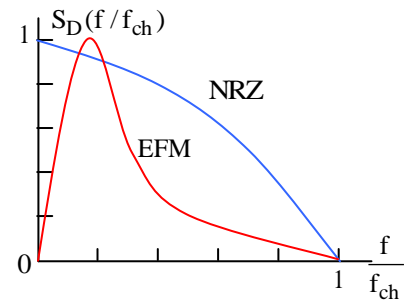


Рис. 7.14. Графики спектральной плотности мощности

Этот спектр состоит из дискретных компонент, число которых определяется параметрами канального кода. В тоже время при записи звуковых программ кодовые комбинации последовательности временных интервалов статистически независимы, и распределение их вероятностей описывается экспоненциальным законом. Это значит, что наиболее вероятны короткие временные интервалы и менее вероятны – длинные. Частота повторения кодовых интервалов может меняться в очень широких пределах, поэтому спектр такой случайной последовательности простирается от 0 до  $f_{ch}$  и канальный код характеризуется

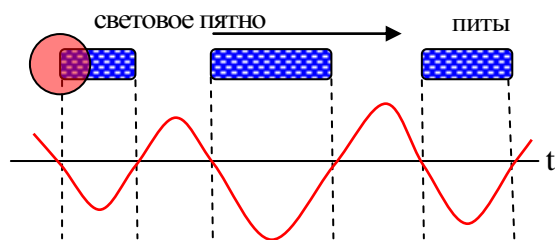


Рис. 7.15. Сигнал считывания

спектральной плотностью мощности  $S_D(f)$ .

На рис. 7.14. для сравнения приведены графики спектров кода NRZ и кода EFM (8/14), который применяется в формате CD. На них по оси абсцисс отложено нормированное значение частоты  $f/f_{ch}$ . Эти спектры отличаются прежде всего значением спектральной плотности в районе нулевых частот.

При воспроизведении записи амплитуды спектральных компонент определяются передаточной функцией головки считывания. Из-за конечных размеров считывающего элемента (зазор магнитной головки или диаметр светового пятна) эта функция всегда имеет завал высоких частот, поэтому в выходном сигнале присутствуют в основном

первые гармоники спектра воспроизводимого сигнала. По этой причине этот сигнал очень мало похож на последовательность прямоугольных импульсов (рис.7.15.). Восстановление исходной последовательности импульсов осуществляется с помощью компаратора или усилителя-ограничителя, которые формируют импульсы по моментам перехода воспроизведенного сигнала через нулевые значения.

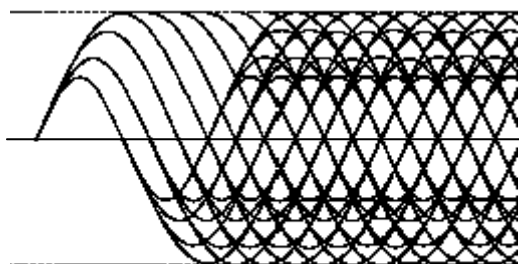


Рис.7.16. Глаз-диаграмма кода EFM

Критерием качества работы системы воспроизведения является глаз-диаграмма (eye-pattern), которая представляет собой осциллограмму воспроизведенного сигнала (рис.7.16.) в режиме синхронизации тактовой частотой  $f_{ch}$ . На этой осциллограмме видны все дискретные частоты спектра, их амплитуды и, главное, моменты перехода через нулевое значение. Спад амплитуд на высоких частотах определяет частотные искажения,

обусловленные передаточной функцией головки считывания. Если все переходы находятся на оси времени и сигнал четкий, ни частотной, ни фазовой коррекции не требуется. Наличие джиттера хорошо видно по размыванию области перехода через нулевое значение.

Во всех групповых канальных кодах существует проблема обеспечения выполнения требований в отношении параметров  $d$  и  $k$  на стыках канальных символов. Всегда возможно, что один канальный символ закончится самой длинной, но допустимой последовательностью «0», а следующий с нее начнется и тогда общая их длина превысит допустимую. В различных кодах эта проблема решается по-разному. Иногда вводят дополнительные разделительные канальные биты между канальными символами, часто используют 2 и более таблиц кодирования. В этом случае исходной кодовой комбинации соответствует несколько вариантов канальных символов. Выбирается тот, при котором значения  $d$  и  $k$  соответствуют заданным.

В табл.7.2. для удобства пользования сведены все расчетные формулы для основных характеристик групповых канальных кодов. Исходными для расчета данными являются период  $T_d$  и длины кодовых комбинаций  $m$  и  $n$ .

Таблица 7.2. Расчетные формулы канального кода

Характеристики	Формулы
Избыточность кода	$R_{ch} = m / (m - n)$
Скорость канального кода	$V_{ch} = m / n$
Период тактовой частоты канального кода	$T_{ch} = V_{ch} T_d$
Коэффициент самосинхронизации	$\vartheta = T_{max} / T_{min}$
Окно детектирования	$T_w = T_{ch}$
Коэффициент линейной плотности записи	$K_d = (d + 1) V_{ch}$
Вероятность не обнаружения ошибок	$P_{sch} = 2^{-(n-m)}$

**Самосинхронизация** – это возможность выделения из воспроизведенного сигнала тактовой частоты канального кода  $f_{ch}$ . Код считается самосинхронизирующимся, если коэффициент самосинхронизации  $\vartheta$  находится в пределах от 2 до 4. Этот коэффициент ограничивает диапазон изменения временных интервалов. Для обеспечения самосинхронизации канального кода при построении таблицы кода следует выбирать канальные символы с наибольшим числом логических «1». Тогда в коде записи частота переходов будет больше и коэффициент  $\vartheta$  увеличится.

**Спектр канального кода** можно менять в широких пределах путем изменения  $d$ ,  $k$ ,  $m$  и  $n$ . Если  $d$  стремится к нулю, то верхняя граница спектра расширяется до  $f_{max} = 1/2T_{ch}$ . Если  $k$  стремится к бесконечности, то нижняя граница спектра приближается к нулю. Если необходимо иметь подъем частотной характеристики в области высоких частот, то при построении таблицы кода следует выбирать канальные символы с наибольшим числом «1». Если требуется подъем низких частот, то, наоборот, число «1» в канальных символах должно быть меньше. Возможности формирования необходимой

формы огибающей спектра канального кода тем больше, чем выше его избыточность. Эти возможности существенно ограничиваются условием, что код должен быть самосинхронизирующимся.

**Окно детектирования** – это минимальный временной интервал, равный периоду канального кода, на который отличается один кодовый интервал от другого. Этот параметр определяет возможность декодирования кода в условиях действия шумов в канале записи. При воспроизведении канального кода шумы канала приводят к временным колебаниям фронтов импульсов, которые принято называть **джиттером**. Декодирование возможно без ошибки пока амплитуда джиттера не превышает половины окна детектирования. Такая возможность обеспечивается тем, что в декодере небольшие качания фронтов импульсов полностью устраняется с помощью специальной электронной схемы, называемой де-джиттер. Чем шире окно детектирования, тем выше помехозащищенность канального кода и меньше вероятность кодовых ошибок, но при этом снижается максимальная плотность записи.

**Коэффициент плотности** записи данных определяет, во сколько раз увеличивается плотность записи данных при использовании канального кодирования. Пока для самых

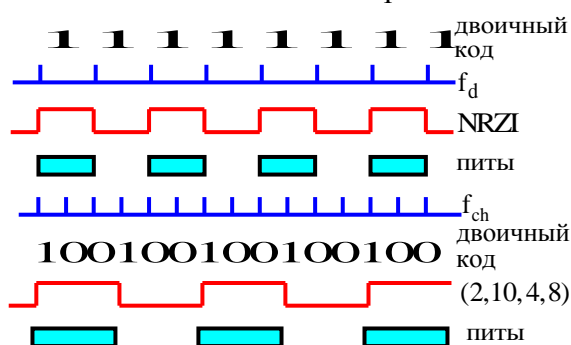


Рис.7.17. К вопросу повышения максимальной плотности записи при канальном кодировании

оптимальных канальных кодов эта плотность повышается не более чем в 2 раза. Следует пояснить, что только использование канального кода не повышает линейной плотности записи, просто возникает возможность записывать более скоростные цифровые потоки, что и приводит к возможности увеличения плотности записи.

Сказанное поясняется **рис.7.17**. На этом рисунке показан вариант записи непрерывной двоичной последовательности 111111... в кодах NRZI и (2,4,4,8). Код NRZI не имеет избыточности и плотность записи он не

повышает. В этом коде запись данных производится с частотой их следования  $f_d$ . В коде (2,4,4,8) тактовая  $f_{ch}$  вдвое выше, двоичный поток принимает вид 1001001001..., а длительность временных интервалов и питов увеличивается в 1,5 раза. При одной и той же разрешающей способности головки считывания, это позволяет также в 1,5 раза увеличить скорость записываемого потока данных. Это и приводит к увеличению линейной плотности записи.

**Цифровая сумма DSV(Digital Sum Value)** характеризует постоянную составляющую спектра и рассчитывается как текущая сумма последовательности временных интервалов

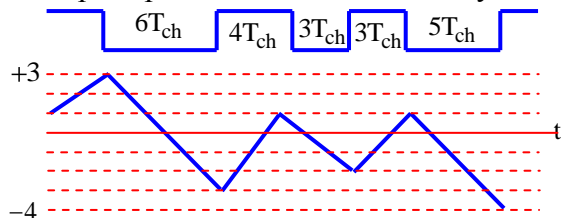


Рис.7.18. DSV в канальном коде (2,10,4,8)

в коде NRZI в предположении, что высокий уровень равен +1, а низкий –1. Считается, что постоянной составляющей нет, если эта сумма не превышает  $\pm 7$ . На **рис.7.18** показано, как меняется DSV в функции времени. Как видно, она приводит к возникновению низкочастотных помех. Во всех системах звукозаписи наличие постоянной составляющей в сигнале

канального кода крайне не желательно. При магнитной записи постоянную составляющую не пропускает магнитная головка записи, в оптических системах частотный диапазон от 0 до 20 кГц используется для передачи сигналов управления систем автоматического регулирования, таких как фокусировка и радиальное слежение за дорожкой. Значительная постоянная составляющая нарушает работу этих систем.

Минимизация постоянной составляющей реализуется несколькими способами. Одним из решений этой проблемы является введение разделительных бит между канальными

символами, как это делается в формате CD. В этом случае производится непрерывный расчет текущего значения DSV, и кодовые значения разделительных бит выбираются так, чтобы постоянная составляющая уменьшалась. Иногда минимизируется среднеквадратическое значение DSV. Если требуется обеспечить подавление постоянной и переменной составляющих в области нулевых частот, то тогда контролируется значение кодовой цифровой суммы CDS (Code Digital Sum) всех канальных символов, как это делается в формате DAT. CDS равна DSV, рассчитанной для каждого канального символа. В этом случае в алфавит кода входят только символы, у которых  $CDS = 0$  или  $CDS \leq \pm 2$ . Постоянная составляющая подавляется тем сильнее, чем больше кодовых комбинаций с  $CDS = 0$ .

*Размножение ошибок* является серьезной отрицательной стороной канального кодирования. Оно происходит из-за того, что если в кодовом символе ошибочен хотя бы один бит, ошибочным считается весь символ.

*Воспроизведение записи.* При воспроизведении сигнала канального кода выделяется его тактовая частота  $f_{ch}$  и формируется импульсная последовательность по моментам перехода воспроизведенного сигнала через ось времени. Этот сигнал синхронизируется тактовой частотой. Пока амплитуда джиттера не превышает половины окна детектирования, равного  $T_{ch}$ , благодаря синхронизации кодовые ошибки не возникают.

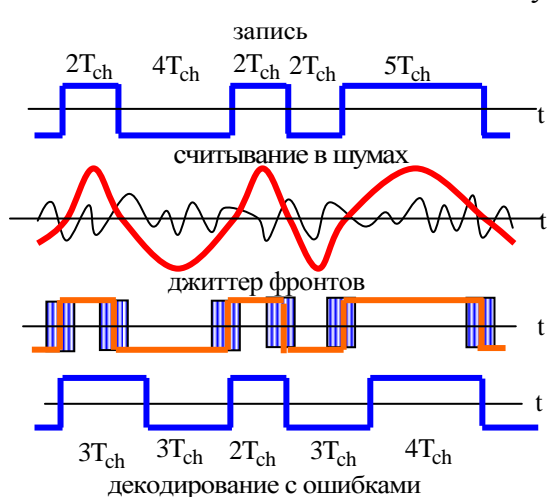


Рис.7.19. Возникновение кодовых ошибок при демодуляции канального кода

Когда это условие не выполняется, тактовая синхронизация приводит к дискретным ошибкам, при которых одна кодовая комбинация переходит в другую (разрешенную или не разрешенную), как это показано на рис.7.19.

*Кодирование канального кода* производится с помощью электронной таблицы. Такая таблица реализуется обычно с помощью ПЗУ (постоянное запоминающее устройство), в котором записано  $m$   $n$ -разрядных разрешенных кодовых комбинаций. ПЗУ имеет  $m$  входов и один выход. Адресами входов являются  $m$  – разрядные кодовые значения алфавита  $A(m,n)$ . Каждому входу соответствует свой канальный

символ длиной  $n$ .

*Декодирование канального кода* производится также с помощью ПЗУ, в котором восстанавливается исходный двоичный код. В этом ПЗУ записаны все кодовые комбинации алфавита  $A(m,n)$ . Адресами входов являются все разрешенные  $n$ -разрядные кодовые комбинации. Кодовая комбинация входного канального символа сравнивается с адресами всех входов и если такая находится, на выходе декодера выдается соответствующая ей кодовая комбинация из  $m$  бит данных. Если она не находится, это значит, что данная комбинация является запрещенной и, следовательно, произошла кодовая ошибка.

Такие ошибки обнаруживаются в канальном декодере с вероятностью 1. В этом случае ошибочный канальный символ «стирается» - ему придается флаг (индекс) ошибки. Он может быть исправлен в декодере корректирующего кода. Если ошибка приводит к разрешенной кодовой комбинации, то при декодировании канального кода она не обнаруживается. Вероятность такого события оценивается значением  $P_{sch}$ . Возможность обнаруживать ошибочные канальные символы определяется избыточностью канального кода  $R_{ch}$ , которая может быть от 0,2 до 0,5. Вероятность их не обнаружения  $\hat{P}_{sch}$  находится в пределах от  $0,25$  до  $1,5 \cdot 10^{-2}$ .



## 8. Оптическая запись звука

### 8.1. Стандарты и форматы оптической записи

В настоящее время цифровая запись звука на оптические диски производится в трех стандартах: CD-DA, Super Audio CD и DVD-Audio, основные параметры которых приведены в табл.8.1. и 8.2. Стандартом определяются характеристики носителя записи, способ аналого-цифрового преобразования звуковых сигналов, корректирующее кодирование, канальная модуляция и организация цифровых данных. К форматам записи относятся оптические диски CD-R, DVD-R, CD-RW, DVD-RW,

Параметры	Стандарты		
	CD-DA	DVD-A	SACD
Число звуковых каналов	2	2... 6 (5.1)	
Модуляция при АЦП	ИКМ		SDM
Частота дискретизации, fs, кГц	44,1	48/96/192 44,1/88,2/176,4	2822
Число разрядов	16	16/20/24	1
Скорость аудио данных, Va, Мб/с	1,41	1,536...9,6	2,822 × 6
Емкость диска, Мбайт	730	4700	4700+730
Канальный код	8/14 (2,10,8,14)	8/16 (2,10,8,16)	
Корректирующий код	2RS (CIRC)	2RS-PC (CIRC plus)	
Компрессия цифрового потока	нет	Нет/есть	
Тактовая частота канального кода, МГц	4,32	26,12	
Линейная скорость записи, м/с	1,3	3,49	
Время записи слоя, мин	60...80	62...816	74

HD-DVD и BR (Blue-Ray). Диски различных форматов отличаются тем, что на них с равным успехом может быть произведена запись звука, видео или просто любых цифровых данных. Запись производится в соответствии со стандартами на эти диски.

Для обеспечения возможности воспроизведения дисков различных стандартов и форматов на одном проигрывателе механические и оптические характеристики всех дисков одинаковы (табл.8.2.). Запись осуществляется по спирали против часовой стрелки с постоянной линейной скоростью. В рамках этих характеристик диски могут быть двухсторонними и двухслойными, что многократно увеличивает время воспроизведения. Запись на первом слое производится от центра к краю, а на втором - от края к центру.

На стандартных оптических дисках осуществляется рельефно-фазовая запись, при

Диаметр, мм	120
Толщина, мм	1,2±0,1
Диаметр отверстия, мм	15
Радиус зоны записи данных, мм	24...58
Масса, Г	13...20
Коэффициент преломления	1,55

которой цифровые данные регистрируются в виде микроуглублений (питы-pits) вдоль спиральной дорожки. Такая запись производится только в заводских условиях, и диски массово тиражируются примерно также как виниловые грампластинки. На оптических дисках различных форматов осуществляется амплитудная запись, при

которой цифровые данные регистрируются в виде локальных изменений отражающей способности (меток) материала покрытия вдоль спиральной дорожки. Такая запись производится на профессиональных или бытовых рекордерах.

Стандарт CD-DA был разработан совместно фирмами Philips и Sony в 1982 году. Это первая в мире система с бесконтактным воспроизведением записи звука с диска отраженным сфокусированным лазерным излучением. Принципиально новыми элементами системы являются оптический диск и оптическая система воспроизведения записи. В настоящее время этот стандарт технически устарел и используемая в нем 16-

ти разрядная ИКМ с частотой дискретизации 44,1 кГц не обеспечивает необходимое качество не только студийной записи, но и бытовой.

Новые стандарты звукозаписи, такие как Super Audio CD и DVD-Audio, в значительной мере базируются на CD системе. Это касается характеристик оптического диска, принципов построения оптической системы воспроизведения, методов канального и корректирующего кодирования. Главное, что в новых системах запись возможна со значительно более высокой плотностью и более высокими техническими характеристиками. Отличия, в основном, касаются методов аналого-цифрового преобразования. Используются более совершенные канальный код 8/16 и корректирующий код 2RS-PC.

Увеличение емкости одного информационного слоя оптического диска до 4,7 Гбайт в новых стандартах достигнуто за счет уменьшения длины волны лазерного излучения и числовой апертуры объектива проигрывателя. Такая емкость дисков позволила реализовать 6-канальную запись звука с пространственным звучанием с динамическим диапазоном 120 дБ и полосой до 100 кГц. Однако, необходимое время записи многоканальных программ 74 мин достигается только с применением компрессии звука.

Т а б л и ц а 8.3. Время записи DVD дисков

Характеристики АЦП	Va, Мб/с	Время, мин компрессия	
		нет	есть
		96кГц-20 бит-5 кан	9,600
192кГц-24 бит-2 кан	9,216	64	128
48кГц-16 бит-2 кан	1,536	408	816
48кГц-24 бит-6 кан	6,912	86	156

Это связано с тем, что максимальная скорость цифрового потока не может превышать половины тактовой частоты канального кода 26,12 МГц, в DVD-A эта скорость составляет 9,6 Мбит/с. Для этих целей разработаны методы компрессии цифровых данных без потерь, как это делается к компьютерной технике. В стандарте SACD используется DST(Direct Stream Transfer) компрессия, а в DVD-Audio – MLP (Meridian Lossless Packing) компрессия. Стандарт DVD-A значительно более гибкий, он допускает возможность использовать и другие способы компрессии: Dolby Digital, DTS и MPEG. Имеется возможность выбирать частоту дискретизации и число разрядов, поэтому длительность записи в одном слое может быть от 62 до 408 минут (табл.8.3.). При двух слоях это время удваивается, а для двухстороннего диска – учетверяется.

Фирмы Philips и Sony предложили новую систему оптических звуковых дисков SACD, которая полностью совместима со старой системой CD. Такое решение связано с тем, что у населения имеется на руках свыше 10 миллиардов компакт-дисков и не менее 500 миллионов проигрывателей. Поэтому диски SACD имеет те же размеры и почти такую же зону записи, что и CD диски, но только они всегда двухслойные. Фактически эта два склеенных диска толщиной 0,6 мм. На одном диске запись стереофонической музыкальной программы в стандарте CD, а на другом – 6 канальная запись звука (5.1) с пространственным звучанием в стандарте SACD. В проигрывателе используются две оптические системы воспроизведения с отдельными объективами и лазерами. Излучение лазера с длиной волны 0,78 мкм фокусируется объективом с числовой апертурой 0,45 на отдаленный слой, а другой лазер с длиной волны 0,65 мкм и апертурой объектива 0,6 фокусируется на ближний слой с высокой плотностью записи.

Главная особенность стандарта SACD в том, что в нем при аналого-цифровом преобразовании используется не импульсно-кодовая модуляция, а одноразрядная сигма-дельта модуляция с частотой дискретизации 2,8 МГц. При такой модуляции очень сильно уменьшаются шумы квантования, полоса звуковых частот расширяется до 100 кГц и обеспечивается SNR до 120 дБ. Новым является предложенная структура организации непрерывного цифрового потока без разделения его на слова, символы и блоки, которая известна под логотипом DSD (Direct Stream Digital). В этом формате для увеличения времени записи также используется специальная компрессия скорости цифрового потока DST .

Главная особенность стандарта SACD в том, что в нем при аналого-цифровом преобразовании используется не импульсно-кодовая модуляция, а одноразрядная сигма-дельта модуляция с частотой дискретизации 2,8 МГц. При такой модуляции очень сильно уменьшаются шумы квантования, полоса звуковых частот расширяется до 100 кГц и обеспечивается SNR до 120 дБ. Новым является предложенная структура организации непрерывного цифрового потока без разделения его на слова, символы и блоки, которая известна под логотипом DSD (Direct Stream Digital). В этом формате для увеличения времени записи также используется специальная компрессия скорости цифрового потока DST .

Форматы записи CD-R и DVD-R предназначены для однократной записи и используются для архивной записи звука. Первоначально стандарты на эти диски существенно отличались по техническим характеристикам от стандартов CD-DA и DVD. В настоящее время различий практически нет. Форматы записи CD-RW и DVD-RW предназначены для многократной записи и используются для редактирования фонограмм.

Завершена разработка еще двух форматов оптической записи звука под названием Blue-ray Disc и HD-DVD. Записываемые диски этих форматов имеют диаметр 120 мм, но запись производится лазером с длиной волны излучения всего 405 нм и используется объектив с числовой апертурой 0,85 и 0,6 соответственно, что позволило существенно уменьшить диаметр считывающего пятна света. В формате HD-DVD емкость одного слоя составляет 15 Гбайт, в формате Blue-Ray Disc - 23,3 Гбайт. Ведутся работы по дальнейшему повышению емкости таких дисков до 50 и выше Гбайт. Для этого обрабатывается технология увеличения числа слоев от 2 до 6.

## 8.2. Физические основы оптической записи

Оптическая запись цифровых данных и воспроизведение этих данных осуществляются лазерным излучением, сфокусированным в пятно микронных размеров. Поляризованное лазерное излучение фокусируется с помощью микро объектива, основным параметром которого является числовая апертура  $NA$  (рис.8.1), равная

$$NA = \varepsilon \cdot \sin \Psi_0,$$

где  $\Psi_0$  – угол сходимости светового пучка при фокусировке,  $\varepsilon$  - коэффициент преломления среды. Как видно, числовая апертура объектива существенно зависит от коэффициента преломления материала НЗ, поэтому в справочных данных всегда указывается апертура объектива с учетом этого коэффициента. В воздухе  $\varepsilon=1$ , для материала подложки оптических дисков  $\varepsilon=1,55$ .

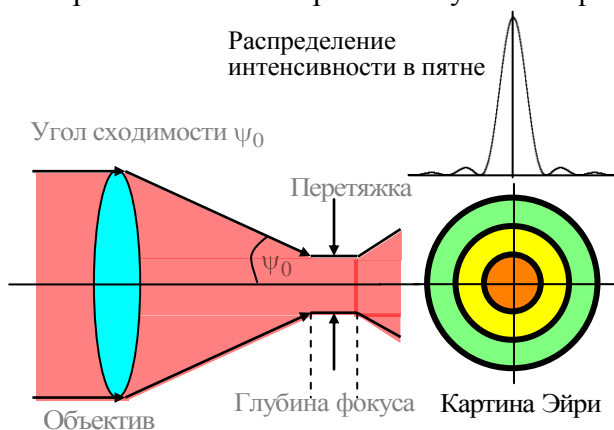


Рис.8.1. Фокусировка лазерного излучения

Сфокусированное излучение лазера имеет распределение интенсивности в поперечном сечении, описываемое функцией Бесселя (рис.8.1). Под микроскопом видна дифракционная картина Эйри в виде

Таблица 8.4. Фокусировка лазерного излучения				
Параметры	CD	DVD, SACD	HD-DVD	Blue Ray
Числовая апертура объектива	0,45	0,6	0,65	0,85
Разрешение, лин/мм	1154	1846	3202	4197
Длина волны лазера, нм	780	650(635)	405	405
Диаметр пятна по уровню 0,5, мкм	1,06	0,66	0,38	0,29
Глубина фокуса, мкм	± 1,9	± 0,9	± 0,48	± 0,28

кольцевых зон с различной интенсивностью света, разделенных черными кольцами, где интенсивность равна нулю. Диаметр сфокусированного пятна  $\phi$  по первому черному кольцу этой картины рассчитывается по формуле

$$\phi = 1,22 \frac{\lambda}{NA}.$$

Более часто диаметр считывающего пятна света определяется на уровне 0,5 по интенсивности, тогда

$$\phi_{0,5} = 0,61 \frac{\lambda}{NA}$$

Результаты расчетов по этим формулам для различных форматов записи приведены в табл.8.4.

Разрешающая способность объектива характеризует какое число линий на миллиметр длины способен различить объектив и она определяется равенством

$$\eta = \frac{2NA}{\lambda}$$

Сфокусированное излучение имеет перетяжку, где излучение является плоской волной и поэтому диаметр светового пятна почти не меняется (рис.8.1). Длина этой перетяжки называется *глубиной фокуса* и она связана с апертурой объектива квадратичной зависимостью

$$G = \frac{\lambda}{NA^2}.$$

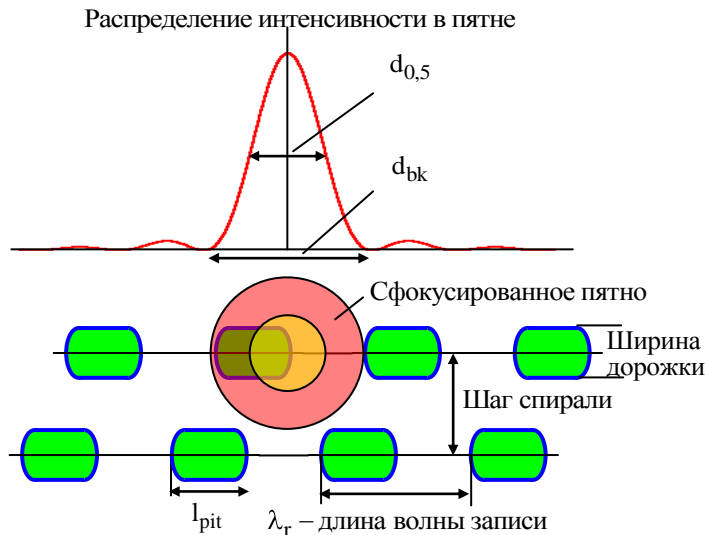


Рис.8.2. Считывание оптической записи с максимальной плотностью

Глубина фокуса определяет необходимую точность работы системы автоматической фокусировки. Когда ошибка фокусировки превышает половину глубины фокуса, диаметр светового пятна начинает быстро увеличиваться. Это приводит к резкому снижению разрешающей способности объектива и, как следствие, к уменьшению глубины модуляции считываемого сигнала.

На рис.8.2.. приведена схема считывания оптической записи сфокусированным лазерным излучением. Диаметр светового пятна по уровню 0,5 лишь немного превышает ширину дорожки и с его помощью осуществляется воспроизведение, если длина питов (меток) не меньше этого диаметра.

Приведенному распределению интенсивности вдоль дорожки соответствует точно такое же распределение и поперек. Поэтому считывание сопровождается помехами от соседних дорожек и соседних пит. Чем выше плотность записи, тем эти помехи больше.

Плотность оптической записи и емкость диска существенно зависят от выбора канального кода, который прежде всего определяет минимальную длину питов

$$l_{pit\ min} = \frac{(d+1) \cdot V_r}{f_T},$$

где  $d$  – параметр канального кода, определяющий минимальный временной интервал кода,  $f_T$  – тактовая частота канального кода,  $V_r$  – линейная скорость записи.

Линейная плотность записи (вдоль дорожки)

$$D_{lin} = \frac{K_{dch}}{l_{pit\ min}}, \quad K_{dch} = \frac{m \cdot (d+1)}{n},$$

где  $K_{dch}$  – коэффициент плотности канального кода,  $m/n$  – скорость канального кода.

Поверхностная плотность зависит от шага спиральной дорожки

$$\Delta_{sp} = (1 \dots 1,5) \cdot d_{0,5}$$

и она определяется равенством

$$D_s = \frac{D_{lin}}{\Delta_{sp}}$$

Существуют два значения емкости оптического диска. Первая называется физической и рассчитывается по формуле

$$C_{dt} = D_s \cdot \pi \cdot (R_{\max}^2 - R_{\min}^2)$$

где  $R_{\max}$  и  $R_{\min}$  - максимальный и минимальный радиусы зоны записи. Вторая называется «емкость пользователя». Она связана с максимальной скоростью цифрового потока аудио данных и временем записи

$$C_{ad} = V_{ad} \cdot T_{rec}$$

Форматы записи	Размеры питов		Плотность записи		Емкость слоя, ГБ	
	Длина min	Шаг Δ	$D_{lin}$ Кб/мм	$D_s$ Мб/мм кв	Data	Audio
CD	0,83	1,66	1,69	1,08	1,1	0,84
DVD-A	0,4	0,74	3,75	5,07	5,55	4,7
HD-DVD	0,23	0,4	6,52	16,3	17,8	15
BlueRay	0,16	0,32	8,75	27,3	29,9	25

В таблице 8.5. приведены данные по плотности записи основных форматов оптических дисков.

Воспроизведение оптических дисков осуществляется отраженным световым пучком. Поэтому основной задачей схемы считывания записи является разделение прямого и отраженного пучков с минимальными потерями. Упрощенная схема оптической системы, предназначенная только для воспроизведения записи с оптических дисков *отраженным световым пучком*, приведена на рис.8.3. Лазерное излучение, которым осуществляется воспроизведение записи, обычно имеет *круговую поляризацию*, при которой составляющие с горизонтальной и вертикальной поляризацией равны и сдвинуты по фазе на  $90^0$ .

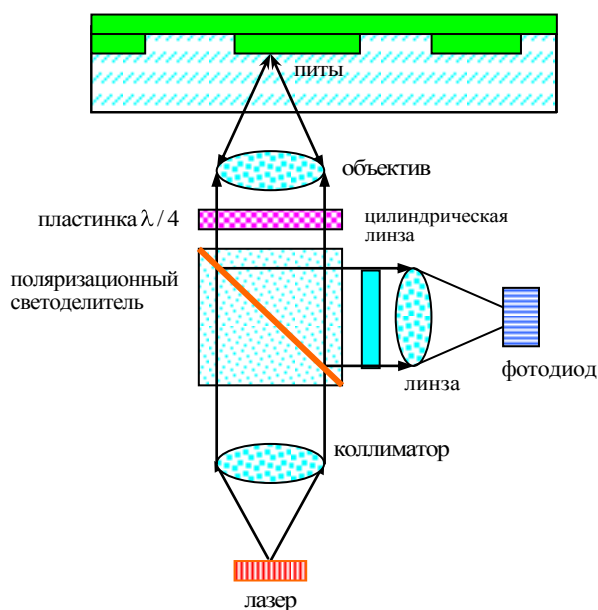


Рис.8.3. Оптическая система проигрывателя

Поэтому разделение световых пучков может осуществляться с помощью *поляризационного светоделителя*. Такой светоделитель имеет почти 100% коэффициент пропускания для одной составляющей поляризованного лазерного излучения и также почти 100% коэффициент отражения для другой составляющей. Использование явления поляризации света для разделения прямого и отраженного световых пучков позволяет резко снизить потери в оптической системе и становится возможным осуществлять воспроизведение оптической записи полупроводниковым лазером мощностью всего 0,5..1 мВт.

Излучение полупроводникового лазера сконцентрировано в эллиптическом конусе. Угловая полуширина в одной плоскости равна  $\pm 5^0$ , а в другой -  $\pm 20^0$ . Коллиматор оптической системы выполняет две функции. Он формирует параллельный пучок света минимальной расходимости и преобразует эллиптический конус излучения в круговой. Четвертьволновая пластинка оптического тракта изменяет угол поляризации светового пучка на  $45^0$ .

В приведенной схеме через четвертьволновую пластинку световой пучок проходит дважды: сначала, когда он фокусируется на носитель записи и затем при отражении. Каждый раз четверть-волновая пластинка меняет угол поляризации излучения на  $45^0$ . В результате она осуществляет поворот угла поляризации на  $90^0$  и таким образом преобразует горизонтальную поляризацию в вертикальную. Поэтому отраженный

световой пучок поворачивается на угол  $90^0$  и через дополнительную линзу фокусируется на фотоприемник.

Для отражения светового пучка на оптический диск наносится металлическое или специальное поляризационное покрытие. При воспроизведении записи с двухслойных дисков обычно производится перефокусировка лазерного излучения с одного слоя на другой. В стандарте SACD для воспроизведения записи используются две оптические головки с разными объективами и лазерами, параметры которых приведены в табл. 8.3. Излучение одного лазера фокусируется на дальний слой, а излучение другого лазера фокусируется на ближний слой. В ближнем слое в качестве отражающего используется поляризационное покрытие на основе кремния, которое излучение с одной длиной волны излучения полностью отражает, а излучение с другой длиной волны полностью пропускает. В дальнем слое используется алюминиевое покрытие с высоким коэффициентом отражения (70-80%), как это предусмотрено в стандарте CD.

**Оптическая передаточная функция** (ОПФ) объектива определяет зависимость глубины амплитудной модуляции выходного сигнала фотоприемника (ФП) от пространственной частоты  $\eta$  воспроизводимой цифровой фонограммы. ОПФ существенно зависит от размера и формы фотоприемника. Если площадь фотоприемника совпадает по размерам с отраженным на него световым пятном, то ОПФ имеет вид 1, приведенный на рис.8.4. При этом ее пространственная частота среза  $\eta_{cut}$  определяется равенством

$$\eta_{cut} = \frac{2NA}{\lambda}$$

Для системы CD  $\eta_{cut} = 1125$  пер/мм. Частота среза воспроизводимого электрического сигнала  $f_{elc}$ , на которой глубина модуляции равна нулю, зависит от линейной скорости записи

$$f_{elc} = \eta_{cut} \cdot V_{rec}$$

При  $V_{rec} = 1,25$  м / с имеем  $f_{elc} = 1,4$  МГц.

Передаточная функция на графике имеет такой вид только при условии, что отсутствуют aberrации оптической системы. Ошибки фокусировки приводят к искривлению формы графической зависимости ОПФ и уменьшению разрешающей способности (кривая 2).

Если размеры ФП малы по сравнению с диаметром светового пятна нулевого порядка, тогда частота среза ОПФ уменьшается вдвое и она определяется графиком вид 3 на рис.8.4. При других размерах ФП графики ОПФ будут иметь вид промежуточный между 1 и 3. Иногда для считывания используются два ФП, которые располагаются в зонах перекрытия 0-го и первых порядков дифракции. В этом случае считываемый сигнал формируется с помощью дифференциального усилителя и тогда ОПФ определяется графиком 4. Такой вариант построения оптической схемы возможен, когда нет необходимости воспроизводить низкие частоты.

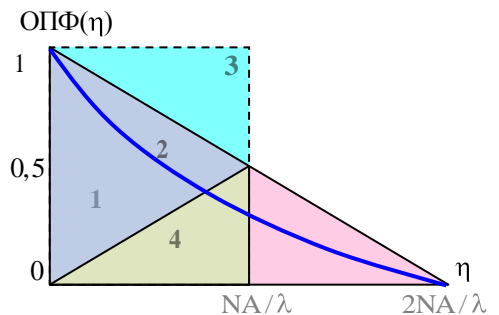


Рис.8.4. Оптическая передаточная функция объектива с фотоприемником

### 8.3. Рельефно-фазовая запись

В соответствии со стандартами CD-A, DVD-A и SACD цифровая звукозапись на оптических дисках производится вдоль воображаемой спиральной дорожки в виде последовательности микро углублений поверхности диска из оптически прозрачного поликарбоната с коэффициентом преломления 1,55. Эти углубления называются *питами*, все они имеют одинаковую ширину и глубину, а цифровые данные регистрируются в виде дискретных изменений длины этих пит и промежутков между ними в соответствии с используемым канальным кодом.

Запись производится с одной стороны диска, а считывание осуществляется с противоположной стороны сквозь достаточно толстую прозрачную основу диска сфокусированным лазерным излучением *на отражение*. При воспроизведении такой записи важную роль играет *глубина пит*, из-за которой возникает отличие длин оптических путей отраженного света от углублений и поверхности диска. Это создает фазовую модуляцию считывающего светового пучка. Так как эта модуляция связана с изменениями рельефа поверхности, то такая запись часто называется *рельефно-фазовой*.

Фактически это механический способ записи, каким производилась первичная запись грампластинок с помощью резца. Он обеспечивает наибольшее время хранения дисков, так как на них не оказывает влияния сильные магнитные или электрические поля, не возникает процессов окисления или коррозии. Практически такая запись реализуется путем тиражирования дисков с использованием никелевой матрицы, как это делалось ранее при изготовлении виниловых грампластинок.

Глубина фазовой модуляции максимальна, когда фазовый сдвиг отраженных лучей от пит и промежутков между ними  $\varphi$  равен  $\pi$ . В соответствии с законами геометрической оптики, требуемая для этого глубина пит  $h_p$  рассчитывается по формуле

$$h_p = \frac{\varphi \cdot \lambda}{4\pi\varepsilon},$$

где  $\varepsilon$  – коэффициент преломления основы оптического диска,  $\lambda$  – длина волны излучения лазера. Если используется лазер с инфракрасным излучением ( $\lambda = 780$  нм) и материал диска имеет коэффициент преломления 1,55, тогда  $h_p = 126$  нм.

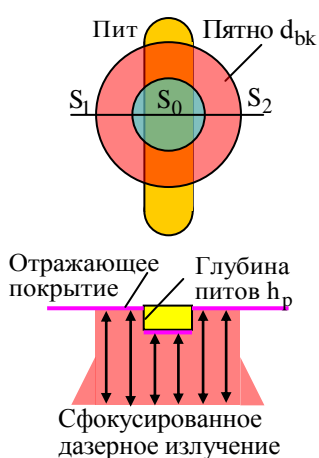


Рис.8.5. Разность хода оптических лучей при отражении от пит

Приведенная формула верна, когда размеры пит существенно больше диаметра считывающего пятна. Когда они соизмеримы, при расчетах надо учитывать волновую природу света и тогда необходимая глубина пит может быть как меньше, так и больше расчетной. Чем меньше их размеры и короче длина волны излучения лазера, тем больше могут быть расхождения. Фактическая глубина пит на CD диска равна 0,11 мкм, а на дисках DVD-A и SACD – 0,16 мкм. Она подбирается экспериментально по максимуму глубины модуляции. На **рис.**

**8.5** запись представлена в виде разреза поперек дорожки, чтобы можно было показать глубину пит и разность хода отраженных световых лучей.

Второе необходимое условие для считывания рельефно-фазовой записи, это равенство световой энергии, отраженной непосредственно от пита ( $S_0$ ) и от окружающей его поверхности диска ( $S_1 + S_2$ ). Когда считывающее пятно находится посередине пита, на фотоприемнике происходит интерференция противфазных световых пучков, и значение тока приближается к нулю. При этом происходит преобразование фазовой модуляции света в амплитудную модуляцию тока фотоприемника с максимальным значением глубины модуляции. Для того чтобы энергии противфазных световых пучков были равны необходимо, чтобы диаметр считывающего пятна света по уровню 0,5 был больше ширины пит почти в два раза. Это объясняется неравномерным распределением интенсивности в сфокусированном пятне света. Если диаметр пятна равен или меньше ширине дорожки, то глубина фазовой модуляции будет очень мала, и таким пятном воспроизвести рельефно-фазовую запись невозможно.

Третьим условием считывания рельефно-фазовой записи является выбор оптимальных размеров и формы фотоприемника (ФП), располагаемого в плоскости изображения объектива. Дорожка с питами представляет собой *нерегулярную фазовую отражающую решетку*, на которой происходит дифракция света. Дифракция вдоль направления дорожки с питами наиболее сильно проявляется, когда световое пятно находит и сходит с пита, а также когда длина пита близка к радиусу светового пятна.

При дифракции светового пучка света возникают 2 и более дополнительных отраженных световых пучков. Они называются порядками дифракции (0, ±1, ±2 и т.д.).

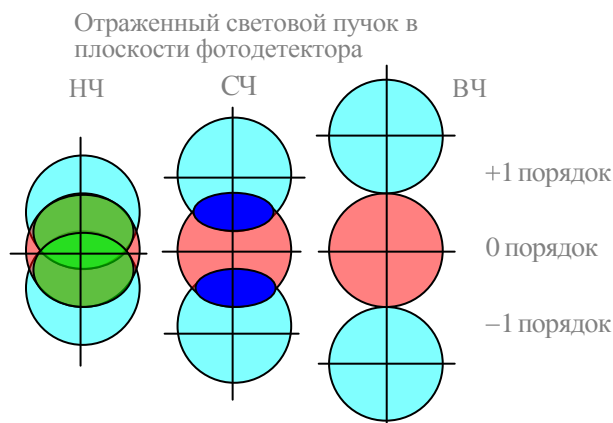


Рис.8.6. Интерференция дифракционных порядков в плоскости фотоприемника

Для фазовой решетки интенсивность этих порядков затухает медленно, на скорость затухания сильное влияние оказывает глубина пит. Угол отклонения этих пучков  $\psi_d$  зависит от пространственной частоты  $\eta$  фазовой решетки и длины волны лазерного излучения

$$\Psi_d = \arcsin \lambda \eta .$$

Чем больше пространственная частота, тем больше угол отклонения. Фазовый сдвиг между 0 и ±1 порядками определяется формулами:

$$\Delta\varphi_{+1} = \pi + 2\pi V_{rec} \cdot \eta \cdot t ,$$

$$\Delta\varphi_{-1} = \pi - 2\pi V_{rec} \cdot \eta \cdot t ,$$

где сдвиг  $\pi$  обусловлен глубиной пит. Как видно, переменные составляющие фазы в двух зонах перекрытия сдвинуты по фазе на угол  $\pi$ .

При больших периодах решетки световые пучки всех порядков почти полностью перекрываются. С уменьшением периода решетки зона перекрытия уменьшается (рис.8.6.) и при некотором значении  $\eta$  она становится равной нулю (перекрытия нет) и запись перестает воспроизводиться. В схеме воспроизведения на отражение объектив пропускает нулевой порядок дифракции, и только те высшие порядки дифракции, которые с ним перекрываются. Остальные высшие порядки пространственно объективом фильтруются.

В зонах перекрытия возникает *интерференция* между 0 и более высокими порядками дифракции (рис.8.6.). При этом фазовая модуляция светового потока преобразуется в амплитудную модуляцию тока фотоприемника. Если размеры фотоприемника совпадают с размерами светового пятна нулевого порядка, происходит интегрирование интерференционной картины по всей площади фотоприемника. При этом переменная составляющая тока фотоприемника, обусловленная интерференцией световых пучков, определяется равенством

$$I_{\approx} = 4A_0 \cdot A_1 \cdot OTF(\eta) \cdot \cos 2\pi \cdot \eta \cdot V_{rec} \cdot t ,$$

где  $A_0^2, A_1^2$  - интенсивности дифракционных порядков,  $OTF(\eta)$  -оптическая передаточная функция (ОПФ) объектива.

Для отражения светового пучка от информационной поверхности CD дисков используется дешевое алюминиевое покрытие толщиной 50...10 нм, у которого коэффициент отражения от 60 до 85%. *Время хранения* таких дисков в нормальных климатических условиях всего 10...15 лет. Это связано с коррозией отражающего покрытия из алюминия.

В двухслойных SACD дисках в качестве отражающего покрытия слоя с низкой плотностью записи также используется пленка алюминия. Отражающее покрытие слоя с высокой плотностью является поляризационным, оно излучение лазера с длиной волны 780 нм пропускает без потерь, а излучение с длиной волны 650 нм полностью отражает. Изготавливается такое покрытие путем вакуумного нанесения очень тонкой пленки кремния, которая не подвержена коррозии.

По опубликованным данным в DVD дисках используются металлические отражающие и полупрозрачные покрытия для однослойных и двухслойных вариантов изготовления с коэффициентом отражения 65 и 30% соответственно. Однако ничего неизвестно о материале этих покрытий. Предполагаемый срок хранения DVD дисков от



50 до 300 лет. Однослойные и двухслойные DVD диски изготавливаются из двух склеенных подложек толщиной 0,6 мм. При указанном сроке хранения наиболее вероятно, что отражение обеспечивается без применения металлических покрытий, например, путем изменения коэффициента преломления материала каждой подложки вблизи ее информационно-поверхности.

*К дефектности CD дисков* особых требований не предъявляется, достаточно чтобы вероятность ошибочных бит не превышала значения  $10^{-3}$ . Требования к качеству DVD(SACD) значительно более жесткие: размеры воздушных пузырьков не должны превышать 100 мкм, дефекты в виде темных пятен должны быть меньше 200...300 мкм, расстояние между дефектами вдоль дорожки – не менее 15...20 мм.

Поликарбонат, из которого изготавливаются оптические диски, достаточно мягкий материал и его легко поцарапать даже ногтем. Поэтому брать диски в руки можно только за края. Их нельзя класть на стол прозрачной поверхностью, нельзя касаться этой поверхности пальцами и оставлять их отпечатки. На этих отпечатках происходит дифракция лазерного излучения и воспроизведение становится невозможным.

Поверхность диска может протираться от пыли и даже полироваться для удаления царапин. Если оптические диски используются в помещениях с сильной запыленностью в виде очень мелкого песка, при длительной работе вся поверхность диска покрывается сеткой очень мелких царапин. Из-за этих царапин возникает дифракция лазерного излучения, и воспроизведение становится невозможным.

Оптический метод воспроизведения записи в значительной мере позволяет исключить влияние царапин и дефектов дисков на качество воспроизведения. Это обеспечивается тем, что сфокусированное излучение лазера имеет большой угол сходимости ( $2\psi$ ), около  $60^\circ$ . При этом диаметр светового пятна на поверхности диска равен 0,8 мм. Поэтому возможные дефекты внешней поверхности в виде мелких царапин и пылинок размером до 0,5 мм находятся далеко вне фокуса объектива, они пространственно фильтруются объективом и не оказывают влияния на качество воспроизведения. Это является одним из важнейших преимуществ оптической записи.

*Изготовление оптических дисков с рельефно-фазовой записью* музыки включает в себя 6 этапов:

- запись цифровой мастер-ленты,
- изготовление диска оригинала,
- запись на диск оригинал первичной фонограммы,
- изготовление *мастер-диска*,
- изготовление никелевой матрицы и
- тиражирование дисков.

Цифровая *мастер-лента* содержит готовую музыкальную программу с тайм-кодом. Необходимая информация пользователя (название и длительности всех составных частей альбома и др.) готовится на компьютере и вводится с него в канал записи, Мастер-лента готовится на студии звукозаписи режиссером с помощью многодорожечных цифровых магнитофонов.

*Диск-оригинал.* Основа этого диска изготавливается из оптического стекла, диаметр его 220 мм, толщина 10 мм. Он тщательно полируется и контролируется, вероятность дефектов на диске не должна превышать  $10^{-6}$ . На поверхность диска наносится слой *фоторезиста* толщиной примерно равной  $\frac{1}{4}$  длины волны излучения лазера, используемого в проигрывателе.

Фоторезист должен иметь максимальную чувствительность в области светового излучения используемого для записи лазера. Запись первичной фонограммы производится с помощью прецизионного лазерного рекордера, функциональная схема которого приведена на **рис.8.7**. Для записи используется гелий-кадмиевый газовый лазер, у которого длина волны излучения 440 нм (синий цвет) и мощность излучения около 15

мВт. Его излучение фокусируется микро объективом с числовой апертурой 0,9 в пятно диаметром 0,3 мкм.

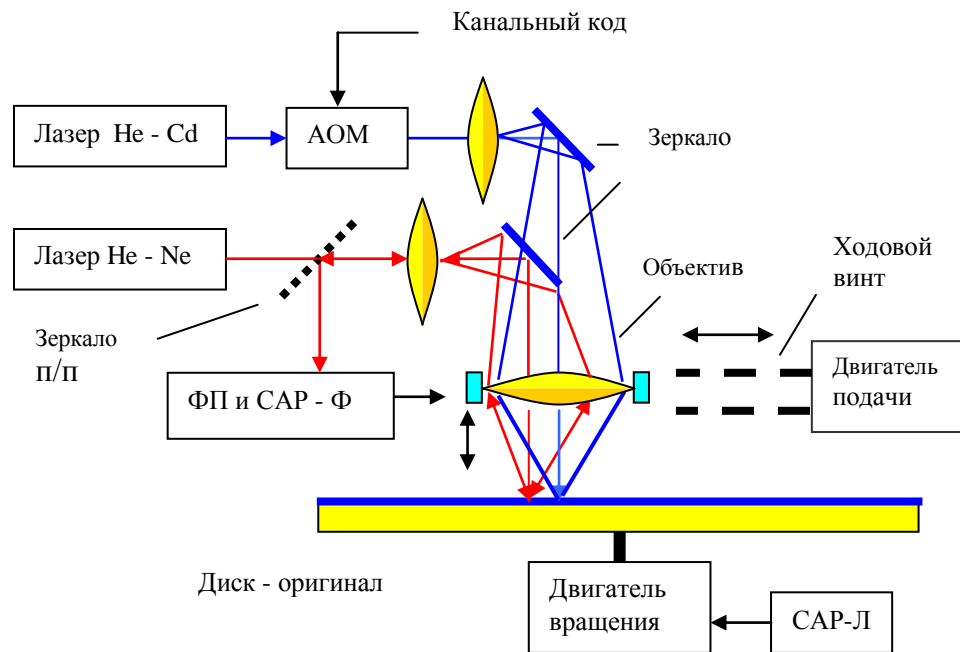


Рис.8.7. Лазерный рекордер для изготовления дисков - оригиналов

Такой объектив имеет глубину фокуса всего 0,5 мкм, поэтому требуется эффективная автоматическая фокусировка (САР-Ф). Она осуществляется с помощью вспомогательного гелий-неонового газового лазера, полупрозрачного (п/п) зеркала и фотоприемника. Длина волны излучения этого лазера 632 нм (красный цвет) и мощность излучения 5 мВт. Фоторезист нечувствителен к такому излучению. Зеркало разделяет падающий и отраженный световые пучки

При записи осуществляется амплитудная модуляция излучения гелий-кадмиевого лазера с помощью акусто-оптического модулятора (АОМ). На вход модулятора подается звуковая информация в канальном коде. Излучения обоих лазеров оптически суммируются и фокусируются одним объективом.

При записи диск-оригинал вращается против часовой стрелки, а оптическая головка с объективом перемещается по радиусу диска. Запись производится с постоянной линейной скоростью  $V_{rec}$ , поэтому частота вращения диска  $\omega$  зависит от радиуса записи  $R_{rec}$

$$\omega = \frac{60 \cdot V_{rec}}{2\pi R_{rec}}, \text{ об/мин.}$$

Скорость радиальной подачи оптической головки определяется заданным шагом между дорожками так, чтобы за один оборот диска головка смещалась на заданный шаг спиральной дорожки. Стабилизации линейной скорости записи и скорости радиальной подачи головки осуществляются с помощью прецизионной механики и систем автоматического регулирования линейной скорости записи САР-Л и радиальной подачи оптической головки САР-РП.

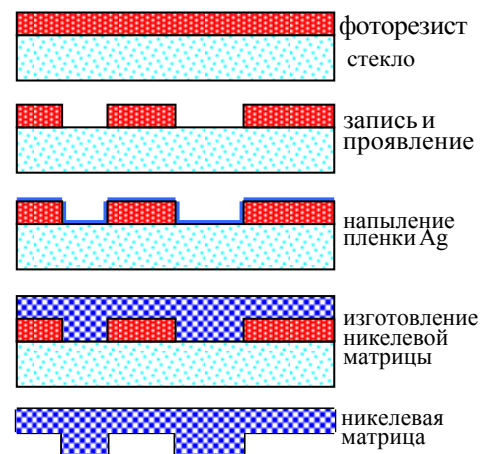


Рис.8.8. Изготовление никелевой матрицы

Под действием модулированного лазерного излучения засвечивается фоторезист. В нем происходят невидимые физико-химические изменения, как в обычной фотопленке. В соответствии с временными интервалами канального кода в фоторезисте формируются невидимые питы. Далее следует этап проявления, при котором в засвеченных местах фоторезист вымывается и образуются углубления, соответствующие питам. Затем диск покрывается тонкой пленкой серебра, чтобы сделать его отражающим. Такой диск с

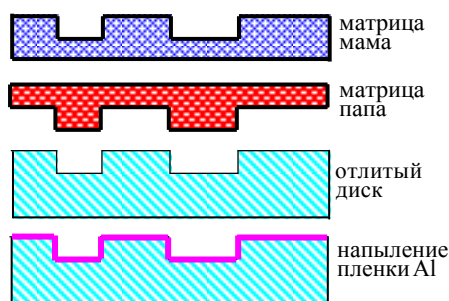


Рис.8.9. Тиражирование оптических дисков

записью фонограммы называется *мастер-диск*. Качество его тщательно контролируется на специальных проигрывателях.

Мастер-диск используется для изготовления никелевой матрицы электролитическим способом (рис.8.8). С помощью этой матрицы изготавливаются рабочие матрицы «мама» и «папа». Тиражирование оптических дисков осуществляется с помощью матрицы «папа» из оптически прозрачного поликарбоната с коэффициентом преломления 1,5. Тиражирование может выполняться литьевым методом или методом прессования. После тиражирования на

информационную поверхность каждого диска наносится отражающее покрытие из алюминия и защитное покрытие (рис.8.9), на которое клеится этикетка.

#### 8.4. Амплитудная запись

Амплитудная запись отличается от фазо-рельефной, тем что в процессе записи на светочувствительном покрытии вдоль спиральной дорожки цифровая информация регистрируется в виде темных штрихов, а не углублений, называемых метками. На отражающей поверхности диска метки имеют пониженный коэффициент отражения. Ширина их одинакова, а длина и расстояние между ними меняются дискретно в соответствии с используемым канальным кодом, который определяет максимальное число таких интервалов, обычно от 2 до 14.



Рис.8.10. Волнообразные канавки и служебные питы на поверхности

Физика воспроизведения амплитудной записи не такая, как при воспроизведении рельефно-фазовой записи. Отраженный световой пучок модулируется метками по интенсивности. При этом глубина модуляции тем больше, чем меньше диаметр считывающего пятна света. Неизбежно возникающая дифракция света на элементах записи в данном случае не имеет принципиального значения, не столь важны и размеры

фотоприемника.

Амплитудная запись производится на записываемых и перезаписываемых дисках *всех форматов*, называемых «болванками». Эти болванки массово тиражируются как и диски с рельефно-фазовой записью. Отличие в том, что при изготовлении диска-оригинала на слое фоторезиста записывается спиральная «немая» дорожка в виде канавки имеющей глубину, ширину и шаг в соответствии с форматом записи. Между канавками при изготовлении дисков записывается служебная информация в рельефно фазовом виде (питы): требуемая мощность лазера, возможная скорость записи, временные и адресные коды.

В процессе записи диска-оригинала с помощью амплитудной модуляции производится периодическое отклонение светового пучка в радиальном направлении. Этот процесс называется *вобуляцией*. Одновременно производится частотная модуляция

вобулирующего сигнала, с помощью которой записывается сигнал времени, позволяющий формировать тактовую частоту при воспроизведении записи с этих дисков.

В результате на штампованных болванках канавка имеет волнообразный характер (рис.8.10). Такая форма канавки позволяет осуществлять радиальное слежение за положением записывающего пятна света относительно центра канавки. При тиражировании на поверхность болванки сначала наносится отражающее покрытие, а потом она вся покрывается тонким светочувствительным покрытием, но амплитудная запись производится только по дну канавки.

В каждом формате используются свои параметры вобуляции и частотной модуляции. Например, на CD-R дисках частота и амплитуда вобуляции соответственно равны 22,5 кГц и 0,03 мкм, а кварцевая частота сигнала времени - 1 кГц. Частота вобуляции DVD-RW дисков является функцией частоты вращения  $\omega$  и определяется произведением  $2880 \times \omega$ , при  $\omega=20$  Гц она равна 57600 Гц.

**Записываемые диски форматов CD-R и DVD-R.** Основой конструкции таких дисков является подложка E из оптически прозрачного поликарбоната толщиной 1,2 мм у CD дисков и 0,6 мм у DVD дисков, на внутренней поверхности, на которой имеется спиральная канавка (рис.8.11). У дисков указанных форматов в качестве светочувствительного покрытия используется тонкая прозрачная пленка фталоцианинового или цианинового органического красителя D. Такая пленка чувствительна к лазерному излучению с длиной волны 635...780 нм и при нагреве этим излучением до температуры выше 100<sup>0</sup>С она меняет свой цвет и становится для него не прозрачной. Поэтому после записи поверхность диска имеет высокий коэффициент отражения и не прозрачные информационные («черные») метки.

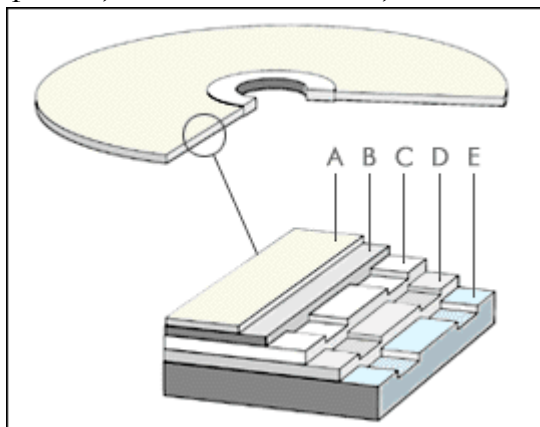


Рис.8.11. Конструкция CD-R диска

Оптический контраст такой записи тем выше, чем больше коэффициент отражения металлического покрытия C. Поэтому толщина покрытия достаточно большая, от 50 до 100 нм и выполняется оно из серебра или его сплавов, у которых коэффициент отражения достигает 65...85%. Для дисков профессионального назначения для этих целей используется золото и даже платина. Они очень дорогие. Для защиты информационного слоя от повреждений и воздействия воздушной среды дисков CD-R применяется лаковое покрытие B. Поверх него клеится этикетка A. Однослойные диски DVD-R отличаются тем, что вместо защитного покрытия приклеивается подложка из поликарбоната толщиной 0,6 мм. Она более надежно защищает информационный слой. В односторонних двухслойных DVD-R дисках в одном слое используется отражающее покрытие из серебра, а в другом – полупрозрачное покрытие из золота.

Цианиновый краситель обладает сине-зеленым (цвет «морской волны») или насыщенно синим оттенком рабочей поверхности, фталоцианин, в большинстве случаев, практически бесцветен, с бледным оттенком салатного или золотистого цвета, из-за чего диски на основе фталоцианинового активного слоя часто называют «золотыми». Сложно сказать, какой из этих двух типов слоев лучше. Цианиновый краситель более терпим к предельным сочетаниям мощности чтения/записи, чем «золотой» фталоцианиновый, поэтому зачастую диски на основе цианинового слоя проще считывать на некоторых дисководы. Фталоцианин - более современная разработка. Диски на основе этого активного слоя менее чувствительны к солнечному свету и ультрафиолетовому излучению, что способствует увеличению долговечности записанной информации и несколько более надежному хранению в неблагоприятных условиях.

При записи с однократной скоростью сканирования используются лазеры с мощностью 4...8 мВт. При 40x повышении скорости требуется увеличение мощности лазера до 40 мВт. Под действием лазерного излучения температура красителя повышается до 250<sup>0</sup>С, в нем происходят химические изменения и коэффициент отражения уменьшается с 75 до 25%. Отношение сигнал\шум при записи составляет около 47 дБ. Вероятность дефектов обычно оценивается значением частоты ошибочных блоков BLER и она равна примерно 20/с, тогда как стандартом CD допускается BLER=220/с. Срок хранения записи на CD-R дисках с алюминиевым покрытием не превышает 10...15 лет.

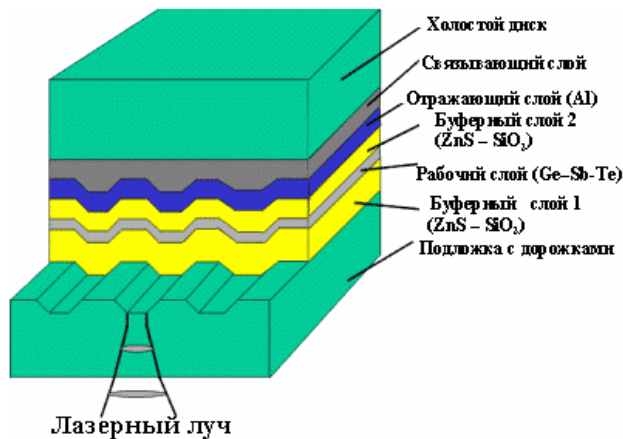


Рис.8.12. Конструкция диска DVD-RW

информационная поверхность диска со спиральной канавкой покрывается тонкой пленкой материала, который может под влиянием температуры переходить из кристаллического состояния в аморфное с заметным изменением коэффициента отражения. В качестве такого материала используется множество сплавов, имеющих низкую температуру кристаллизации. Например, сплав серебра, индия, сурьмы и теллура имеет температуру кристаллизации 200<sup>0</sup>С. У него в кристаллическом состоянии

Если используется отражающее покрытие из серебра это время увеличивается до 75...100 лет. Еще больше время хранения записи на очень дорогих профессиональных CD-R дисках с золотым или платиновым покрытием, которое достигает 200 лет. Примерно такой же срок хранения у DVD-R дисков. Критерием времени хранения является увеличение вероятности ошибочных блоков выше допустимого значения.

*Перезаписываемые диски форматов CD-RW и DVD-RW.* У дисков CD-RW и DVD-RW



Рис.8.13. Запись, стирание и считывание информации с CD-RW и DVD-RW дисков

коэффициент отражения около 35%, а в аморфном состоянии он уменьшается до 25%. Похожие характеристики имеет сплав германия, сурьмы и теллура, у него температура кристаллизации 150...200<sup>0</sup>С, а температура плавления –600<sup>0</sup>С.

Подложка у дисков этих форматов такая же, как и у записываемых дисков. Термочувствительное покрытие наносится между двумя слоями диэлектриков (рис.8.12.). Химический состав активного слоя определяет такие важные показатели, как время и температура кристаллизации, толщина слоя – коэффициент отражения, а состав и толщина окружающих диэлектрических слоев – теплопроводность и, соответственно, время охлаждения слоя в процессе термической записи. Слой алюминия также выполняет буферную роль.

Диски CD-RW также имеют отражающее покрытие из серебра, золота или платины. Цвет у них серо-коричневый. Контраст записи на этих дисках значительно меньше, чем на CD-R дисках и они читаются не на всех дисководов.

В исходном состоянии светочувствительный слой находится в кристаллическом состоянии. Запись производится импульсным сфокусированным лазерным излучением, которое в течение длительности короткого управляющего импульса около 12 нс нагревает его выше температуры плавления 500...700<sup>0</sup>С (рис.8.13). По окончании действия этого импульса происходит быстрое охлаждение активного слоя до температуры

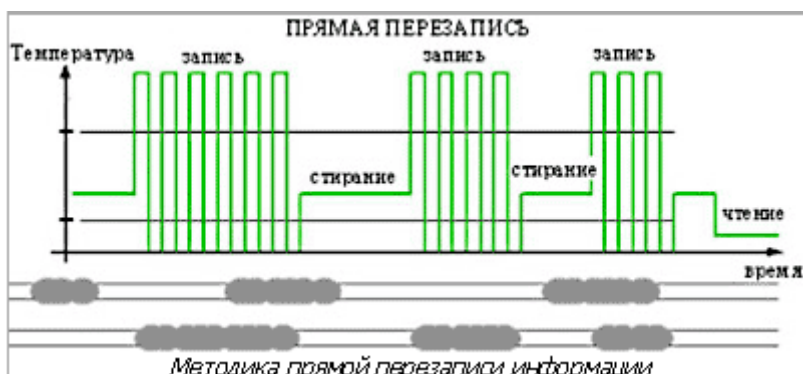


Рис.8.14. Запись временных интервалов 3Т-11Т

ниже кристаллизации, поэтому он переходит в аморфное состояние. Быстрое охлаждение достигается благодаря высокой теплопроводности, окружающих покрытий. Результат охлаждения – предотвращения образования центров кристаллизации, поэтому роста кристаллической фазы не происходит и материал

переходит в аморфное состояние.

Для стирания записи и возвращения активного слоя в кристаллическое состояние необходим нагрев ниже температуры плавления, но выше температуры кристаллизации около 200<sup>0</sup>С. Время этого нагрева должно быть больше времени кристаллизации. Таким образом, запись производится коротким импульсом большой мощности, а стирание – более длительным импульсом меньшей мощности.

Запись сигналов производится в канальном коде, когда цифровая информация передается в виде различной длины временных интервалов от 3Т до 14Т. Этим временным интервалам соответствуют серии коротких мощных световых импульсов от 3 до 14 подряд (рис.8.14). В стандарте CD  $T_{min}$  равно 694 нс, а в стандарте DVD –38 нс, им и соответствует длительность световых импульсов записи около 12 нс.

Для считывания записи с дисков CD-RW и DVD-RW достаточна мощность излучения 0,5..0,6 мВт. В профессиональных рекордерах для повышения плотности записи используются лазеры с длиной волны 635 ...405 нм мощностью до 18 мвт. Число циклов записи - стирания от 1000 до 100000, число циклов чтения - не менее 100000. Время хранения записи на DVD-RW дисках от 25 до 100 лет.

## 8.5. Системы автоматического позиционирования

В процессе воспроизведения записи оптический диск вращается, световое пятно медленно смещается по радиусу и сканирует дорожку с питами. Для обеспечения постоянства линейной скорости сканирования частота вращения изменяется в зависимости от положения считывающего пятна по радиусу. Частота вращения максимальна на минимальном радиусе и минимальна на максимальном радиусе зоны записи звука. В стандарте CD при линейной скорости 1,3 м/с она меняется в пределах от 496 до 214 об/мин, в стандарте DVD(SACD) линейная скорость записи выше и равна 3,49 м/с, а частота вращения значительно меняется от 1389 до 575 об/мин.

Т а б л и ц а 8.6. Ошибки позиционирования		
Параметры	CD	DVD,SACD
Эксцентриситет записи, мкм	140	100
Радиальная ошибка, мкм	0,03	0,022
Вертикальный бой, мм	± 0,5	
Ошибка фокусировки, мкм	± 1 мкм	± 0,23 мкм

В процессе сканирования записи световое пятно должно быть сфокусировано на дорожку с питами и находится все время в центре этой дорожки. Существующие технологии тиражирования допускают возникновение покоробленности дисков, из-за

которой при вращении дисков возникают вертикальные биения, достигающие величины  $\pm 0,5$  мм. Эксцентриситет зоны записи обусловлен не точностью изготовления центрального отверстия диска. Он вызывает радиальные биения до 140 мкм. Допустимая ошибка фокусировки не должна превышать половины глубины фокуса объектива, а радиальная ошибка позиционирования должна быть меньше по крайней мере 0,1 ширины дорожки. Это значит, что без системы автоматического регулирования (САР) воспроизведение оптической записи невозможно. В табл.8.6. приведены данные по допустимым ошибкам, предусмотренные стандартами.

Для двух координатного позиционирования сканирующего пятна используется три САР: САР фокусировки (САР-Ф), САР радиального слежения за дорожкой (САР-РС) и САР радиальной подачи оптической головки (САР-РП). Для обеспечения постоянства скорости воспроизводимого цифрового потока используется трехступенчатая САР-С, которая обеспечивает кварцевую стабильность частоты воспроизводимого сигнала.

Все проигрыватели и рекордеры включают в себя следующие оптико-механические узлы: оптическая головка, актюатор оптической головки, узел радиальной подачи оптической головки и узел вращения компакт-диска. В оптическую головку входит полупроводниковый лазер, объектив, поляризационный светоделитель, четвертьволновая пластинка, много площадочный фотоприемник (ФП), датчики ошибок фокусировки и радиального слежения за дорожкой. Масса такой головки от 0,9 до 2,5 г.

*Актюатор* – это исполнительный двух координатный механизм САР, который перемещает объектив в вертикальном и радиальном направлениях. Обычно он имеет два линейных двигателя в виде катушки головки громкоговорителя с постоянным магнитом. Диапазон отработки ошибок фокусировки находится в пределах  $\pm 0,5$  мм и лишь в некоторых моделях проигрывателей до  $\pm 2,5$  мм. Чувствительность по входу 1,5...4 мм/В. Диапазон отработки радиальных ошибок  $\pm 0,3$ ... $\pm 0,5$  мм, чувствительность по входу 0,2...0,6 мм/В.

*Автоматическая фокусировка.* Ошибка фокусировки приводит к увеличению диаметра считывающего пятна. Как следствие, искажается форма ОПФ и уменьшается глубина модуляции воспроизводимого сигнала. Когда диаметр считывающего пятна становится равным длине волны записи, модуляция уменьшается до нуля и

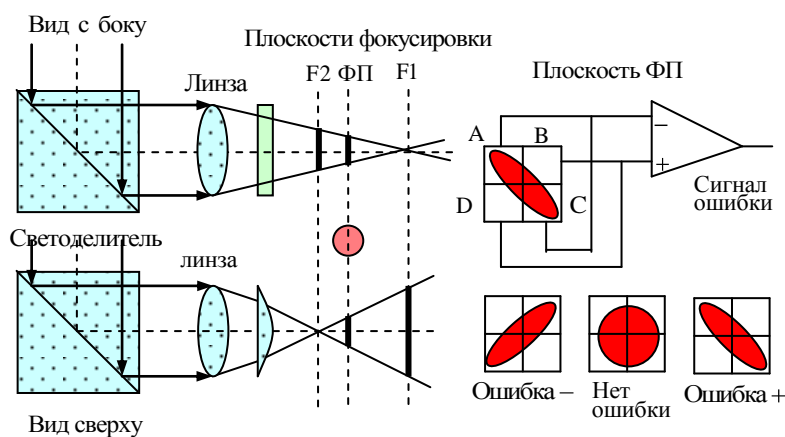


Рис.8.15. Схема датчика сигнала ошибки фокусировки с цилиндрической линзой

воспроизведение становится невозможным. Ошибка фокусировки может быть статической и динамической. Статическая ошибка возникает из-за плохой юстировки оптической схемы. Динамическая ошибка возникает из-за вертикальных биений диска при вращении. Эта ошибка приводит к возникновению паразитной амплитудной модуляции с частотой вращения диска, которая создает кодовые ошибки в виде длинных пакетов и выпадений.

Исходя из данных табл.8.6. по допустимым ошибкам позиционирования можно рассчитать, что необходимый коэффициент регулирования САР-Ф в стандарте CD равен 500, а в стандарте DVD(SACD) – 2170. Это говорит о более высоких технических параметрах стандарта DVD.

Основным элементом САР-Ф является датчик, формирующий сигнал, величина которого пропорциональна ошибке фокусировки, а полярность – указывает на знак ошибки. В настоящее время разработаны и применяются десятки разных датчиков. Наиболее широко применяется датчик с цилиндрической линзой (ЦЛ), схема которого приведена на рис.8.15. ЦЛ устанавливается в схеме воспроизведения в отраженном световом пучке после сферической линзы.

Цилиндрическая линза ведет себя в вертикальной плоскости как плоскопараллельная стеклянная пластина, поэтому в этой плоскости она не изменяет световой пучок, сформированный сферической линзой, и он фокусируется в плоскости F1. На рис.8.15. эта плоскость представлена как вид с боку. В горизонтальной плоскости (вид сверху) ЦЛ формирует сходящийся пучок света с фокусом в плоскости F2. В результате световые пятна в плоскостях F1 и F2 преобразуются в прямые линии, взаимно перпендикулярные. На схеме они обозначены темными линиями с длиной равной сечению световых пучков в этих плоскостях.

Всегда существует некоторая плоскость ФП, в которой проекции световых пятен в горизонтальной и вертикальной плоскостях равны. Это значит, что в этой плоскости сечение светового пучка имеет вид окружности. В этой плоскости размещают 4-х площадочный ФП датчика. Площадки А,В,С и D фотоприемника соединяют между собой, как это показано на схеме. Сигналы А+С и В+D подаются на дифференциальный усилитель, формирующий сигнал ошибки.

Если ошибка фокусировки равна нулю, световое пятно на ФП имеет вид окружности, все его площадки освещены одинаково и сигнал ошибки равен нулю. Ошибка фокусировки приводит к смещению плоскостей фокусировки F1 и F2 относительно плоскости ФП влево или вправо в зависимости от знака ошибки. При этом световое пятно на ФП принимает вид эллипса и его ось поворачивается влево или вправо в зависимости от знака ошибки. В этом случае площадки ФП освещены не одинаково и поэтому на выходе усилителя формируется сигнал, пропорциональный величине ошибок.

Рассмотренный датчик ошибки фокусировки широко используется в проигрывателях системы CD и он стандартизирован для системы DVD.

*Радиальное слежение за дорожкой.* Ошибка положения пятна на дорожке приводит к паразитной амплитудной модуляции сигнала воспроизведения, уменьшению его глубины модуляции вплоть до нуля, когда пятно уходит с дорожки и воспроизведение прекращается. Исходя из данных табл.8.6. по допустимым ошибкам позиционирования можно рассчитать необходимый коэффициент регулирования САР-РС. В стандарте CD он равен 4670, а в стандарте DVD(SACD) – 4545.

В настоящее время разработаны и применяются десятки разных датчиков сигнала ошибки, определяющего радиальное положение считывающего пятна света. Наиболее широко в стандарте CD применяется датчик с использованием двух дополнительных световых пучков, как это показано на рис.8.16. Дополнительные пятна отстоят от основного на 20 мкм. Один из них смещен влево от дорожки, другой - вправо. Для

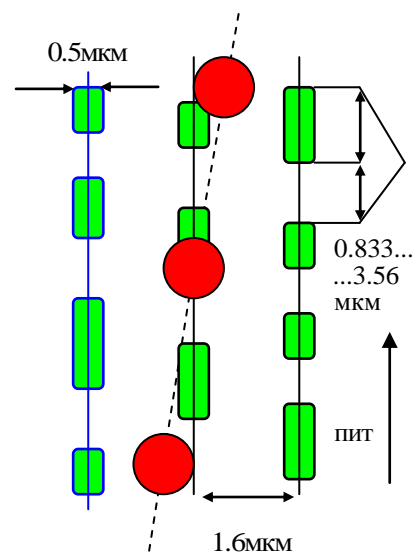


Рис.8.16. Слежение за дорожкой



каждого светового пятна в схеме воспроизведения имеется отдельный ФП. Выходы двух дополнительных ФП подключены к дифференциальному усилителю, который формирует сигнал ошибки. Если ошибки положения нет, оба ФП освещены одинаково и выходной сигнал усилителя равен нулю. Ошибка приводит к большей освещенности одного или другого ФП, тогда формируется сигнал ошибки.

Формирование 3 световых пятен в схеме воспроизведения осуществляется несколькими способами. Часто для этого используют специальную фазовую решетку, которая размещается в световом пучке лазера. Дифракция излучения на этой решетке приводит к возникновению 3 световых пучков 0 и  $\pm 1$  порядков. Интенсивность пучка 0 порядка обычно больше. Все эти световые пучки фокусируются одним объективом на носитель записи, как это показано на рис.8.16. В другом варианте используются 3 отдельных лазера, изготовленных в одном корпусе. Все 6 ФП всегда изготавливаются на одной подложке.

В стандарте DVD для штампованных дисков стандартизирована другая более простая система CAP-PC с использованием явления дифракции света на дорожке записи. На рис.8.17. поясняется принцип работы такой системы. Когда считывающее пятно

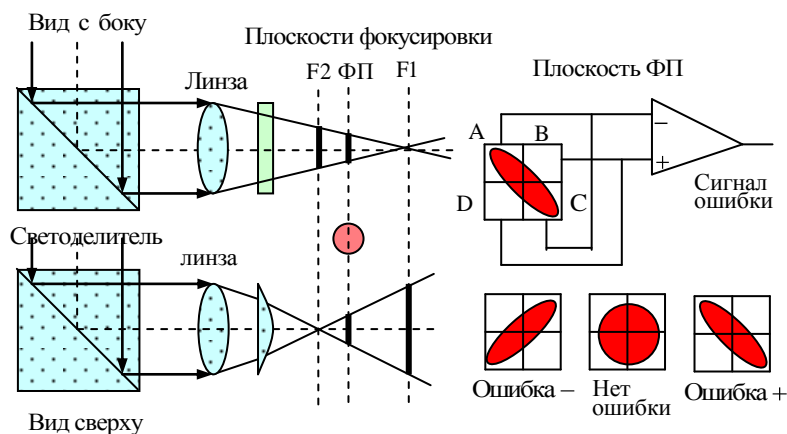


Рис.8.17. Схема датчика сигнала ошибки фокусировки с цилиндрической линзой

сканируется по дорожке записи с пятами происходит дифракция света как вдоль дорожки, так и в радиальном направлении. На рисунке показаны порядки дифракции: 0, 1 и  $-1$ . В областях перекрытия нулевого и первых порядков дифракции на ФП (заштрихованные области) возникает интерференция света, вызывающая амплитудную модуляцию выходного сигнала ФП.

Если радиальной ошибки позиционирования нет, интенсивности первых порядков дифракции в радиальном направлении равны, поэтому равны и токи с левой и правой половин 4-х площадочного ФП. В этом случае дифференциальный усилитель вырабатывает сигнал ошибки, равный нулю. Радиальная ошибка приводит к тому, что увеличивается сигнал с левой или правой половины ФП. Соответственно дифференциальным усилителем формируется сигнал ошибки, полярность которого указывает на ее знак.

Формирование сигнала ошибки фокусировки осуществляется с помощью цилиндрической линзы и 4-х площадочного ФП, как это рассмотрено выше. В рамках приведенной схемы для выработки этого сигнала суммируются токи площадок ФП по диагоналям и тогда дифференциальный усилитель формирует сигнал ошибки. Считываемый сигнал формируется как сумма токов со всех площадок ФП.

**САР радиальной подачи.** Схема радиальной подачи приведена на рис.8.18. В этой схеме актюатор с оптической головкой перемещается в радиальном направлении с помощью ходового винта, редуктора и двигателя вращения. В режиме воспроизведения актюатор должен смещаться на величину шага спиральной дорожки за каждый оборот

диска. Так как частота вращения диска зависит от положения по радиусу считывающего пятна света, то радиальная скорость подачи также должна меняться с радиусом считывания в соответствии с равенством

$$V_h(R_{rec}) = \frac{h_p V_{rec}}{2\pi R_{rec}},$$

где  $V_h(R_{rec})$  – скорость движения актюатора по радиусу. В стандарте CD движение актюатора происходит со средней скоростью от 13,2 до 5,7 мкм. Такая маленькая скорость движения обеспечивается редуктором.

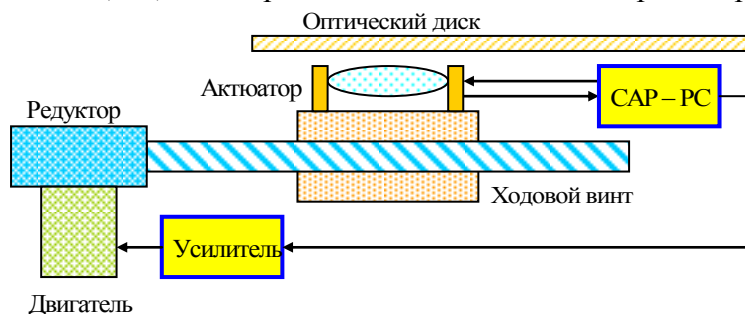


Рис.8.18. Схема радиальной подачи актюатора

Датчиком положения в этой схеме является оптический диск с записью информации о номерах треков и секторов. Двигатель работает в старт-стопном режиме и управляется CAP-PC. В этой

CAP датчик радиальной ошибки, входящий в оптическую схему проигрывателя, вырабатывает два сигнала. Один, относительно высокочастотный, связан с радиальными ошибками положения считывающего пятна на дорожке. Другой, очень низкочастотный, связан с ошибкой радиальной подачи.

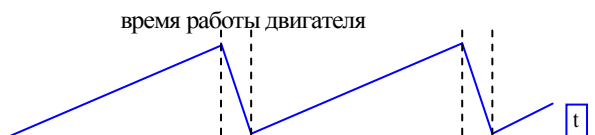


Рис.8.19. НЧ сигнал ошибки подачи

В нормальном состоянии двигатель подачи не работает и ошибка радиальной подачи увеличивается со временем (рис.8.19.). Когда эта ошибка достигнет заданной величины включается двигатель подачи и ошибка начинает быстро уменьшаться. При уменьшении ее до значения близкого к нулю работа двигателя прекращается.

Время работы двигателя от 2 до 6 с.

*CAP стабилизации скорости воспроизводимого цифрового потока.* Эта CAP включает в себя три ступени регулирования. Первые две ступени обеспечивают постоянство линейной скорости сканирования. Для этого необходимо во время воспроизведения

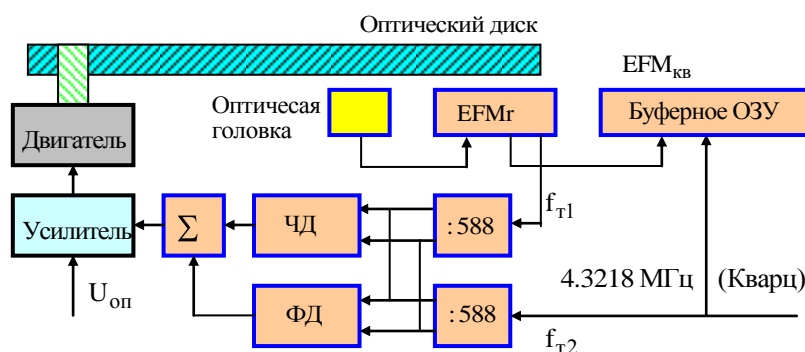


Рис.8.20. Схема стабилизации скорости считывания

менять число оборотов двигателя вращения диска в соответствии с равенством

$$N = \frac{60 \cdot V_{rec}}{2\pi R_{rec}},$$

где значения скорости  $V_{rec}$  и радиуса  $R_{rec}$  могут принимать значения в пределах, предусмотренных стандартом.

На рис.8.20. приведена упрощенная схема CAP-C, используемая в стандарте CD. На таких же идеях строится CAP в стандарте DVD. В схему входит двигатель вращения

постоянного тока и три ступени регулирования. В первой ступени осуществляется стабилизация частоты вращения и изменение ее в функции радиуса сканирования. Функции датчика частоты вращения выполняет сам оптический диск с информацией о тактовой частоте канального кода, равной 4,32 МГц. При воспроизведении записи сигнал с оптической головки подается на формирователь считываемого сигнала EFM<sub>r</sub>, у которого тактовая частота  $f_{T1}$  может отличаться от 4,32 МГц. Канальный код обладает свойством самосинхронизации, поэтому с помощью формирователя из него эта частота выделяется. В качестве опорной тактовой частоты используется сигнал от кварцевого генератора  $f_{T2}$ , равно точно 4,3218 МГц. Сравнение этих частот производится на более низкой блочной частоте 7350 Гц, поэтому оба сигнала делятся на 588.

В первую ступень стабилизации скорости сканирования входит частотный детектор (ЧД). На два его входа подаются сигнал блочной частоты, воспроизведенный с диска, и сигнал от кварцевого генератора. Если эти частоты не равны, детектор вырабатывает сигнал ошибки, который изменяет напряжение на входе двигателя и, следовательно, число его оборотов. Это грубая ступень регулирования, она работает только при разгоне двигателя и обеспечивает равенство тактовых частот  $f_{T1}$  и  $f_{T2}$ .

Во вторую ступень регулирования входит фазовый детектор (ФД), который регулирует число оборотов двигателя так, чтобы фазы этих частот совпадали с точностью до нескольких градусов. Опорное напряжение усилителя  $U_{оп}$  определяет номинальное число оборотов двигателя.

Третья электронная ступень регулирования устраняет фазовые качания воспроизведенного сигнала EFM. Для этого сигнал с формирователя записывается в буферное ЗУ со своей тактовой частотой  $f_{T1}$ , а считывается из него с кварцевой частотой  $f_{T2}$ . Чем больше емкость этого ЗУ, тем меньше требования ко 2-й ступени регулирования. Обычно емкость его около 3000 канальных бит, что позволяет устранять временные колебания в пределах  $\pm 500$  мкс. В последних моделях проигрывателей используется память несколько мегабайт, на основе которых строится защита от случайных ударов, которые вызывают изменение положения считывающей головки.

## 8.6. Стандарты корректирующего кодирования CD и DVD

Для обнаружения и исправления ошибок в стандарте CD используется 2-х каскадный код RS в поле Галуа( $2^8$ ) под названием CIRC (Cross Interlived Reed-Solomon Code - код Рида-Соломона с перемежением). Число элементов поля 256, они задаются порождающим полиномом:

$$G(X) = X^8 + X^4 + X^3 + X^2 + 1$$

и примитивным элементом  $\alpha = X$  (00000010).

Перед кодированием мультиплексированный поток аудио данных левого (L) и правого (R) каналов в виде 16-ти разрядных слов с частотой дискретизации 44,1 кГц организуется

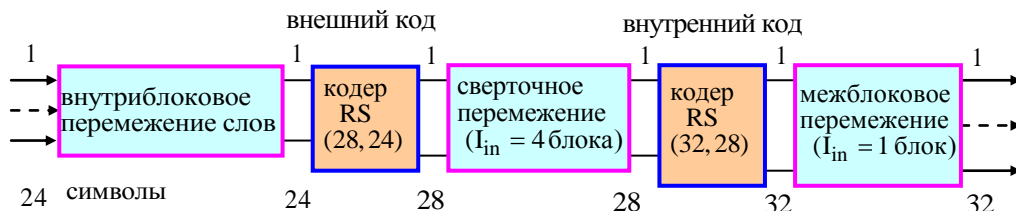


Рис.8.21. Блок-схема кодера CIRC

в блоки по 6 слов левого и правого каналов. Затем каждое слово разбивается на 2 символа по 8 бит, отдельно старшие и младшие разряды. В результате блок данных содержит 24 символа (по 12 символов левого и правого звуковых каналов). Необходимость разбиения слов на символы связана с тем, что код RS недвоичный и все операции в нем производятся с символами, а не битами. На вход

кодера, блок-схема которого приведена на рис.8.21., аудио данные поступают с блочной частотой 7350 блоков/с и скоростью 1,411 Мбит/с.

Кодер включает в себя двухкаскадный RS код и 3 ступени перемежения. Внешний код – код RS (28,24) имеет 4 проверочных символа и избыточность 14,3%. Кодирование выполняется с помощью проверочной матрицы  $H_q$  размером  $28 \times 4$ . Внутренний код – код RS (32,28) также имеет 4 проверочных символа, но избыточность его меньше-12,5%. Кодирование осуществляется с помощью другой проверочной матрицы  $H_p$  размером  $32 \times 4$ . Эти проверочные матрицы заданы стандартом и имеют следующий вид:

$$H_q = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 & 1 \\ \alpha^{27} & \alpha^{26} & \dots & \alpha^2 & \alpha^1 & 0 \\ \alpha^{54} & \alpha^{52} & \dots & \alpha^4 & \alpha^2 & 0 \\ \alpha^{81} & \alpha^{78} & \dots & \alpha^6 & \alpha^3 & 0 \end{bmatrix}, \quad H_p = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 & 1 \\ \alpha^{31} & \alpha^{30} & \dots & \alpha^2 & \alpha^1 & 0 \\ \alpha^{62} & \alpha^{60} & \dots & \alpha^4 & \alpha^2 & 0 \\ \alpha^{93} & \alpha^{90} & \dots & \alpha^6 & \alpha^3 & 0 \end{bmatrix}$$

Коды RS удобны тем, что при использовании полусиндромов  $S_{qj}$  и  $S_{pj}$  проверочные символы  $Q_0, Q_1, Q_2, Q_3$  в первом каскаде и  $P_0, P_1, P_2, P_3$  во втором каскаде могут быть рассчитаны с помощью матриц небольшого размера:

$$\begin{bmatrix} Q_0 \\ Q_1 \\ Q_2 \\ Q_3 \end{bmatrix} = \begin{bmatrix} \alpha^{212} & \alpha^{153} & \alpha^{152} & \alpha^{209} \\ \alpha^{156} & \alpha^2 & \alpha^{135} & \alpha^{152} \\ \alpha^{158} & \alpha^{138} & \alpha^2 & \alpha^{153} \\ \alpha^{218} & \alpha^{158} & \alpha^{156} & \alpha^{212} \end{bmatrix} \times \begin{bmatrix} S_{q0} \\ S_{q1} \\ S_{q2} \\ S_{q3} \end{bmatrix}, \quad \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix} = \begin{bmatrix} \alpha^{212} & \alpha^{141} & \alpha^{128} & \alpha^{173} \\ \alpha^{156} & \alpha^{245} & \alpha^{111} & \alpha^{116} \\ \alpha^{158} & \alpha^{126} & \alpha^{233} & \alpha^{117} \\ \alpha^{218} & \alpha^{146} & \alpha^{132} & \alpha^{176} \end{bmatrix} \times \begin{bmatrix} S_{p0} \\ S_{p1} \\ S_{p2} \\ S_{p3} \end{bmatrix}$$

Внутренний код предназначен прежде всего для обнаружения ошибочных символов, исправления ошибок малой кратности и стираний. Внешний код используется для исправления ошибок большой кратности и стираний.

В первой ступени кодирования осуществляется внутриблоковая перестановка слов левого и правого каналов, а также межблоковое перемежение слов с интервалом 2 блока. Целью этого короткого перемежения является увеличение числа интерполируемых слов при длинных выпадениях, после того как все возможности исправления ошибок в 2 каскадах исправления ошибочных символов в декодере исчерпаны. Во второй ступени кодирования производится длинное межблоковое сверточное перемежение с интервалом 4 блока. Оно предназначено для исправления длинных выпадений в декодере. При этом данные каждого блока систематично размещаются в 112 блоках. Сверточное перемежение позволяет исправлять выпадения сигнала длиной до 8 блоков и маскировать ошибки при выпадениях длиной до 51 блока. В третьей ступени кодирования производится перестановка четных символов из одного блока в соседний. Этим обеспечивается возможность исправления коротких пакетов ошибок уже в 1 ступени декодирования.

После кодирования блок данных состоит из 24 символов звуковых данных и 8

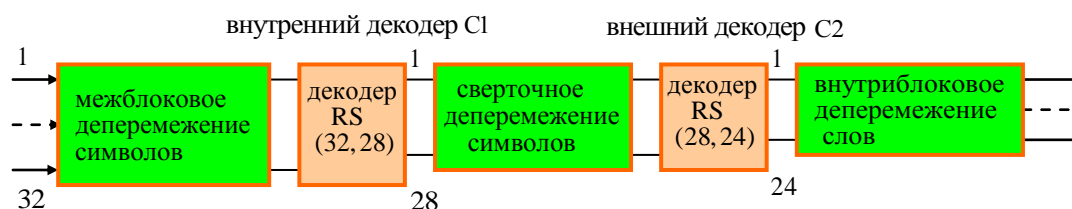


Рис.8.22. Блок-схема декодера кода CIRC

проверочных символов. В него дополнительно включается еще один символ, в котором передается информация для индикации ее на экране дисплея проигрывателя: наименование треков, имена исполнителей, длительность программ и т.д.. При этом скорость цифрового потока возрастает до 1,94 Мбит/с. Далее 98 блоков по 33 символа

объединяются в сектор объемом 3234 бит, из которых только 2352 бита несут звуковую информацию. Это сделано для того, чтобы можно было записывать цифровой звук на диски CD-R с секторной организацией данных. Частота повторения секторов 75 Гц.

Кодер представляет собой специализированный процессор, который по заданной программе производит вычисление в поле Галуа проверочных символов в двух ступенях. Перемежение осуществляется с помощью оперативного ОЗУ емкостью 16 Кбит. Потенциальные возможности кода: полное исправление ошибок при выпадениях длиной до 2,38 мм и маскирование ошибок с помощью интерполяции при выпадениях длиной до 8,16 мм.

Блок-схема декодера кода CIRC приведена на [рис.8.22](#). Короткое межблоковое деперемежение четных символов из одного блока в соседний позволяет в зависимости от стратегии декодирования исправлять от 2 до 8 символов в пакете. Если использовать обнаруживающую способность канального декодера, то в этой ступени можно исправлять стирания, введенные при канальном декодировании. Обычно эта ступень декодирования используется для обнаружения ошибок и исправляется всего 1 ошибка, а остальные символы стираются. Стирание всех символов в блоке производится также, если число ошибок в блоке больше исправляющей способности ступени.

При деперемежении символы пакетов ошибок разбрасываются по 28 блокам с интервалом 4 блока. Это позволяет исправлять длинные выпадения. Длина исправляемых выпадений зависит от исправляющей способности второй ступени декодирования C2. В зависимости от стратегии декодирования она может быть от 2 до 8 блоков подряд. Поэтому декодер 2 ступени должен иметь максимально возможную исправляющую способность как по ошибкам, так и стираниям. Если число ошибок в блоке больше 2 или число стираний больше, чем возможно исправить, все символы блока стираются. Это необходимо для последующей маскировки неисправленных ошибок. Последнее деперемежение производится на уровне слов левого и правого каналов. Оно определяет длину выпадений, исправляемых с помощью интерполяции.

При декодировании в каждой ступени выполняются следующие математические операции:

- деперемежение символов и слов,
- вычисление синдромов ошибок,
- обнаружение ошибок с помощью синдромов,
- идентификация числа ошибочных символов и стираний в блоке,
- определение локаторов ошибочных символов,
- исправление одиночных и двойных ошибок,
- исправление стираний,
- копирование или введение стираний.

Декодер представляет собой специализированный звуковой процессор, который по заданным расчетным формулам и программе выполняет все необходимые операции по деперемежению, обнаружению и исправлению ошибочных символов за время, равное длительности 1 блока – 136 мкс. Для деперемежения используется ОЗУ емкостью от 16 до 32 Кбит.

Если кодер выполняется строго по стандарту, то декодеры различных фирм существенно отличаются стратегиями декодирования и, соответственно, и качеством работы проигрывателей. В зависимости от стратегии декодирования по разному исправляются случайные ошибки и выпадения. Вероятность не обнаружения ошибок в декодере кода CIRC обычно равна  $2^{-19}$ .

В настоящее время применяются в основном 3 стратегии декодирования, краткое описание которых приведено в [табл.8.7](#). В этой таблице представлены данные, характеризующие исправляющую способность по ошибкам в двух ступенях декодирования.

Т а б л и ц а 8.7. Стратегии декодирования		
Сложность стратегии	Декодер С1	Декодер С2
	Если идентификация установила, что в блоке нет или всего одна ошибка:	
Простая	Исправляется 1 ошибка, остальные символы стираются	Исправляется 1 ошибка, остальные символы стираются
Средняя	Исправляется 1 ошибка, остальные символы стираются	Исправляется 1 ошибка и 2 стирания остальных символов стираются
Супер	Исправляются 1 и 2 ошибки, остальные символы стираются	Исправляются 1 ошибка и 2 стирания, 2 ошибки и 1 стирание, 4 стирания, остальные символы стираются

На рис. 8.23 приведены графики этих стратегий, характеризующие исправляющую способность всего декодера по случайным ошибкам. На этих графиках по оси абсцисс отложена вероятность ошибочных символов на входе декодера, а по оси ординат - вероятность ошибочных слов (16 разрядных выборок) на выходе декодера. Область работы декодера лежит правее графиков стратегий: *A*, *B*, *C* и *D*. Прямая *D* определяет обнаруживающую способность кода. В области левее этого графика декодер не работает.

Для всех стратегий декодирования декодер может исправлять ошибки при вероятности появления ошибочных символов на входе декодера меньше, чем  $10^{-3}$ . При этом ошибочные слова на выходе декодера возникают не чаще, чем 1 раз в минуту. Стандартом CD допускается значение BLER = 0,032, что соответствует частоте повторения ошибочных блоков 235 бл/с (2800 слов/с). Для большинства дисков вероятность ошибочных символов около  $10^{-4}$  и тогда с равным успехом они могут воспроизводиться проигрывателями с любыми стратегиями декодирования.

Для супер-стратегии декодирования при вероятности ошибочных символов на входе декодера  $P_s = 10^{-3}$  вероятность ошибочного слова на выходе не превышает  $10^{-16}$  (1 ошибка за 10000 лет). При этих же условиях для стратегии *B* ошибки на выходе декодера будут происходить через 40 мин, а для стратегии *C* – через минуту. Чем более эффективна стратегия декодирования, тем сложнее процессор и требуется большее его быстродействие. Именно стратегиями декодирования и отличаются проигрыватели компакт

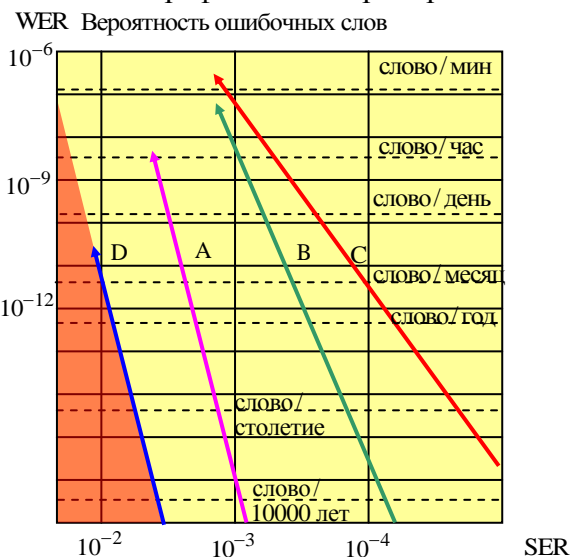


Рис. 8.23. Графики исправляющей способности декодера кода CIRC

Т а б л и ц а 8.8. Исправление выпадений в С2			
Стратегия	Длина исправлений	Полное исправление ошибок	Маскировка выпадений
Упрощенная (С)	в блоках	3	48
	длина, мм	0,51	8,16
Средней сложности (В)	в блоках	8	48
	длина, мм	1,36	8,16
Супер-стратегия (А)	в блоках	14	48
	длина, мм	2,38	8,16

В табл.8.8. приведены характеристики декодера по исправлению длинных выпадений и пакетов ошибок.

*Корректирующее кодирование в стандарте DVD* включает в себя отдельные коды для обнаружения и коррекции ошибочных символов. Для обнаружения ошибок

используется двоичный циклический код CRC с одним проверочным словом длиной 32 бита. Порождающий полином кода

$$G(x) = X^{32} + X^{31} + X^4 + 1.$$

Для вычисления проверочного слова 2048 байт звуковых данных организуются в блок из 2064 байт, из которых 12 символов относятся к служебной информации и 4 байта являются проверочным словом кода CRC. Проверочные 32 бита (EDC- error detection code) вычисляются в соответствии с равенствами:

$$EDC(x) = \sum_{i=0}^{31} a_i x^i = I(x) \cdot x^{32} \bmod G(x), \quad \text{где} \quad I(x) = \sum_{i=32}^{16511} a_i x^i,$$

$a_i$  – весовой множитель двоичного кода,  $x$  – степенной множитель полинома,  $i$  – номер двоичного символа в блоке. При этом все операции умножения выполняются по модулю порождающего полинома  $G(x)$ .

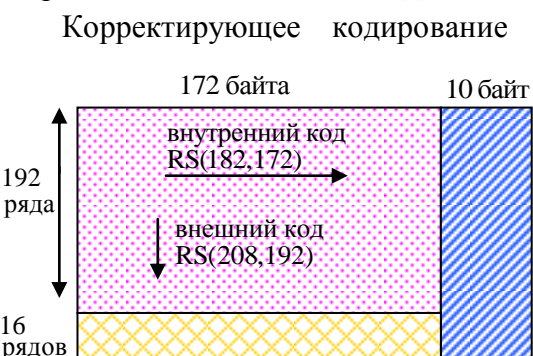


Рис.8.24. Код 2RS-PC в формате DVD

Корректирующее кодирование осуществляется с помощью кода-произведения (Product Code) на основе RS кодов, который называют кодом RS-PC (рис.8.24). Этот код, как и код CIRC недвоичный в поле Галуа  $GF(2^8)$ . Он отличается тем, что вычисление проверочных символов производится не с помощью проверочных матриц, а с помощью порождающих полиномов. Эффективность его значительно выше, чем у кода CIRC.

Для кодирования организуется блок данных в виде матрицы, в которой длина строки равна 172 байта и число рядов - 192 (всего 33024 байт данных). Внешний код (по столбцам) RS(208,172) с минимальным кодовым расстоянием 17 имеет 16 проверочных символов, которые вычисляются по следующим формулам:

$$R_j(x) = \sum_{i=192}^{207} B_{i,j} x^{207-i} = I_j(x) \cdot x^{16} \bmod G_{po}(x), \quad \text{где} \quad I_j(x) = \sum_{i=0}^{191} B_{i,j} x^{191-i},$$

$i, j$  – номера символов по строкам и столбцам,  $B_{i,j}$  – символ в поле  $GF(2^8)$ . Порождающий полином определяется равенством

$$G_{po}(x) = \prod_{k=0}^{15} (x + \alpha^k),$$

где  $\alpha^k$  – примитивный элемент примитивного полинома  $P(x) = X^8 + X^4 + X^3 + X^2 + 1$ .

Внутренний код (по строкам) RS(182, 172) с минимальным кодовым расстоянием 11 имеет 10 проверочных символов, которые вычисляются по формулам:

$$R_i(x) = \sum_{j=172}^{181} B_{i,j} x^{181-j} = I_i(x) \cdot x^{10} \bmod G_{p1}(x), \quad \text{где} \quad I_i(x) = \sum_{j=0}^{171} B_{i,j} x^{171-j},$$

$$G_{p1}(x) = \prod_{k=0}^9 (x + \alpha^k) \text{ - порождающий многочлен.}$$

В стратегии декодирования предусмотрена возможность многократных попыток исправления ошибок. Поэтому используемый код позволяет осуществлять декодирование при вероятности ошибочных бит (BER) равной  $10^{-2}$ . Вероятность случайных ошибок на выходе не превышает  $10^{-15}$ , тогда как в для дисков CD она в 10 раз больше и равна  $10^{-14}$ . Максимальная длина полностью исправляемых выпадений равна 500 байт, что соответствует длине дефекта 2,4 мм. С помощью интерполяции маскируются выпадения длиной до 2200 байт (4,6 мм).

## 8.7. Стандарты канальной модуляции CD и DVD

В стандарте CD используется канальный код EFM (Eight to Fourteen Modulation). Он относится к группе RLL кодов вида  $(d,k,m,n)$ . В этом коде  $d = 2$ ,  $k = 10$ ,  $m = 8$ ,  $n = 14$ , поэтому он часто обозначается как  $(2,10,8,14)$ . При таких значениях параметров кода  $d$  и  $k$  минимальное число последовательных 0 в кодовой таблице равно 2, а максимальное-10. Это значит, что в цифровом потоке минимальное число последовательных 0 или 1 равно 3,

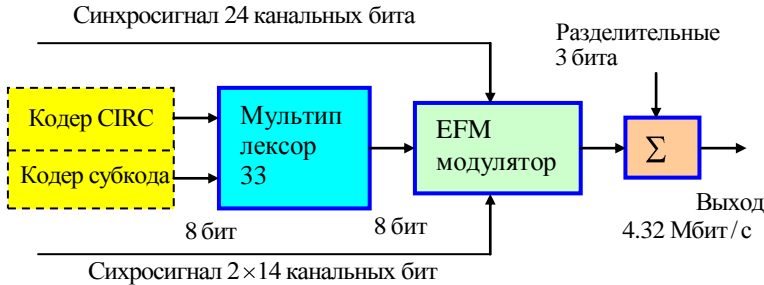


Рис.8.25. Схема канальной модуляции

а максимальное –11.

Канальное кодирование заключается в преобразовании последовательности символов по 8 бит данных в символы по 14 канальных бит. При этом число возможных кодовых комбинаций определяется как  $2^n = 2^{14}$ . Алфавит кода включает в себя только  $2^m = 256$

комбинаций, которые выбираются из 16384 возможных. Выбираются только те комбинации, которые удовлетворяют параметрам кода  $d$  и  $k$ , определяющих допустимое число последовательных 1 и 0. Таких комбинаций 267, поэтому из них 11 остаются неиспользованными.

Блок-схема канального модулятора приведена на рис.8.25. Кодирование в нем осуществляется с помощью электронной таблицы с 256 входами и 1 выходом. После корректирующего кодирования мультиплексор формирует непрерывный поток символов по 8 бит, который подается на вход модулятора. Данные служебной информации не кодируются кодом RS, в блок данных не входят и подаются на вход модулятора отдельно.

Каждой входной кодовой комбинации соответствует только одна комбинация из 14 канальных бит. Для примера приводим небольшой фрагмент этой таблицы (табл.8.9). Считывание данных из этой таблицы производится обязательно канальном коде NRZI (Non Return to Zero Inverted), в котором изменение уровня выходного сигнала модулятора происходит при каждой 1 в таблице.

При использовании такой таблицы кодирования возникают проблемы на стыках кодовых слов, когда один символ заканчивается длинной последовательностью нулей, а следующий начинается с нулей. При этом нарушаются правила кода и декодировать такую последовательность невозможно.

Таблица 8.9. Фрагмент таблицы		
Десятичное значение	Биты данных 8 бит	Канальные биты 14 бит
0	00000000	01001000100000
1	00000001	10000100000000
2	00000010	10010000100000
3	00000011	10001000100000

Поэтому в коде EFM предусмотрено введение разделительных 3 канальных бит между всеми символами. Дополнительные биты могут формировать следующие кодовые комбинации: 000, 010 и 001. Эти же биты используются для минимизации текущей цифровой суммы DSV.

Фрейм данных на выходе канального модулятора формируется путем преобразования 33 символов блока данных после корректирующего кодирования в 588 канальных бит. Из них информацию несут только 462 бита, 24 формируют синхросигнал фрейма и 102 являются разделительными битами, которые никакой информации не несут. При декодировании все разделительные биты сразу отбрасываются.

Необходимость минимизации постоянной составляющей спектра сигнала в канальном коде связана с двумя обстоятельствами. Во первых, грязь, царапины и пыль на поверхности компакт диска создают шумы в области очень низких частот. Их можно фильтровать, только при условии, что в этом спектре нет основного сигнала. Во-вторых, область низких частот, вплоть до постоянной составляющей, используется для работы



САР. Если не минимизировать основной сигнал в области НЧ будет оказывать мешающее действие на работу систем регулирования.

Минимизация осуществляется по специальному алгоритму путем подбора кодовых комбинаций разделительных бит. При этом подборе контролируется текущее значение DSV, которое рассчитывается как текущая сумма 1 и 0 (0 принимается за минус 1). Считается, что в спектре нет постоянной составляющей, если  $DSV < 7$ .

В табл.8.10. приведены для сравнения основные характеристики кодов EFM и 8/16, используемого в стандартах DVD(SACD). Как видно они практически одинаковы.

Таблица 8.10. Параметры кодов		
Параметры	EFM	8/16
(d,k,m,n)	2,10,8,14	2,10,8,16
Избыточность кода, %	42,8	50
Число разделительных бит,	3	0
Число таблиц кодирования	1	5
Скорость канального кода	8/17	8/16
Коэф. плотности канального кода	1,41	1,5
Коэф. самосинхронизации,	3,66	
Число кодовых интервалов времени	9	

В связи с тем, что в коде EFM вводятся дополнительные биты, не несущие информации, его коэффициент плотности рассчитывается по формуле

$$K_d = \frac{(d+1) \cdot m}{n+j}$$

где  $j$  – число разделительных бит в символе. При этом  $K_d = 1,41$  и благодаря этому при скорости 1,2 м/с достигается максимальная линейная плотность 1693 бит/мм. В коде 8/16 разделительных бит нет, поэтому коэффициент плотности больше.

Сравнительные временные характеристики кодов приведены в табл.8.11. Они отличаются примерно в 6 раз, что связано только с различием скоростей цифровых потоков аудиоданных.

Таблица 8.11. Временные характеристики кодов		
Параметры	EFM	8/16
Скорость аудиоданных, $V_a$ , Мб/с	1,41	9,6
Скорость данных, $V_d$ , Мб/с	2,03	13,1
Канальная тактовая частота, МГц	4,3218	26,16
Окно детектирования, нс	231	38,3
Мин. временной интервал, нс	694	115
Макс. временной интервал, нс	2540	421
Макс. частота спектра, МГц	0,72	4,35
Мин. частота спектра, МГц	0,196	1,19

Параметры  $d$  и  $k$ , определяющие спектр канального кода, выбраны для кода EFM так, чтобы верхняя и нижняя границы спектра находились в середине частотного диапазона ОПФ (рис.8.4.). В спектре кода присутствуют 9 частот в

диапазоне от 196 до 720 кГц. Период этих частот меняется дискретно с интервалом 231 нс. Частота среза ОПФ

составляет 1,4 МГц и частотный диапазон канального кода находится ниже половины частоты среза. Поэтому частотные искажения не очень большие. Для оптимальной фильтрации важно, что огибающая спектра кода EFM и ОПФ достаточно близки по форме, как это видно из рис.8.26.

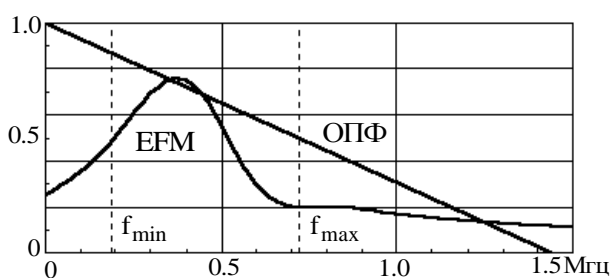


Рис.8.26.Спектральная плотность кода EFM

Помехозащищенность кода EFM определяется окном детектирования шириной 231 нс. Это значит, что пока амплитуда джиттера не превысит значения 115 нс, декодирование осуществляется без ошибок. Декодирование канального кода может осуществляться с помощью электронной таблицы или логической схемы. Такая таблица имеет 256 входов и 1 выход. Разрешенные кодовые комбинации

являются адресом входа.

Если при декодировании произошла ошибка, которая разрешенную кодовую комбинацию преобразовала в запрещенную, то в коде EFM она обнаруживается с вероятностью 1. Если ошибка преобразовала одну разрешенную комбинацию в другую, но тоже разрешенную, то такая ошибка не обнаруживается. Средняя вероятность обнаружения ошибок в коде EFM равна 0,9844. Высокая избыточность канального кода позволяет использовать его как первую ступень обнаружения ошибок. Все обнаруженные

ошибки стираются и эти стирания могут быть исправлены уже в 1 ступени декодера, исправляющего кодовые ошибки.

В стандартах DVD-Audio и Super Audio CD при канальной модуляции используется код 8/16, который иногда называют EFM plus. Это групповой канальный код (2,10,8,16), относящийся к группе RLL кодов. В этом коде каждые 8 бит данных преобразуются в 16 канальных бит. В коде 8/16  $d=2$  и  $k=10$ , как и в коде EFM. Это значит, что число нулей подряд в кодовой таблице не может быть меньше 2 и больше 10. При выполнении этого условия, в канальном коде не будет подряд меньше 3 и больше 11 последовательных 1 или 0.

Как видно из табл.8.10 у кодов 8/16 и EFM почти одинаковые параметры, относящиеся непосредственно коду. Только благодаря отсутствию разделительных бит у кода 8/16 несколько больше коэффициент плотности записи. В стандарте DVD-A значительно больше скорость записываемого потока цифровых данных (9,6 Мбит/с), поэтому тактовая частота канального кода увеличена до 26 МГц. В результате ширина окна детектирования уменьшилась до 38 нс и следовательно, уменьшилась помехозащищенность. Это значит, что для обеспечения надежности декодирования необходимо предъявлять очень высокие требования к SNR тракта и, особенно, к качеству записи на оптических дисках. Перед записью данные организуются в виде фрейма, в который входят 1488 канальных бит, из них 32 формируют синхросигнал и 1456 относятся к общим данным.

Главное отличие кода 8/16 в том, что в нем нет разделительных бит между канальными символами. Поэтому для модуляции приходится использовать не 1, а 5

Таблица 8.12. Таблица состояний символов кода 8/16

Канальный символ X(t)	Состояние X(t+1)	Канальный символ X(t+1)
На конце 0 или 1 «0»	1	В начале от 2 до 9 «0»
На конце от 2 до 5 «0»	2	В начале от 1 до 5 «0»
На конце от 2 до 5 «0»	3	В начале нет или от 1 до 5 «0»
На конце от 6 до 9 «0»	4	В начале нет или 1 «0»

таблиц кодирования. Для этого все выходные кодовые комбинации модулятора из 16 канальных бит разделяются на 4 состояния, приведенные в табл.8.12.

В соответствии с этой таблицей, для того чтобы не нарушались правила кодов RLL после символа, имеющего состояние 1, следующий символ может начинаться с 2 и до 9 нулей подряд. Если символ заканчивается от 2 до 5 нулей подряд (состояние 2), то следующий символ может иметь на входе от 1 до 5 нулей. По расчетам число символов, относящихся к 1 состоянию – 352, к 2, 3 и 4 состоянию относятся – 352, 351 и 415 соответственно, т.е. минимальное число равно 351. В коде 8/16 используется 32-битовое синхрослово, поэтому часть возможных кодовых комбинаций теряется и могут использоваться всего 344 символа. Общее число полезных кодовых комбинаций равно 256, поэтому остаются еще 88 слов, которые могут использоваться для минимизации спектральной плотности мощности на низких частотах, как это делается в коде EFM, путем контроля текущей цифровой суммы DSV.

Таблица 8.13. Фрагмент таблицы кода 8/16

i	Состояние 1	i	Состояние 1	i	Состояние 3	i	Состояние 4	i
0	0010000000001001	1	0100000100100000	2	0010000000001001	1	0100000100100000	2
1	0010000000010010	1	0010000000010010	1	1000000100100000	3	1000000100100000	3
2	0010000100100000	2	0010000100100000	2	1000000000010010	1	1000000000010010	1
3	0010000001001000	2	0100010010000000	4	0010000001001000	4	0010000001001000	4

В кодер 8/16 входят 5 таблиц. Главная таблица предназначена именно для кодирования, она имеет 256 входов ( $i=0...255$ ), на которые подаются 256 возможных кодовых комбинаций из 8 бит. Эти комбинации одновременно являются адресом входа таблицы. Таблица кодирования 8.13. состоит из 4 подтаблиц, содержащие канальные символы для возможных 4 состояний. Символы представлены в коде NRZI.

В качестве примера канального кодирования по табл.8.13. допустим, что предыдущий символ имел состояние 1 и следующий символ подается на вход 3. Тогда выходной канальный символ модулятора выбирается из 1 подтаблицы и будет 0010000001001000 с индексом 2, указывающим на состояние этого символа. Это значит, что следующий символ необходимо выбирать из 2 подтаблицы.

Вспомогательная таблица используется для замены только первых 88 символов главной таблицы при необходимости коррекции DSV. В алгоритме коррекции предполагается, что исходная DSV равна нулю. В случае отличия этой суммы от нуля она минимизируется через 2 канальных символа, тогда как в коде EFM это достигается через 4 символа. Первые 88 символов главной таблицы имеют отрицательное значение DSV, а подстановки из вспомогательной таблицы имеют положительное значение. Если расчет DSV производится по трем символам, то достигается снижение спектральной плотности мощности в области низких частот на 3 дБ по сравнению со спектром кода EFM. Постоянная составляющая кода 8/16 лишь немного больше, чем у кода EFM.

## 9. Магнитная запись звука

### 9.1. Физические основы цифровой магнитной записи

Цифровая магнитная звукозапись отличается от аналоговой прежде всего сильным расширением требований к частотному диапазону носителя записи. Так при частоте дискретизации 44,1 кГц и 16 разрядном кодировании полоса частот для записи стереофонического сигнала увеличивается с 20 кГц до 1 МГц. Если производится 6-канальная запись с использованием 24 разрядного кодирования с частотой дискретизации 192 кГц, то требуется частотный диапазон выше 20 МГц и емкость носителя для часовой записи больше 13 Гбайт. Последние 25 лет поверхностная плотность магнитной записи удваивается каждые 2 года и в настоящее время достигала  $6 \cdot 10^7$  бит/мм<sup>2</sup>. При этом происходит не менее  $3 \cdot 10^4$  изменений намагниченности на каждый миллиметр длины фонограммы.

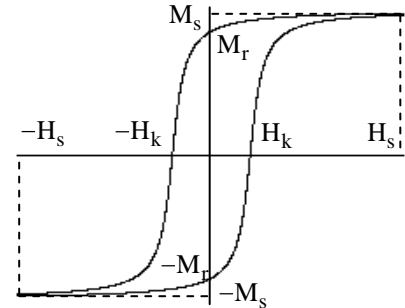


Рис.9.1. Петля гистерезиса

Магнитная запись звука основана на свойствах ферромагнитных материалов намагничиваться под действием внешнего магнитного поля, и оставаться намагниченными после выхода из этого поля. Такое поле возникает при протекании тока по обмотке катушки с сердечником из магнитного материала, который имеет воздушный зазор. В системах звукозаписи обычно используются замкнутые *кольцевые сердечники*, в которых вблизи зазора возникает внешнее *поле рассеивания*, намагничивающее НЗ, движущийся в этом поле. При этом изменения тока катушки превращаются в изменения остаточной намагниченности фонограммы.

Перемагничивание при цифровой звукозаписи происходит по предельной петле гистерезиса (рис.9.1.), поэтому магнитные носители характеризуются 4 основными параметрами: коэрцитивная сила  $H_k$ , напряженность насыщения  $H_s$ , остаточная намагниченность  $M_r$  и прямоугольность петли  $K_p$ . Это значит, что при записи напряженность магнитного поля  $H$  должна превышать напряженность насыщения материала носителя  $H_s$ .

Носитель магнитной записи состоит из немагнитной подложки 1 и магнитного покрытия 2 (рис.9.2.). Между ним и поверхностью головки всегда существует некоторый

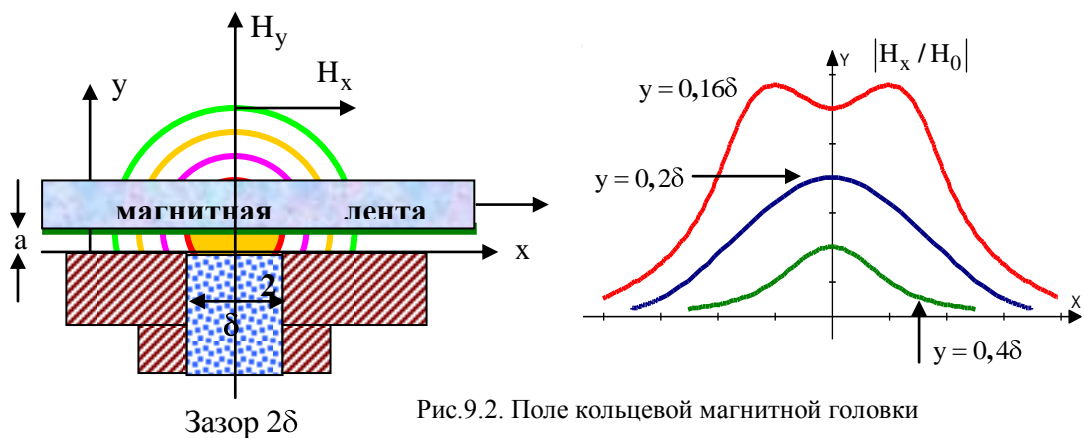


Рис.9.2. Поле кольцевой магнитной головки

зазор  $a$ . Внешнее магнитное поле кольцевой головки имеет вид окружностей, поэтому оно раскладывается на две составляющие – горизонтальную  $H_x$  и вертикальную  $H_y$ . Напряженность поля внутри зазора  $H_\delta$  тем выше, чем больше число витков катушки

магнитной головки  $W$  и ток записи  $I_{rec}$ , а также и чем меньше ширина зазора головки  $2\delta$

$$H_{\delta} = \frac{I_{rec} \cdot W}{2\delta \left(1 + \frac{R_s}{R_{\delta}}\right)},$$

где  $R_s$  – магнитное сопротивление сердечника,  $R_{\delta}$  – магнитное сопротивление зазора.

Под действием напряженности магнитного поля катушки  $H$  носитель записи получает намагниченность  $M$ :

$$M = \rho H_x,$$

где  $\rho = \mu - 1$ ,  $\rho$  – магнитная восприимчивость ленты,  $\mu$  – магнитная проницаемость материала рабочего слоя (РС). При записи в этом слое создается остаточный магнитный поток, определяемый равенством

$$\Phi_r = \mu_0 K_p M_g b_m d_m K_v, \quad K_p = \frac{M_r}{M_s},$$

где  $b_m$  – ширина дорожки записи,  $d_m$  – толщина РС магнитной ленты,  $M_g$  – намагниченность частиц,  $K_v$  – коэффициент объемного заполнения, показывающий, какую часть РС составляет магнитное вещество,  $K_p$  – коэффициент прямоугольности петли гистерезиса. Из формулы следует, что увеличить остаточный магнитный поток можно увеличением коэффициента прямоугольности, намагниченности частиц и коэффициента объемного заполнения.

Горизонтальная  $H_x(x, y)$  и вертикальная  $H_y(x, y)$  составляющие напряженности магнитного поля определяются равенствами

$$H_x(x, y) = -\frac{H_0}{\pi} \cdot \left[ \arctg \left( \frac{1 + \frac{2x}{2\delta}}{\frac{2y}{2\delta}} \right) + \arctg \left( \frac{1 - \frac{2x}{2\delta}}{\frac{2y}{2\delta}} \right) \right], \quad H_y(x, y) = \frac{H_0}{2\pi} \cdot \ln \left[ \frac{\left(1 + \frac{2x}{2\delta}\right)^2 + \left(\frac{2y}{2\delta}\right)^2}{\left(1 - \frac{2x}{2\delta}\right)^2 + \left(\frac{2y}{2\delta}\right)^2} \right]$$

где  $x, y$  – координаты. При расстояниях между головкой и носителем больше 0,1 ширины зазора горизонтальная составляющая существенно больше вертикальной, что

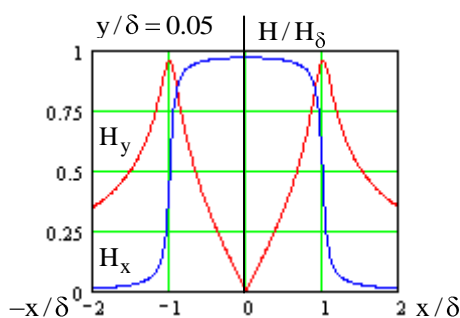


Рис.9.3. Вертикальная и горизонтальная составляющие напряженности магнитного поля

характерно для аппаратуры аналоговой звукозаписи. В современной аппаратуре цифровой звукозаписи величина зазора между плавающей магнитной головкой и НЗ может быть всего 25 ангстрем. На рис.9.3. приведены графики составляющих напряженности для  $y/\delta = 0.05$ , когда их максимальные значения равны. Это позволяет использовать три вида цифровой звукозаписи.

*Запись с продольным намагничиванием* является основой современной аналоговой звукозаписи. К сожалению она сопровождается достаточно сильным эффектом саморазмагничивания, возникающим уже в звуковом диапазоне частот. Этот эффект заключается в возникновении после записи поля внутри РС магнитной ленты, которое направлено навстречу полю фонограммы, Оно вызывает уменьшение остаточной намагниченности и, следовательно, отдачу ленты. Это объясняется тем, что с уменьшением длины волны

записи расстояние между противоположными полюсами намагненности становится меньше, эффект саморазмагничивания увеличивается, а уровень считываемого сигнала

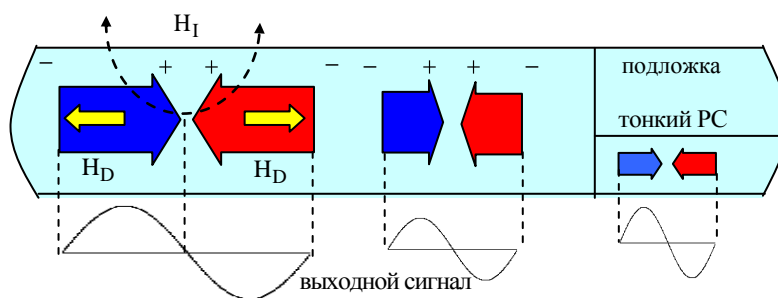


Рис.9.4. Запись с продольным намагничиванием - эффект саморазмагничивания

уменьшается (рис.9.4). Напряженность этого поля определяется равенствами:

$$H_{\approx} = K_{dm} \cdot M_r, \quad K_{dm} = \frac{1}{1 + \frac{\lambda_{rec}}{2d_m}}$$

где  $K_{dm}$  - коэффициент размагничивания,  $\lambda_{rec}$  - длина волны записи.

Из приведенного равенства следует, что саморазмагничивание проявляется сильнее на высоких частотах, когда величина  $\lambda_{rec}$  приближается к толщине PC  $d_m$ . Поэтому для

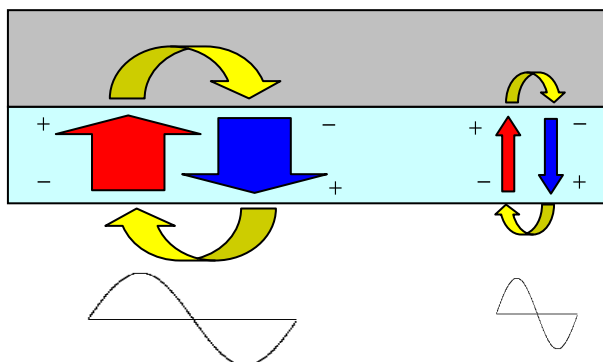


Рис.9.5. Запись и считывание с вертикальным намагничиванием

расширения частотного диапазона необходимо использовать носители с очень тонким металлическим покрытием, величина которого составляет доли длины волны записи. Важно, что с уменьшением толщины PC коэффициент саморазмагничивания быстро уменьшается, и уже при  $d_m = 0,5 \cdot \lambda_{rec}$  он равен 0,5.

Носители с тонким PC изготавливаются на основе металлических покрытий с достаточно сложной технологией. У таких носителей велико

значение  $SNR$ , из-за уменьшения уровня сигнала. Все дефекты подложки переходят в дефекты PC, они подвержены коррозии, и велика опасность повреждения из-за контакта с магнитной головкой.

**Запись с вертикальным намагничиванием** применяется только в цифровой аппаратуре с очень малым зазором между носителем и магнитной головкой. При этой записи на носителе регистрируется вертикальная составляющая напряженности магнитного поля. В качестве основного магнитного покрытия в этом случае используется тонкая пленка на основе сплава кобальта и хрома ( $CoCr$ ) с анизотропией перпендикулярной его поверхности с размером частиц около 0,5 мкм. Дополнительное относительно толстое покрытие из сплава никеля с железом ( $NiFe$ ) предназначено для записи горизонтальной составляющей поля намагничивания (рис.9.5.) и уменьшения влияния дефектов подложки. Полезная толщина этого покрытия, в котором заключено 90% считываемой информации, при записи с

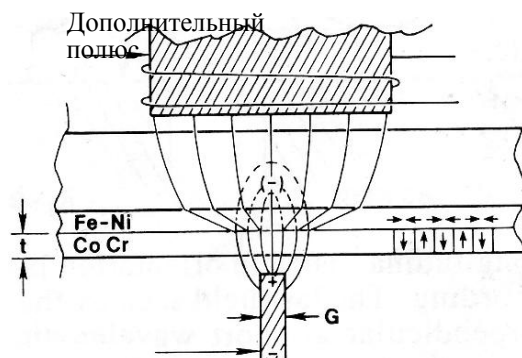


Рис.9.6. Запись с вертикальным намагничиванием 2 головками

вертикальным намагничиванием составляет всего  $\lambda/e$ . С увеличением продольной плотности записи свыше  $2 \cdot 10^6$  бит/мм эта толщина уменьшается до  $\lambda/2$ .

При данном способе записи эффект саморазмагничивания сильнее на низких частотах и он существенно уменьшается с частотой, что связано с возрастающей ролью горизонтальной составляющей намагниченности в формировании сигнала считывания (рис.9.5.).

Обычно для записи используются *две однополюсные головки без зазоров*, с разных сторон носителя (рис.9.6.). Записываемый сигнал подается на вспомогательную головку с полюсом достаточно больших размеров. Запись производится однополюсным магнитом очень малых размеров. Известны разработки, когда оба полюса расположены с одной стороны. Запись с вертикальным намагничиванием может производиться и с помощью кольцевой магнитной головки, как это делается по стандарту MD (Mini-Disk).

*Запись с вертикальным и горизонтальным намагничиванием (изотропическая запись)* применяется только в цифровой аппаратуре с очень малым зазором между носителем и магнитной головкой. В этом случае используется изотропический носитель, у которого магнитные свойства одинаковы во всех направлениях. В этом способе важно, чтобы при считывании записи составляющие напряжения, обусловленные разными

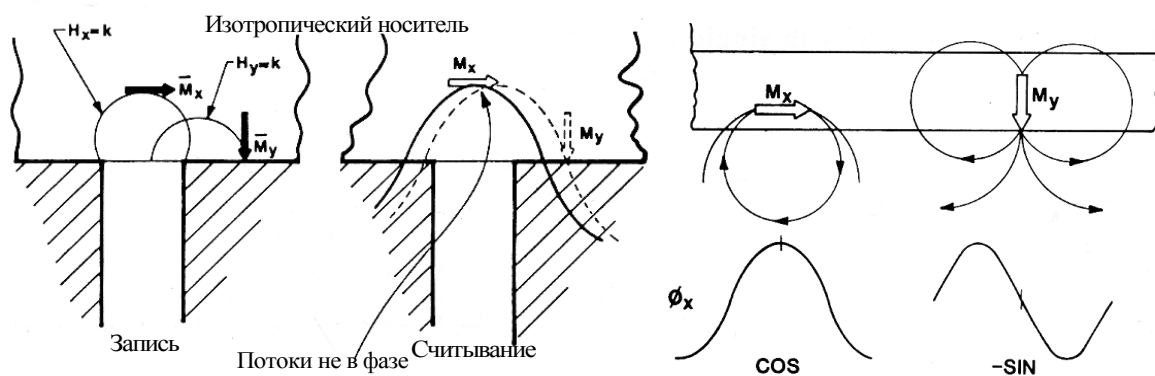


Рис.9.7. Изотропическая запись

векторами намагниченности, были не в противофазе. В процессе записи вертикальная составляющая поля стремится стереть все горизонтальные составляющие вблизи поверхности головки.

Поэтому возникают два центра намагниченности (рис.9.7), сдвинутые относительно друг друга. При считывании такой записи в выходном сигнале присутствуют составляющие, обусловленные различными векторами намагниченности, которые связаны между собой преобразованием Гильберта. От одной составляющей возникает сигнал в виде четной функции (cos), а от другой – нечетной (sin). Они близки по фазе и усиливают друг друга, когда запись производится на высоких частотах с продольной плотностью записи выше  $3 \cdot 10^6$  бит/мм, используется магнитное покрытие с высокой коэрцитивной силой с коэффициентом прямоугольности около 0,85, а зазор головки не более 0,15 мкм.

*Физика цифровой магнитной записи.* Принципиальных различий в физических основах аналоговой и цифровой магнитной записи нет. Разница только в том, что цифровой звуковой сигнал, подлежащий записи, представляет собой последовательность двухполярных импульсов в канальном коде с частотой во много раз превышающей звуковой диапазон. В этом коде цифровая информация передается путем дискретного изменения длительности положительных и отрицательных импульсов. Логическая «1» соответствует положительной полярности импульса, а логический «0» – отрицательной. Логические «1» и «0» намагничивают РС носителя записи до насыщения:  $+M_s$  и  $-M_s$  (рис.9.8.),

При такой записи информации нелинейность намагничивания от напряженности магнитного поля здесь не имеет значения. Благодаря этому нет необходимости использовать ВЧ подмагничивание. Не нужно производить стирание записи отдельной магнитной головкой, так как эта операция выполняется автоматически во время записи за счет намагничивания РС до насыщения. Для записи и воспроизведения могут использоваться или отдельные головки, или одна универсальная.

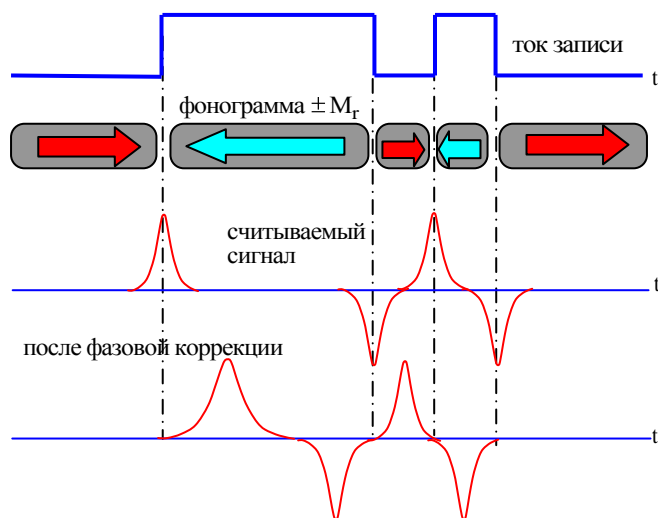


Рис.9.8. Цифровая магнитная запись и считывание

При воспроизведении цифровой записи с помощью индукционной ГВ ЭДС на обмотке головки пропорциональна производной магнитного потока по времени (закон Фарадея). Вследствие такой операции дифференцирования на выходе ГВ формируются двухполярные узкие импульсы, называемые характеристическими. Форма этих импульсов достаточно хорошо аппроксимируется законом Гаусса. Важно обратить внимание на то, что максимумы этих импульсов совпадают с временем смены намагниченности. Восстановление же прямоугольной формы воспроизводимых импульсов обычно производится по моментам перехода воспроизводимого сигнала через нулевые значения с помощью компаратора нуля.

Поэтому воспроизводимый сигнал сдвигается по фазе

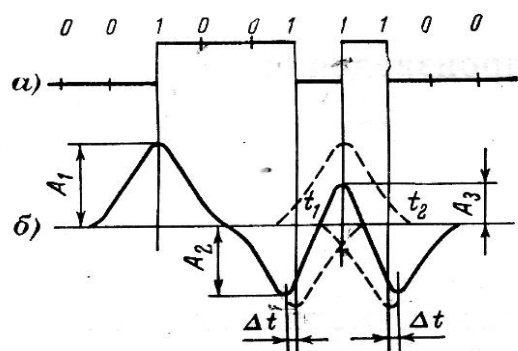


Рис.9.10. Межсимвольная интерференция

относительно сигнала записи на  $\pi/2$ , и моменты перехода через нуль записываемого и воспроизводимого сигналов не совпадают. Этот фазовый сдвиг можно исключить с помощью еще одной операции дифференцирования или фазовой коррекции и тогда моменты перехода через нуль этих сигналов совпадают (рис.9.8).

Для воспроизведения цифровой звукозаписи с магнитной ленты обычно используются индукционная ГВ, которая в области низких частот имеет АЧХ в виде прямой линии с крутизной подъема +6 дБ/октаву (рис.9.9). Поэтому в аналоговой записи всегда используется частотная коррекция для выравнивания АЧХ в области низких частот. Так как цифровой сигнал воспроизводится в виде импульсов, то низкочастотная коррекция тракта воспроизведения в цифровых магнитофонах не нужна.

При цифровой магнитной записи спектр записываемого сигнала может простирается до 10 ... 250 МГц, поэтому используются и более высокие скорости записи до 5 м/с. Тем не менее, при воспроизведении такой записи возникают все виды волновых потерь. В отличие от аналоговой записи здесь важно иметь в виду, что частотные и фазовые

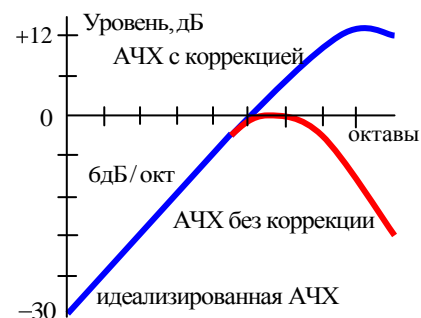


Рис.9.9. АЧХ магнитной головки



искажения при цифровой звукозаписи никакого влияния на частотную характеристику звукового тракта не оказывают. Они приводят к возникновению межсимвольной интерференции, вызывающей смещению моментов перехода выходного сигнала через нулевое значение и, как следствие, к кодовым ошибкам (рис.9.10.). Чем больше частотные и фазовые искажения в цифровом тракте, тем больше джиттер в виде дрожания фронтов импульсов и, следовательно, больше вероятность возникновения ошибок, вызывающих трески. Поэтому в цифровых магнитофонах волновые потери и фазовые искажения надо также компенсировать с помощью коррекции.

**Магнитные головки** для цифровой звукозаписи сначала изготавливались только из феррита. Однако, ферриты имеют очень маленькую индукцию насыщения и не могут использоваться для записи цифровой информации на материалы с коэрцитивной силой более 80 кА/м, поэтому в настоящее широко используются другие материалы, характеристики которых приведены в табл.9.1. Сендаст (AlFeSi –сплав) и пермаллой (Ni<sub>80</sub>Fe<sub>20</sub>) применяются в двух типах головок: с металлизированным зазором MiG (Metal-in-Gip) и тонкопленочными, у которых индукция насыщения вдвое больше, чем у феррита. Наиболее перспективно использование ламинированных материалов с изолирующими слоями такие как FeAlN/SiO<sub>2</sub> и FeN/SiO<sub>2</sub> с частотным диапазоном до 200 МГц.

Т а б л и ц а 9.1. Материалы магнитных головок

Материал	B <sub>s</sub> (Тесла)	H <sub>s</sub> (А/м)
Сендаст	1	0,15
Пермаллой	1	30
Аморфный CoZr	1,4	4
Нитриды железа	2	40
CoFe	2,4	800
Fe <sub>16</sub> N <sub>2</sub>	2,8	5

Для воспроизведения цифровой записи с магнитных дисков в последнее время все чаще используются головки воспроизведения с гигантским магниторезистивным эффектом, известные под названием GMR (Giant Magnetoresistive). В этих головках под

действием магнитного поля фонограммы в электрическом проводнике возникает явление сверхпроводимости, поэтому они обладают очень высокой чувствительностью. Важно, что АЧХ таких головок не имеет завала в области низких частот.

**Магнитные носители** для цифровой звукозаписи подразделяются на порошковые и металлические. Порошковый РС обычно изготавливается из гамма-окисла железа, у которого  $H_k = 22,5$  кА/м, а  $B_r = 0,09...0,14$  Т. Применяется также феррит-кобальт, у которого  $H_k = 25...80$  кА/м, и двуокись хрома с  $H_k = 40$  кА/м. Толщина РС от 0,5 до 1 мкм. Для металлических покрытий используются сплавы: Co – W, Co – Ni, и др.. У них  $H_k$  может быть от 16 до 80 кА/м, Толщина таких покрытий 0,1...0,3 мкм, коэффициент прямоугольности равен 0,45...0,75, а магнитная индукция около 1 Т. Для защиты очень тонких металлических покрытий используются немагнитные покрытия из родия толщиной 0,3...0,7 мкм.

## 9.2.Многодорожечные цифровые магнитофоны стандарта DASH

**Катушечные многоканальные многодорожечные магнитофоны** стандарта DASH были разработаны в начале 70-х годов совместно фирмами Sony и Studer. В магнитофонах этой системы используются магнитные ленты стандартной ширины. Число звуковых каналов может быть от 2 до 48, в зависимости от ширины ленты и скорости записи, предусмотрены также два канала аналоговой записи звука для звукорежиссера. По скорости записи возможны 3 режима работы: Fast(F), Medium(M) и Slow(S). Основная частота дискретизации равна 48 кГц. На этой частоте скорость движения лент такая же как в аналоговых студийных магнитофонах: 19,05, 38,1 и 76,2 см/с. Возможна запись с частотами дискретизации 44,1 и 32 кГц. Основные характеристики этого стандарта приведены в табл.9.2 и 9.3.

Т а б л и ц а 9.2. Носитель записи

Режим	F			M			S
Ширина ленты	6,3	12,7	25,4	6,3	12,7	25,5	6,3
Число каналов	8	24	48	4	12	24	2
Число дорожек	8	24	48	8	24	48	8
Ширина дорожки, мкм	240						
Защитный интервал, мкм	40						
Скорость ленты, см/с	76,2			38,1			19,05

В магнитофонах стандарта DASH осуществляется многоканальная продольная запись сразу на большом числе дорожек на движущуюся с большой скоростью магнитную ленту. На каждой дорожке установлены свои ГЗ и ГВ, которые неподвижны, как в аналоговых магнитофонах. Расстояние между головками 0,52 мм, ширина полюсного наконечника ГЗ - 0,37 мм и этой головкой записывается дорожка шириной 0,24 мм. Для того чтобы не предъявлять жестких требований к положению ГВ на дорожке, ее полюсный наконечник сделан значительно меньше ширины дорожки (0,15 мм). Ширина зазоров ГВ равна 0,35 мкм, что обеспечивает возможность воспроизведения записи с высокой плотностью. Головки всех дорожек всегда объединяются в единый блок.

В стандарте DASH осуществляется 16-ти разрядное аналого-цифровое преобразование с использованием линейной ИКМ. Для коррекции ошибок применяется простейший двоичный блоковый код с перемежением CIC (Cross

Т а б л и ц а 9.3. Характеристики стандарта DASH			
Параметры записи		Характеристики цифрового тракта	
Зазор магн. головки, мкм	0,35	Число разрядов на отсчет, q	16
Мин. длина волны записи, мкм	1,9837	Избыточность кода, %	25
Мин. длина бита, мкм	0,6612	Скорость бит данных, Мбит/с	1,74
Лин. плотность записи, бит/мм	1514	Число слов данных в блоке,	17
Длительность записи, мин	60...65	Длительность блока(16/48), мкс	156,2
		Блочная частота, кГц	64

Interleave Code), в котором производится двойная проверка на четность, отдельно для четных и нечетных номеров слов в блоке данных. Общая избыточность кода 25%. Обнаружение ошибок производится с помощью циклического кода CRC с порождающим полиномом

$$G(x) = X^{16} + X^{12} + X^5 + 1 .$$

В блок данных входят: 16 слов звуковой информации, 2 проверочных слова на четность, одно проверочное слово кода CRC и слово со служебной информацией. Длина блока на носителе равна 0,19 мм.

При кодировании используется внутриблочное и межблочное сверточное перемежения с интервалом 204 блока отдельно для четных и нечетных слов. В результате такого перемежения четные и нечетные слова одного блока записываются на носитель отдельно, каждые в области 22,7 мм. Расстояние между четными и нечетными словами на ленте 38,9 мм. Это обеспечивает надежное исправление длинных пакетов ошибок.

При записи данных на носитель используется каналный код HDM-1 (High Density Modulation). Этот код не относится к группе NRZ кодов и не является групповым, но по свойствам похож на групповой код (2,8). Кодирование осуществляется по определенным правилам, легко реализуемым на основе цифровых логических схем. В основе этих правил лежит возможность изменения уровня сигнала посередине между двумя последовательными 1 или 0. В табл.9.4. приведены некоторые данные этого кода, рассчитаны для скорости цифрового потока бит

Т а б л и ц а 9.4. Канальный код HDM-1		
Канальный код (d, k)	(2, 8)	
Скорость канального кода	0,5	
Период данных, нс	T	574,7
Окно детектирования, нс	0,5T	287
Дискретность интервалов	0,5T	287
Число интервалов	6	
Мин. интервал, мкс	1,5 T	0,86
Макс. интервал, мкс	4,5 T	2,59
Макс. частота спектра, кГц	1 / 3 T	580
Мин. частота спектра, кГц	1 / 9 T	174
Кэф. самосинхронизации	3	
Плотность записи, бит/мм	1511	
Длина волны записи/ мкм	1,98	

данных 1,74 Мбит/с. Из таблицы видно, что характеристики этого канального кода почти точно такие же, как у канального кода EFM в стандарте CD. У него только не контролируется постоянная составляющая, которая достаточно велика. В магнитной записи эта проблема не является столь важной, как в оптических системах.

Магнитофоны стандарта DASH разрабатывались в 70-х годах, когда еще не было известно групповое канальное кодирование и не было технической возможности реализации кодов RS. Тем не менее, их технические характеристики удовлетворяют потребителей, и они достаточно широко используются в студийной технике. В значительной мере это связано с тем, что использованные методы корректирующего кодирования и канальной модуляции позволяют осуществлять ручной монтаж фонограмм с помощью ножниц. Именно это обстоятельство было одним из основных требований при разработке магнитофонов этого стандарта, так как в то время еще не был освоен электронный монтаж.

Исторически первой моделью магнитофона системы DASH был аппарат Sony PCM-3324. Это 24-дорожечный магнитофон с 4 дополнительными режиссерскими дорожками. При записи на штатную ленту в катушке диаметром 14 дюймов уместается 65 минут записи звука. Дальнейшим развитием системы стал стандарт DASH-F, использованный в модели Sony PCM-3324S. В этой модели используется 16-разрядные АЦП и 18-разрядные ЦАПы. Позже эта фирма выпустила в продажу 48-дорожечные магнитофоны PCM-3348.

В настоящее время выпускается магнитофон модели PCM-3348HR с 24-разрядными АЦП и ЦАП. Фирма Studer также выпускает 24-разрядные магнитофоны модели D827 MCH. Высокое качество записи такого большого числа дорожек с высокой надежностью пока не обеспечивает ни один конкурирующий стандарт. Магнитофоны такого класса – техника не дешевая, элитная. Для серьезной профессиональной работы магнитофоны стандарта DASH являются пока наиболее оптимальным вариантом.

### 9.3. Стереофонические магнитофоны стандарта DAT Наклонно-строчная запись

Этот стандарт записи был предложен фирмой Sony и раньше он назывался R-DAT (Rotary Digital Audio Tape). В магнитофонах этого стандарта *стереофоническая звукозапись* производится с мультиплексированием двух звуковых каналов на *медленно движущийся* ленточный НЗ с помощью двух вращающихся магнитных головок.

Т а б л и ц а 9.5. Характеристики стандарта DAT			
Магнитная лента		Параметры записи	
- толщина магнитного покрытия, мкм	2,6...3	Рабочий зазор магн. головки, мкм	0,25...0,3
- ширина ленты, мм	3,81	Линейная скорость записи, м/с	3,1333
- длина ленты, м	58,6	Мин. длина бита на ленте, мкм	0,416
Формат записи- наклонно-строчный		Линейная плотность записи, бит/мм	2400
- ширина дорожки, мкм	13,591	Скорость движения ленты, мм/с	8,15
- длина дорожки, мм	23,5	Длительность записи, мин	120
- угол наклона дорожки, град	6°22'		

Эти же головки используются для считывания записи. По стандарту магнитофон может работать в режиме записи или воспроизведения. В профессиональных вариантах исполнения осуществляется одновременная запись и считывание с помощью 4-х вращающихся головок.

Характеристики ленты и формата записи приведены в табл.9.5. Запись производится на ленте шириной 3,81 мм и толщиной 13 мкм, находящейся в кассете размером 75 x 54 x 10,5 мм. При ускоренной перемотке этой ленты скорость ее движения увеличивается в 200 раз. По краям ленты имеются 2 дополнительные аналоговые звуковые дорожки для звукорежиссера. Цифровая звукозапись осуществляется в центральной зоне ленты виде

наклонных дорожек. На **рис.9.11.** направления движения ленты и записи дорожек указаны стрелками. Дорожки на ленте

расположены с наклоном в направлении движения ленты *без промежутков*, такая запись называется *наклонно-строчечной*.

Ширина полюсного наконечника головки существенно больше ширины дорожки, равной 13,5 мкм, поэтому запись дорожек производится со значительным наложением.

Запись без промежутков между дорожками возможна благодаря тому, что зазоры двух записывающих головок имеют противоположные наклоны относительно перпендикуляра к направлению записи. Одна имеет наклон  $+20^\circ$ , а другая  $-20^\circ$ . Это исключает помехи от наложения дорожек.

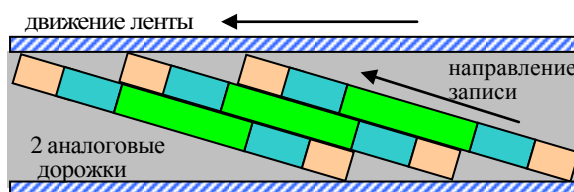


Рис.9.11. Наклонно-строчечная запись

Обе универсальные головки записи-воспроизведения имеют рабочий зазор шириной 0,25...0,3 мкм, поэтому при скорости записи 3 м/с частота среза  $f_o = 10...12$  МГц. Другие характеристики записи приведены в **табл.9.5.** Головки укреплены на барабане диаметром 30 мм под углом  $180^\circ$  (**рис.9.11.**), скорость вращения барабана 2000 об/мин. Магнитная лента прижимается к барабану в пределах угла  $90^\circ$ . Блок головок устроен так, что за время одного оборота барабана равного 30 мс, 50 % времени ни одна из головок А и В не касается магнитной ленты. Контакт одной из головок с магнитной лентой происходит через каждые 15 мс. Длительность контакта равна времени записи одной дорожки 7,5 мс.

Каждая дорожка имеет длину 23,5 мм и содержит информацию о звуке, служебную информацию и специально записанный сигнал с частотой 130 КГц для обеспечения работ системы слежения за дорожкой при воспроизведении. Вся эта информация размещена в 5 зонах дорожки, звуковые данные записываются в центре дорожки. В стандарте DAT предусмотрено 6 вариантов его реализации в режиме стереофонической записи звука, отличающиеся частотой дискретизации, числом разрядов и видом ИКМ. В основном формате используется линейная 16-ти разрядная ИКМ с частотой дискретизации 48 кГц.

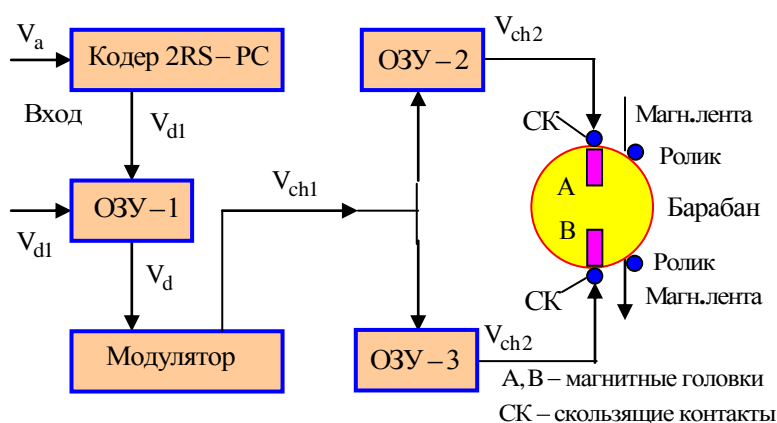


Рис. 9.12. Функциональная схема цифрового тракта магнитофона системы DAT в режиме записи

Функциональные схемы магнитофонов системы DAT, выпускаемые различными фирмами, существенно отличаются. На **рис.9.12.** приведена очень упрощенная функциональная схема цифрового тракта магнитофона в режиме записи, поясняющая принцип работы. В этой схеме на вход кодера 2RS подается мультиплексированный стереофонический сигнал в виде единого цифрового потока со скоростью  $V_{da} = 1,536$  Мб/с. Для осуществления кодирования аудиоданные организуются в блоки длительностью 135,4 мкс, включающие в себя по 26 символов. После корректирующего кодирования блок данных состоит из 32 символов и затем в него добавляется 4 символа служебной информации. В результате скорость цифрового потока  $V_d$  увеличивается до значения  $V_{d1} = 2,4576$  Мб/с, приводимого в технических характеристиках. В стандарте DAT очень большой объем различной дополнительной информации для потребителя. Для

этих данных особой защиты от ошибок не требуется, поэтому их цифровой поток формируется помимо кодера RS и его скорость равна  $V_{d2} = 1,3056$  Мбит/с.

В ОЗУ-1 формируется массив данных записываемых на магнитную ленту в одну строку: 128 блоков информации с выхода кодера RS и 68 блоков дополнительной информации. Общая скорость цифрового потока  $V_d$ , подаваемого с ОЗУ-1 на вход канального модулятора составляет 3,7632 Мбит. Тактовая частота канального модулятора 8/10 рассчитывается по формуле

$$V_{ch1} = \frac{V_d \cdot n}{m}$$

и она получается равной 4,704 МГц. Выходной сигнал с модулятора записывается в буферные ОЗУ-2 и ОЗУ-3 с тактовой частотой  $V_{ch1}$ , а считывается с них на головки записи поочередно с другой частотой  $V_{ch2}$ . Это связано с тем, что запись каждой строки на ленте происходит за время в пол оборота барабана, равное 7,5 мс, а эта информация накапливается в ОЗУ-1 за время всего оборота, 15 мс. Это значит, что запись должна производиться с удвоенной канальной скоростью  $V_{ch2} = 9,408$  МГц.

Корректирующее кодирование и канальная модуляция производятся, когда ни одна головка не находится в контакте с НЗ. Ввод данных в одно из ОЗУ производится в то

	7,5	7,5	7,5	7,5
	есть контакт ленты с головкой А	нет контакта ленты с головками А и В	есть контакт ленты с головкой В	нет контакта ленты с головками А и В
ОЗУ-1	запись		ввод с АЦП	кодирование
ОЗУ-2	ввод с АЦП	кодирование	запись	
ОЗУ-1	считывание	декодирование	вывод на ЦАП	
ОЗУ-2	вывод на ЦАП		считывание	декодирование

время, когда из другого ОЗУ осуществляется запись. Каждая из этих операция длится по 7,5 мс. В режиме воспроизведения к выходам головок подключаются усилители считывания, с которых данные вводятся поочередно в ОЗУ-2 и ОЗУ-3. В схему вводятся демодулятор и декодер. Воспроизведение цифровой записи производится поочередно головками А и В, демодуляция и декодирование осуществляются в промежутки времени по 7,5 мс, когда нет контакта ни с одной головкой. Порядок выполнения этих операций поясняется табл.9.6. В реальной схеме операции кодирования и декодирования осуществляются одним процессором по разным программам и функции всех ОЗУ выполняет одна память.

### Корректирующее кодирование.

Для коррекции ошибок в стандарте DAT используется код-произведение (Product Code) на основе кодов двух RS кодов в поле Галуа ( $2^8$ ), имеющий очень высокую исправляющую способность. Для кодирования аудиоданные организуются в сектора из 28 блоков по 26 символов (рис.9.13.), всего 728 символов. Кодирование осуществляется по строкам и столбцам сектора. В результате вместе с проверочными символами объем сектора увеличивается до 1024 символа, что согласуется с нормами, принятыми в вычислительной технике.

Внешний код RS(32,26), обозначаемый  $P$ , имеет 6 проверочных символов и при избыточности 18,7% позволяет исправлять до 3 ошибок или до 6 стираний. Внутренний код RS(32,28), обозначаемый  $Q$ , имеет 4 проверочных символа и при избыточности 12,5% может исправлять до 2 ошибок в блоке или 4 стирания.

Порождающие полиномы кодов:

$$G(x)_P = \prod_{i=0}^5 x - \alpha^i, \quad G(x)_Q = \prod_{i=0}^3 x - \alpha^i.$$

Проверочные матрицы кодов, по которым вычисляются полусиндромы, имеют вид:

$$H_p = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ \alpha^{31} & \alpha^{30} & \dots & \alpha^2 & \alpha & 0 \\ \alpha^{62} & \alpha^{60} & \dots & \alpha^4 & \alpha^2 & 0 \\ \alpha^{93} & \alpha^{90} & \dots & \alpha^6 & \alpha^3 & 0 \\ \alpha^{124} & \alpha^{120} & \dots & \alpha^8 & \alpha^4 & 0 \\ \alpha^{155} & \alpha^{150} & \dots & \alpha^{10} & \alpha^5 & 0 \end{bmatrix}, \quad H_q = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 & 1 \\ \alpha^{31} & \alpha^{30} & \dots & \alpha^2 & \alpha & 0 \\ \alpha^{62} & \alpha^{60} & \dots & \alpha^4 & \alpha^2 & 0 \\ \alpha^{93} & \alpha^{90} & \dots & \alpha^6 & \alpha^3 & 0 \\ \alpha^{124} & \alpha^{120} & \dots & \alpha^8 & \alpha^4 & 0 \\ \alpha^{155} & \alpha^{150} & \dots & \alpha^{10} & \alpha^5 & 0 \end{bmatrix}.$$

Как видно, здесь используется проверочная матрица  $H_q$  точно такая же, как в стандарте CD.

Если в стандарте CD допускается вероятность ошибок не более  $10^{-3}$ , то в стандарте DAT

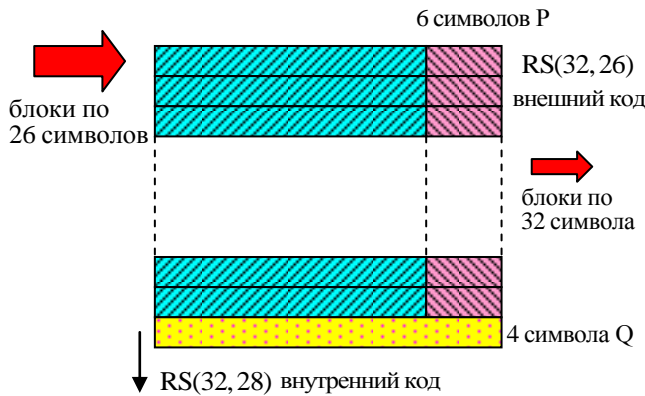


Рис.9.13. Корректирующий код 2RS-PC

корректирующее кодирование значительно эффективнее и рассчитано на исправление ошибочных символов с вероятностью до  $10^{-2}$ . При вероятности ошибочных символов  $10^{-3}$  вероятность ошибки на выходе не превышает  $10^{-15}$ . На рис.9.14. приведены графики потенциальной возможности кода в стандарте DAT по исправлению и обнаружению ошибочных символов в зависимости от их вероятности на входе декодера. При вероятности ошибок на входе  $P_S = 10^{-2}$  вероятность их не исправления равна  $10^{-10}$ . Для перехода к

частоте ошибок необходимо вероятность ошибок умножить на скорость цифрового потока аудиоданных, равную в этой системе 192 Кбайт/с. Тогда получим, что такой вероятности соответствуют 1-2 ошибки в день.

В стандарте DAT используются очень сложные алгоритмы перемежения. Оно осуществляется внутри блоков, по строкам, по столбцам, между секторами и между дорожками на носителе записи. Для осуществления такого перемежения звуковые данные, подлежащие записи, объединяются в матрицу из 4 секторов, представленных на (рис.9.13.). Матрица состоит из 28 строк и 128 столбцов объемом 131072 байт. Каждая строка этой матрицы соответствует данным, относящимся к звуковой информации, записываемым на одной дорожке (128 блоков). Вдоль строки осуществляется сверточное перемежение с интервалом 4 блока, при котором данные одного блока распределяются по 32. По столбцам осуществляется межблочное перемежение с интервалом 1 блок, при котором происходит перестановка всех нечетных символов в соседнюю строку. Это соответствует перемежению данных между соседними дорожками. Система перемежения обеспечивает полное исправление пакетов ошибок при длине дефектов до 2,64 мм и их маскировку при длине до 8,9 мм.

На рис.9.15. приведены алгоритмы декодирования внутреннего в внешнего кодов, реализующие возможности кода по исправлению одиночных случайных ошибок и выпадений. В первой ступени декодирования осуществляется обнаружение ошибочных символов и исправление одиночных и двойных ошибок в блоке. Если число ошибок в блоке превышает возможности декодирования, то все символы в блоке стираются и им придаются специальные флаги  $F1$ . Во 2 ступени декодирования, алгоритм работы

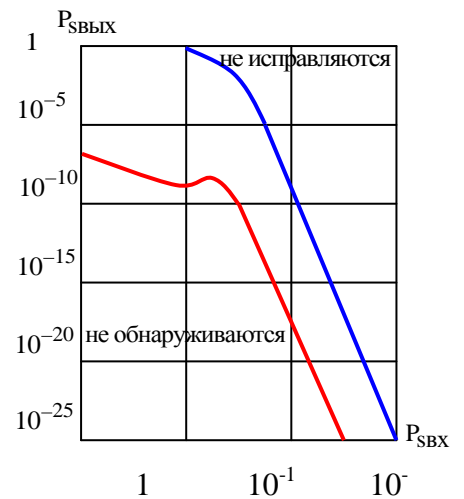


Рис.9.14. Графики вероятности случайных ошибок

значительно сложнее. В ней исправляются ошибки и стирания, введенные в 1 степени. Максимально может исправляться 2 ошибки и 6 стираний.

### Канальная модуляция

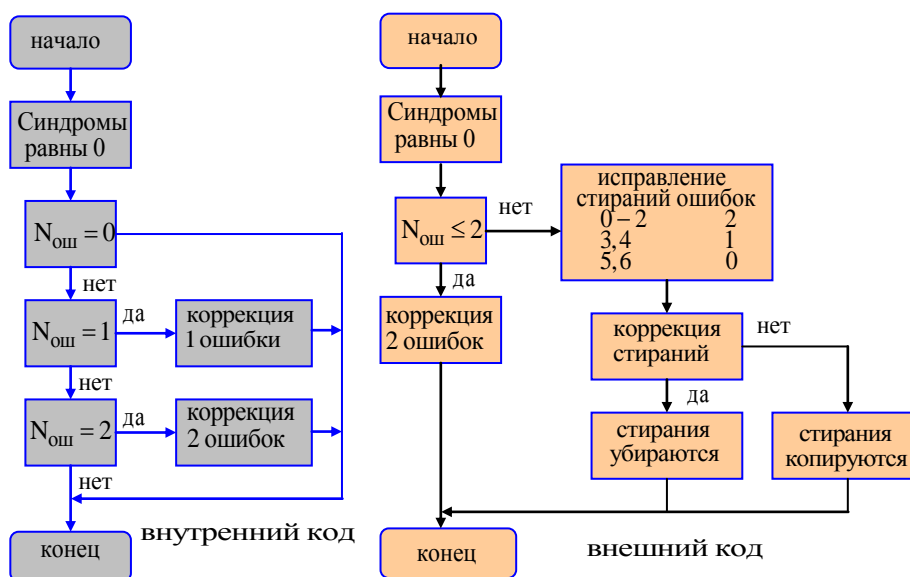


Рис.9.15. Алгоритмы декодирования кода 2RS

В магнитофонах стандарта DAT используется каналный код 8/10, в котором  $d = 0$  и  $k = 3$ . Он относится к группе RLL кодов и часто обозначается как (0,3,8,10). В отличие от кода EFM в коде 8/10 нет разделительных бит и полностью отсутствует постоянная составляющая. Кодирование заключается в преобразовании последовательности символов по 8 бит данных в символы по 10 канальных бит. Из  $2^{10}$  кодовых комбинаций выбираются 256 кодовых комбинаций, которые удовлетворяют параметрам  $d$  и  $k$  кода, определяющих допустимое минимальное и максимальное число последовательных 0 в таблице кодирования.

Для кода 8/10 число последовательных нулей в таблице может быть от 1 до 4. Кодовых комбинаций, удовлетворяющих параметрам  $d$  и  $k$ , которые не имеют постоянной составляющей, всего 153. Оставшиеся 103 комбинации приходится выбирать из тех, которые имеют ненулевое значение DSV. Поэтому, вместо одной 10-разрядной комбинации, каждому из 103 8-разрядных символов поставлены в соответствие две, отличающиеся друг от друга только знаком DSV. Одна имеет значение  $DSV = +2$ , другая –  $DSV = -2$ . Конфигурация их отличается только в 1 разряде, одна имеет на входе 1, другая - 0.

Таблица 9.7. Фрагмент таблицы канального кода 8/10						
$Q = -1$				$Q = 1$		
8 бит	10 бит	DSV	$Q'$	10 бит	DSV	$Q'$
00010011	0101110010	0	1	0101110010	0	-1
00010100	1101110001	2	1	0101110001	-2	-1
00010101	1101110011	2	-1	0101110011	-2	-1

По этим причинам канальная модуляция осуществляется с помощью 2 таблиц по 256 символов, их фрагменты приведены в табл.9.7. Условно эти таблицы названы  $Q = -1$  и  $Q = 1$ . В этих таблицах приведены как преобразование 8 разрядного числа в 10 канальных бит, так и указывается значение DSV каждой кодовой комбинации, которое

может быть 0, 2 или минус 2. В зависимости от текущего значения DSV следующая кодовая комбинация выбирается из одной таблицы или другой. Если, например, на входе модулятора повторяется одна комбинация 00010101, то на выходе модулятора будут чередоваться комбинации 1101110011 и 0101110011 из разных таблиц с DSV равным +2 и -2 и таким образом достигается поддержание DSV на нулевом уровне.



Рис.9.16. График спектральной плотности мощности кода 8/10

разделительные биты не используются, а минимизация постоянной составляющей осуществляется в процессе кодирования. Спектр кода 8/10 простирается от 0 до 10 МГц и спад в районе нулевых частот равен минус 20 дБ (рис.9.16.). Его основная мощность сосредоточена в узком диапазоне 2,35...4,7 МГц и к 10 МГц она уменьшается на 40 дБ.

Основные технические характеристики кода 8/10 приведены в табл.9.8. Как видно, он имеет очень низкий коэффициент плотности записи, равный 0,8. Это значит, что канальное кодирование плотность записи уменьшает. Тем не менее, линейная плотность записи в этом коде выше, чем на оптических дисках в стандарте CD. В коде 8/10 очень маленькая избыточность, поэтому он не может использоваться для обнаружения ошибок.

Декодирование канального кода осуществляется с помощью электронной таблицы или логической схемы. Таблица имеет 256 входов и 1 выход. Разрешенные кодовые комбинации являются адресом входа. Если в процессе воспроизведения произошла ошибка, которая разрешенную кодовую комбинацию преобразовала в запрещенную, то она обнаруживается с вероятностью 1. Если ошибка преобразовала одну разрешенную комбинацию в другую, но тоже разрешенную, то такая ошибка не обнаруживается. Помехозащищенность канального кода определяется окном детектирования. Пока величина джиттера не превысит половины окна, равного 53 нс, декодирование производится без ошибок.

Магнитофоны стандарта DAT могут быть студийными, репортерскими и бытовыми. При этом они существенно отличаются по характеристикам и техническим возможностям. Студийные магнитофоны фирмы Sony PCM R-300, R-500 и R-700 имеют 20-ти разрядное кодирование. Магнитофоны PCM 7000 и 7040 имеют 4 головки и поэтому обеспечивают воспроизведение во время записи. Фирма TASCAM также выпускает магнитофоны этой системы: модели DA-30, 40 и 45HR. Последняя модель имеет 24 разрядное кодирование и удвоенную скорость движения ленты (16,3 мм/с). Магнитофон модели DA-60MKII является репортерским.

В заключение настоящего раздела приведем такой печальный факт- цифровые кассеты долго не живут из-за очень высокой плотности записи. Принцип «да-нет» здесь предстает в невыгодном свете, поскольку относится и к работоспособности цифровой аппаратуры и к пригодности цифровых записей для работы, Статистика такова: около 10...30 лет хранятся самые лучшие экземпляры DAT кассет и около 5 лет - основная

таблиц с 256 входами и 1 выходом. Адресами таблицы являются 256 выбранных 10-разрядных кодовых комбинаций. Записываемый фрейм канального кода, включает в себя 320 канальных бит данных, относящихся к звуку и проверочным символам, 30 бит служебной информации и синхросигнал из 10 канальных бит. Длительность фрейма равна 135,4 мкс.

В отличие от кода EFM в модуляторе 8/10 нет проблемы стыковки канальных символов и

Т а б л и ц а 9.8. Канальный код 8/10	
Избыточность кода, %	20
Число таблиц кодирования	2
Скорость канального кода	8/10
Коеф. плотности записи	0,8
Коеф. самосинхронизации	4
Скорость канальных бит, Мб/с	9,408
Окно детектирования, нс	106,4
Дискретность интервалов, нс	106,4
Мин временной интервал, нс	106,4
Макс. временной интервал, нс	425,6
Макс. частота спектра, МГц	4,7
Мин. частота спектра, МГц	2,35
Линейная плотность, бит/мм	2388



масса. После этого срока катастрофические выпадения приводят запись в полную негодность. На этом пострадало большое количество известных вещательных фирм, которые перевели свои архивы на кассеты DAT. Заметим, что аналоговая запись на качественной магнитной ленте хранится около 50 лет. Потом начинает осыпаться магнитный слой, теряет прочность основа. Но запись держится «до последнего».

Есть еще один важный вопрос о многократной перезаписи цифровых данных на магнитной ленте. Даже при очень высоком качестве ленты из-за очень высокой скорости записи – воспроизведения часть чисел теряется. Для борьбы с этим используется избыточность корректирующих кодов и с помощью этих кодов при декодировании обнаруживаются ошибки. Иногда они полностью исправляются, иногда для их исправления используется интерполяция, при которой пропавший сигнал «дорисовывается».

Из этого следует важный вывод, что при цифровой перезаписи фонограмм всегда вмешивается система коррекции ошибок. При последующем воспроизведении система коррекции вмешивается еще раз и так далее. Звук будет меняться. Поэтому цифровыми перезаписями злоупотреблять не стоит. Если система коррекции не справляется со своими обязанностями, то появляются цифровые искажения, а они звучат на порядок хуже, чем аналоговые. Зачастую мы слушаем не запись, а творение системы коррекции.

#### 9.4. Многоканальные дисковые магнитофоны формата HDD

Основой дискового магнитофона HDD является магнитный накопитель на жестких дисках, который часто называют винчестером или просто жестким диском. Эти магнитофоны предназначены для многоканальной студийной цифровой звукозаписи с числом каналов от 4 до 24 и выше. Все каналы мультиплексируются в единый цифровой поток, и запись производится на обе стороны пластин винчестера. Для записи звука большей частью используются только винчестеры с пластинами со стандартным диаметром 3,5 дюйма, скорость их вращения может быть 5400, 7200 и 10000 об/мин, а емкость равна 40, 60 и 80 Гбайт соответственно (Гбайт – Гигабайт =  $10^9$  байт).

Из-за секторной организации данных полезная площадь этих пластин не превышает 56%, поэтому поверхностная плотность звукозаписи на них очень велика и составляет: 30, 46 и 60 Мбит/мм<sup>2</sup> соответственно. Это почти в 500 раз больше поверхностной плотности записи на компакт диске.

Столь высокая плотность записи достигается несколькими путями. Во-первых, основа пластины изготавливается из алюминия толщиной 1...2 мм с очень тщательной полировкой обеих поверхностей. На них вакуумным напылением сначала наносится буферное покрытие толщиной около 10 мкм на основе соединения никеля и фосфора (Ni-P), которое существенно уменьшает шероховатость поверхности. Далее наносится двухслойное мелкодисперсионное ферромагнитное покрытие на основе хрома (Cr), кобальта (Co) и платины (Pt) общей толщиной около 0,25 мкм. Защитное покрытие выполняется в виде алмазоподобной пленки толщиной несколько сотен ангстрем (рис.9.17.). Такой магнитный носитель обладает сверхнизким уровнем шума с вероятностью ошибочных бит при записи не более  $10^{-12}$ ... $10^{-14}$ . Сравним с тем, что вероятность ошибочных бит при изготовлении компакт дисков находится в пределах  $10^3$ ... $10^4$ .

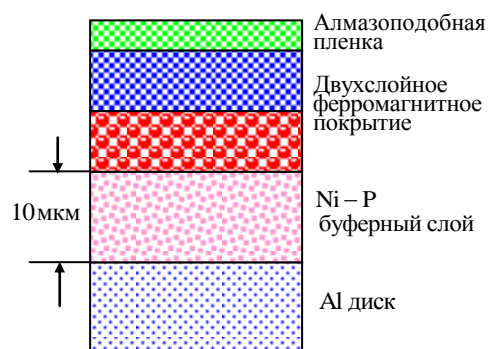


Рис.9.17. Покрытия жесткого диска

Запись на каждую сторону пластины производится магнитными головками с рабочим зазором шириной 0,1...0,5 мкм в виде *вертикальной намагниченности*. Скорость записи современных магнитных головок достигает 300 Мбит/с. Для обеспечения такой скорости записывающие головки изготавливаются на основе тонких пленок пермаллоя. Чтение



Рис.9.18. Индукционная записывающая головка и считывающая головка GMR

записи на жестких дисках осуществляется отдельными головками, которые конструктивно объединены с записывающими головками (рис.9.18). В рабочем состоянии головка парит над поверхностью вращающего жесткого диска, поддерживаемая на небольшом расстоянии аэродинамическими силами в потоке увлекаемого пластинами газа-наполнителя винчестера. В настоящее время зазор между головкой и поверхностью диска составляет всего 25... 10 нм.

В современной аппаратуре используются два метода считывания : на основе магниторезистивного эффекта и на основе спин-эффекта электронов. В головках первого типа под действие магнитного поля фонограммы изменяется электрическое сопротивление резистивного датчика. Эти изменения не превышают 1...2%. В 1997 году фирмой IBM была разработана считывающая головка с гигантским магниторезистивным эффектом, при котором под действием магнитного поля сопротивление датчика меняется до 100% и, следовательно, во много раз увеличивается выходной сигнал головки. Такие головки получили специальное название GMR (Giant Magnetoresistive). Благодаря им за 10 лет поверхностная плотность записи повысилась в 100 раз, достигнув 100 Гбит/квадратный дюйм, а скорость считывания возросла до 500 Мбит/с.

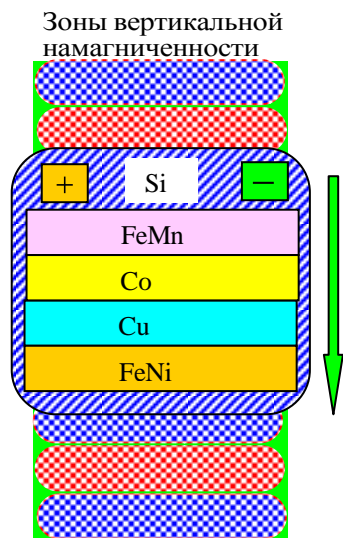


Рис.9.19. Структура GMR элемента

В датчиках этих головок имеет место квантово-механический эффект, наблюдаемый в тонких плёнках, состоящих из чередующихся ферромагнитных и немагнитных слоёв. Он проявляется в существенном уменьшении электрического сопротивления в присутствии внешнего магнитного поля. Этот эффект возникает вследствие квантовых свойств электронов, которые имеют такую характеристику, как "спин". Спин электронов может иметь только два направления - вверх и вниз. Электроны проводимости со спином, направление которого совпадает с направлением магнитного поля внутри GMR-среды, испытывают меньшее сопротивление при движении и имеют большую свободу перемещения, что проявляется в уменьшении электрического сопротивления. Электроны со спином, ориентированным против внутреннего магнитного поля испытывают большее сопротивление при движении и чаще сталкиваются с атомами среды, поэтому электрическое сопротивление возрастает.

На этом эффекте и построена работа GMR-головок.

Датчик GMR головки состоит из четырех тонкопленочных слоев (рис.9.19) . Первый активный слой, выполняющий функции считывающего элемента, представляет собой тонкую пленку на основе сплава никеля и железа (FeNi), которая легко намагничивается внешним полем перпендикулярной ориентации. Ширина его 0,3...0,5 мкм, а толщина - 10 нм. Таким магнитным датчиком теоретически можно считывать биты с линейной плотностью свыше  $10^{10}$  бит/мм.

Тонкая пленка кобальта выполняет функцию постоянного магнита, она создает постоянное магнитное поле вертикальной направленности по отношению к НЗ. Два

основных магнитных слоя FeNi –Co разделяются очень тонкой немагнитной медной прослойкой (Cu). Она препятствует переходу намагниченности слоя кобальта на активный слой. Толщина этого слоя около 0,9 нм (атом меди имеет размер 0,3 нм). Четвертый слой – антиферромагнитный на основе сплава железа с марганцем (FeMn). Он поддерживает намагниченность слоя кобальта постоянной ориентации.

Толщина четырехслойного датчика около 30 нм, он изготавливается методом фотолитографии на кремниевой подложке толщиной 1 мм и представляет собой сложный



Рис.9.20. GMR головка фирмы ALPS

электрический проводник тока, к концам которого подводится постоянное напряжение. Площадь, занимаемая такой конструкцией, равна примерно 1 мм. У моделей GMR головок фирмы Read-Rite ширина считывающего элемента равна 0,45 мкм, а его толщина – около 15 нм.

При считывании цифровая фонограмма вызывает намагниченность активного слоя датчика GMR головки. Если эта намагниченность совпадает с направлением намагниченности слоя кобальта (постоянного магнита), все электроны, спин которых направлен так же, как и эти магнитные поля, свободно передвигаются в обоих слоях и электрическое сопротивление малое. Если направления этих полей противоположны, то движение электронов со спином вверх будет сдерживаться одним из этих двух слоев, а со спином вниз - другим, вследствие чего общее сопротивление датчика возрастает. Таким образом, датчик головки работает по принципу сравнения векторов собственного магнитного поля и поля фонограммы.

Еще одним фактором, повышающим плотность записи, является использование разработанного фирмой IBM канального группового кода (2,7). В табл.9.9. приведены основные характеристики этого кода. Код имеет переменную длину кодовых слов  $n$ ,

Таблица 9.9.		Канальный код 2/7	
Характеристики кода		Таблица кодирования (значения m и n)	
Канальный код (d, k, m, n)	(2, 7)	Вход кодера	Выход кодера
Скорости канального кода,	0,5	10	0100
Число кодовых интервалов	6	11	1000
Коеф. самосинхронизации	2,67	000	000100
Коеф. плотности канального кода	1,5	010	100100
Избыточность канального кода, %	100	011	001000
Вероятность не обнаружения ошибок	0,25...0,0625	0010	00100100
Число таблиц кодирования	1	0011	00001000

которая меняется от 2 до 8 канальных бит и постоянную скорость  $V_{ch} = 0,5$ . Главное, у него большой коэффициент плотности  $K_d$ , равный 1,5. Код отличается простой схемой реализацией, так как он не имеет разделительных бит. Кодирование/декодирование производится с помощью простой таблицы из 2 столбцов и 7 строк (табл.9.9) или

логической матрицы. При этом код (2/7) обладает высокой обнаруживающей способностью кодовых ошибок.

Цифровая запись звука на жесткие диски осуществляется в виде концентрических дорожек с зоной записи по радиусу от 15 до 43,5 мм. Запись производится с примерно постоянной линейной плотностью и для этого частота вращения меняется от одной группы секторов к другой. При емкости одной стороны диска 40 Гбайт шаг дорожки равен 0,49 мкм, ширина дорожки 0,34 мкм и линейная плотность составляет 30000 бит/мм, что соответствует длине пита на диске всего 0,05 мкм. Воспроизведение этой записи осуществляется GMR элементом шириной 0,45 мкм, у которого ширина считывающего элемента равна 150 ангстрем. Реализации такой высокой линейной и поверхностной плотности записи в значительной мере способствует полная изоляция пластин от внешней пыльной среды и использование в производстве дисков высоких технологий.

В дисковых рекордерах для обнаружения и исправления ошибок используется код-произведение на основе двух кодов Рида Соломона с алгоритмом исправления ошибок 3 поколения. Этот код исправляет одиночные ошибки в последовательности из 121 бита и выпадения длиной до 185 бит. Скорость записи данных на диск составляет до 300 Мбит/с, а скорость считывания выше – до 525 Мбит/с.

Жесткие диски в рекордер для звукозаписи устанавливаются по выбору потребителя. В зависимости от его емкости можно изменять шаг дорожки в широких пределах. По желанию потребителя в рекордере может быть также установлено до 6-ти жестких внешних дисков и съемный диск. В качестве жестких дисков используются винчестеры моделей E-IDE или Ultra-DMA IDE типа Kingston. В некоторых моделях встраиваются оптические рекордеры системы CD, а вместо жесткого диска устанавливается магнитооптический диск. Возможна синхронная работа 8, 16 и 32 рекордеров, что позволяет одновременно иметь от 64 до 768 каналов для записи звуковых сигналов.

Время записи зависит от емкости жесткого диска, числа каналов и формата цифрового сигнала. При 24-канальной записи в формате 24/48 на диске емкостью 20 Гбайт длительность записи равна 100 мин. Время записи увеличивается обратно пропорционально числу каналов. Максимальная длительность одного файла принята равной 24 часам.

Почти во всех рекордерах предусмотрена возможность программного изменения состава оборудования и его конфигурации. Это позволяет избежать быстрого морального старения аппаратуры. Имеется также возможность выполнять монтажные функции такие как: автоматический поиск нужного фрагмента, вставка без монтажных переходов, выполнение многих монтажных операций (вырезать/копировать/вставить/стереть), а также временной сдвиг отдельных дорожек или всех сразу. При выполнении этих операций число виртуальных звуковых каналов может быть до 99. Длительность монтажного перехода не превышает 10 мс. Обычно имеется и внутренний микшер на все каналы. В рекордерах предусмотрена программная установка значения частоты дискретизации: 44,1; 48 и 96 кГц. В некоторых моделях имеется возможность задания частоты дискретизации равной 32 и 192 кГц.

Дисковые магнитные рекордеры обычно имеют 8 или 16 звуковых каналов. Их технические характеристики и возможности много лучше, чем у цифровых многоканальных многодорожечных магнитофонов, особенно это касается монтажных функций и в частности времени поиска. У рекордеров оно не превышает 10 мс, тогда как время ускоренной перемотки часовой магнитофонной кассеты составляет 2 мин. Эти рекордеры по существу являются специализированными компьютерами. По-видимому, в скором будущем они полностью заменят цифровые многодорожечные магнитофоны.

## 10. Магнитооптическая запись звука

### 10.1. Физические основы магнитооптической записи

Основой магнитооптической записи является использование эффекта сильной зависимости свойств магнитных материалов от температуры  $T$ . С ее повышением существенно изменяется форма предельной петли гистерезиса, а именно: уменьшаются

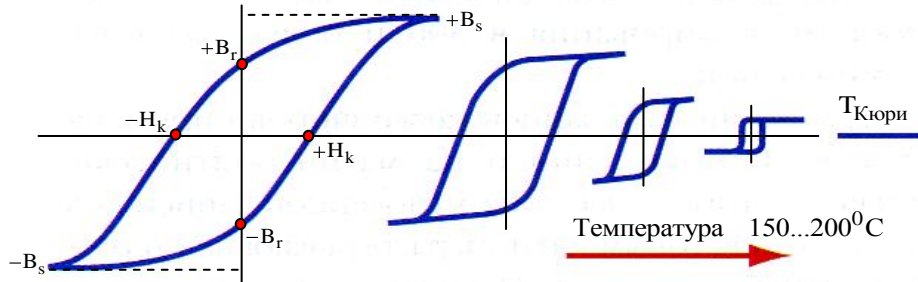


Рис. 10.1. Изменение петли гистерезиса с повышением температуры  
коэрцитивная сила и остаточная намагниченность, увеличивается крутизна наклона петли гистерезиса, магнитная проницаемость и отношение  $B_r / H_k$  (рис.10.1.). Все это приводит к тому, что петля гистерезиса становится прямоугольной и запись может осуществляться слабым магнитным полем. В пределе при некотором значении температуры  $T_k$ , называемой точкой Кюри, происходит переход вещества из ферромагнитной фазы в парамагнитную. При понижении температуры магнитные свойства восстанавливаются.

Этот термомагнитный эффект подчиняется закону Вейса-Кюри

$$\mu = C_k / (T_k - T)$$

где  $C_k$  - константа Кюри, зависящая от материала,  $T_k$  - температура (точка) Кюри,  $\mu$  - магнитная проницаемость материала, являющаяся коэффициентом пропорциональности между напряженностью магнитного поля  $H$  и намагниченностью  $M$ . При приближении к точке Кюри значение  $\mu$  стремится к бесконечности. Это соответствует моменту переходу вещества из одной фазы в другую. Практически при нагревании до температуры близкой к точке Кюри коэрцитивная сила магнитного слоя уменьшается от 4 до 10 раз.

На основе описанного эффекта может осуществляться многократная запись и стирание цифровой информации. Фактически это магнитный способ записи, так как цифровая информация регистрируется в форме намагниченности НЗ. Когда же для нагрева носителя используется сфокусированное лазерное излучение, она называется магнитооптической. Предел плотности такой записи определяется не размерами магнитных доменов, а диаметром сфокусированного пятна. Материалы, которые могут использоваться для термомагнитной записи, должны обладать достаточно высокой коэрцитивной силой, чтобы исключить возможность случайного размагничивания внешним магнитным полем. Важно, чтобы температура фазового перехода была не очень высокой, чтобы не требовалась большая мощность лазерного излучения. Наименьшей точка Кюри  $128^\circ\text{C}$  у диоксида хрома, у гамма-оксида железа эта температура равна  $650^\circ\text{C}$ .

Для считывания магнитооптической записи используется эффект Керра, который возникает когда участки фонограммы с различной намагниченностью меняют угол поляризации отраженного лазерного излучения. В результате отраженный световой поток модулируется по углу поляризации. Таким свойством обладают тонкие аморфные пленки на основе сплава с редкоземельными элементами, такие как  $\text{Y}_3\text{Fe}_5\text{O}_{12}$ ,  $\text{GdFe}_3\text{O}_{12}$ ,  $\text{CdTbFe}$

или TbFeCo. Запись на них может производиться только с вертикальной намагниченностью.

## 10.2. Стерефоническая звукозапись по стандарту MD

Существует множество вариантов реализации магнитооптической записи. Магнитное поле может быть постоянным, а модулируется лазерное излучение. Могут модулироваться одновременно магнитное поле и лазерное излучение. Для цифровой звукозаписи используется стандарт MD (Mini Disk), в соответствии с которым модулируется магнитное поле, а лазер осуществляет только нагрев магнитного слоя. Это позволяет производить стирание и запись одновременно.

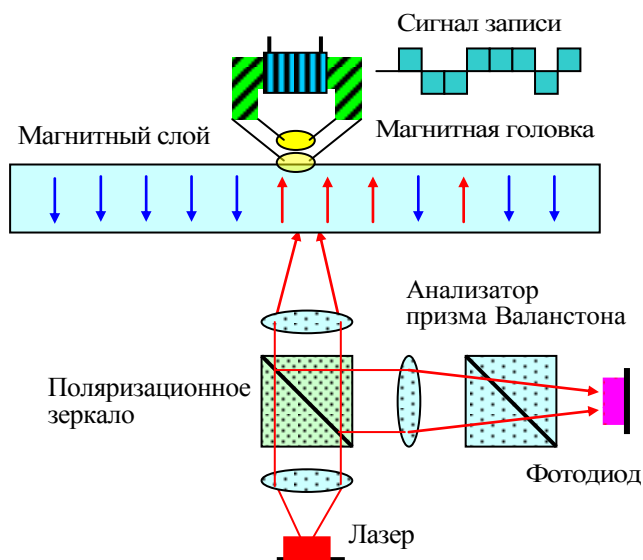


Рис.10.2. Магнито-оптическая запись и считывание

В этом стандарте запись стерео сигнала производится с постоянной линейной скоростью 1,2...1,4 м/с на прозрачный магнитооптический диск на основе поликарбоната диаметром 64 мм и толщиной 1,2 мм. Магнитная головка, расположенная с одной стороны диска, создает достаточно широкое поле намагничивания. Сфокусированное лазерное излучение направляется с другой стороны (рис.10.2.) и запись производится в точках, где температура

нагрева выше точки Кюри. Длина волны излучения лазера 780 нм, мощность при записи 2,5...5 мВт.

Запись осуществляется в канальном коде EFM, как и в стандарте CD, код коррекции несколько отличается. Регистрация цифровой информации производится в виде последовательности магнитных меток различной длины, как и при амплитудной оптической записи. Плотность такой записи определяется диаметром светового пятна.

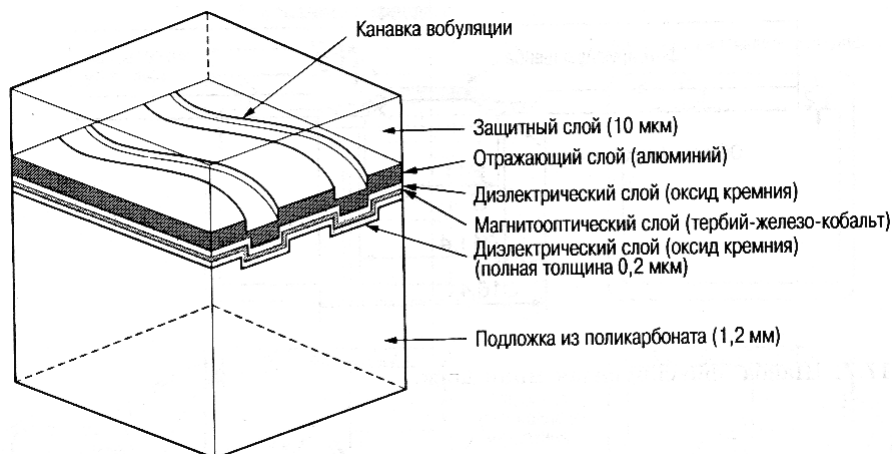


Рис.10.3. Структура магнитооптического MD

Под действием нагрева происходит и стирание старых данных.

Длительность записи на диске равна 74 мин, что достигается использованием очень эффективной системы сжатия цифрового потока ATRAC. Она уменьшает эту скорость почти в 5 раз с 1,4 МГц до 232 кГц, практически без ухудшения качества звучания.

Конструкция MD почти такая же, как у CD-RW дисков (рис.10.3.). Запись производится по дну спиральной канавки с вобуляцией, обеспечивающей радиальное слежение за дорожкой как при записи, так и считывании. Ширина канавки 1,1 мкм, глубина  $\lambda/8$ . В качестве магнитного покрытия используется сплав тербий-железо-кобальт с температурой Кюри  $185^{\circ}\text{C}$  и коэрцитивной силой всего 6,4 кА/м. Это позволяет осуществлять быстрый нагрев носителя и запись слабым магнитным полем. Магнитный слой с обеих сторон изолирован диэлектрическими покрытиями на основе оксида кремния, защищающими его от коррозии. Алюминиевое покрытие выполняет

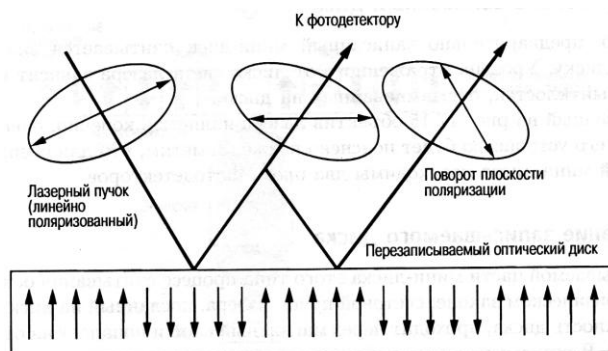


Рис.10.4. Принцип считывания магнитооптической записи (Эффект Керра)

отражающую роль по тепловой энергии, что позволяет снизить мощность лазера.

Воспроизведение магнитооптической записи осуществляется сфокусированным лазерным излучением с использованием эффекта Керра. Изменение угла поляризации отраженного светового пучка от участков с различной намагниченностью поясняется рис. 10.4..

Разделение падающего и отраженного световых пучков производится, как обычно, с помощью поляризационного светоделителя (рис.10.2). Отраженный свет направляется на анализатор поляризации, в качестве которого используется призма Валанстона. С помощью этой призмы модуляция светового пучка по углу поляризации преобразуется в амплитудную модуляцию тока фотоприемника. Угол Керра не превышает  $0,5^{\circ}$ , но при этом отношение сигнал/шум не менее 25 дБ, что вполне достаточно для надежного считывания записи.

### 10.3. Форматы магнитооптической записи

Магнитооптические диски широкого назначения выпускаются 2-х диаметров: 3,5" и 5,25". Они могут быть односторонними и двухсторонними и предназначены для работы в компьютерных накопителях и в специальных рекордерах. Запись звука на них осуществляется в канальном коде (2,7) или (1,7). Основные характеристики МО приведены в табл.10.1. В студиях радиовещания МО диски используются для монтажа фонограмм с многократной перезаписью без сжатия звука.

Фирма HBB выпускает МО диски специально для профессиональной звукозаписи. Ее модель 3,5" диска МО 640MB имеет очень высокое отношение сигнал-шум, низкий уровень BLER и время хранения 30 лет. Другая модель диска диаметром 5,25" МО 2,6 GB/5,2GB имеет еще большую емкость и время хранения до 100 лет. Гарантируется не менее 1 млн. циклов записи-воспроизведения. В студиях наиболее широко используются стереофонические МО рекордеры фирмы STUDER модели D424. Запись производится на двухсторонние МО диски емкостью 1,2 и 2,6 Гбайт с разрядностью от 16 до 24 бит и частотой дискретизации 48 кГц.

МО диски устойчивы к сильным внешним магнитным полям, они не боятся перепадов температуры, солнечного света и вибраций. Гарантированное время жизни около 40 лет. У дисков этого типа есть потенциальные возможности повышения

плотности записи. Уже сейчас разработана и внедряется технология магнитного суперразрешения MSR (Magnetic Super Resolution), при которой запись осуществляется на сложное многослойное покрытие. Благодаря этому покрытию при воспроизведении записи с высокой плотностью существенно уменьшается межсимвольная интерференция и поэтому значительно повышается надежность ее считывания без ошибок.

Другим направлением повышения емкости MO дисков является использование более коротковолновых лазеров и высокоразрешающих микрообъективов. Недавно, корпорация Sony объявила о выпуске MO-привода BW-F101, рассчитанного на работу с дисками PDD (Professional Disc for DATA). В приводе используется лазер с длиной волны излучения 405 нм, который фокусируется объективом с числовой апертурой 0,85 в световое пятно диаметром 467 нм. Емкость диска PDD составляет 23 Гбайта. Скорость чтения - 11 Мбит/с, записи - 9 Мбит/с. Это примерно соответствует параметрам оптических дисков форматов HD-DVD и BlueRay.

Т а б л и ц а 10.1. Параметры MO дисков		
Диаметры дисков, мм	130 и 90	
Толщина, мм	1,2	
Шаг дорожки, мкм	1,1...1,6	
Длина бит, мкм	0,31...0,52	
Скорость вращения, об/мин	1800...4500	
Коэф. преломления подложки	1,57	
Коэф. отражения, %	15...24	
Мощность лазера: при записи, мВт	8	
при считывания	1,5...2	
Число циклов перезаписи	$>10^7$	
BER: при записи	$< 2 \times 10^{-5}$	
после коррекции	$< 2 \times 10^{-12}$	
SNR, дБ	$> 47$	
Время жизни, лет	$> 40$	
Емкость MO дисков:		
Плотность	Диск 130 мм	Диск 90 мм
1x	560/650 Мбайт	128 Мбайт
2x	1,2/1,3 Гбайт	290 Мбайт
4x	2,3/2,6 Гбайт	
5x		540/640 Мбайт
5x		540/640 Мбайт
10x	4,6 Гбайт	1,3 Гбайт



# 11. Цифровые звуковые интерфейсы

## 11.1. Общие сведения

*Цифровой звуковой интерфейс* – это линия передачи звуковой информации в кодированном виде с использованием *цифровой канальной модуляции*, которая включает в себя передающую часть, кабельную линию передачи и приемную часть. В процессе модуляции последовательный двоичный код аудио данных преобразуется в двухуровневую *последовательность дискретных временных интервалов*. Такая модуляция часто называется *канальным кодированием*.

Звуковые интерфейсы используются для передачи цифровых звуковых сигналов от одного до другого объекта аудио техники, например, от пульта звукорежиссера к многоканальному магнитофону. Такая связь может осуществляться внутри студии или аппаратной, а также между студиями, находящимися на разных этажах здания или в отдельных зданиях. Длина таких линий связи может быть от 1 до 1000 м. Передача информации может осуществляться с помощью витой пары проводов, коаксиального телевизионного кабеля или волоконно-оптической линии.

Звуковые интерфейсы разделяются на две группы. К первой группе относятся интерфейсы, предназначенные для передачи цифровых аудиоданных (моно, стерео или многоканальный звук), возможно вместе с информацией о корректирующем кодировании и со служебной информацией. Ко второй группе относятся интерфейсы, предназначенные для передачи синхросигналов или сигналов управления.

Наибольший интерес представляют интерфейсы первой группы. В них цифровая информация всегда передается *в последовательном коде* с использованием канальной модуляции. Для передачи стерео и многоканального звука производится предварительное *мультиплексирование* звуковых каналов.

Схема такого интерфейса приведена на **рис.11.1**. В этой схеме в передающей части на вход кодера подается уже мультиплексированный звуковой сигнал. В кодере производится формирование двоичного цифрового потока в виде последовательности

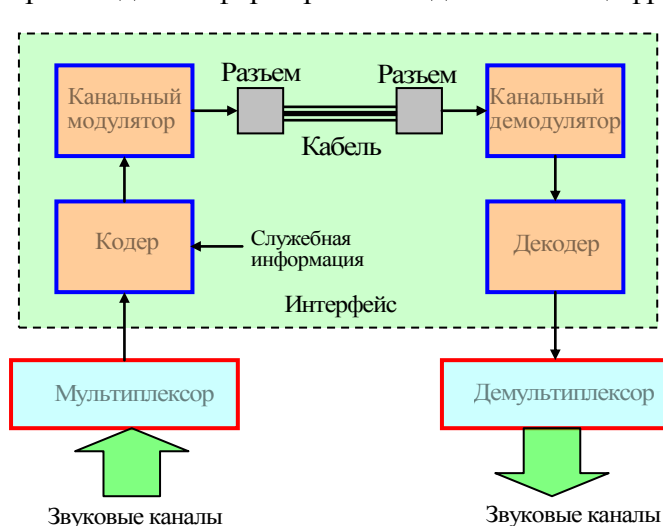


Рис.11.1. Цифровой аудио интерфейс

блоков данных, в которые включается служебная информация, полностью отражающая все характеристики звуковых каналов. Это позволяет на выходе демультимплексора приемной стороны полностью восстановить исходный цифровой поток.

При формировании блоков данных производится *корректирующее кодирование* звуковой и служебной информации, способное обнаруживать кодовые ошибки. Ошибки обнаруживаются в декодере приемной части и маскируются. Для обнаружения ошибок в блоках, содержащих звуковые данные, используется простейшая

проверка на четность, при которой, проверочный бит равен нулю, если число бит в блоке четное, и он равен единице, если число бит нечетное. Для обнаружения ошибок в служебной информации применяется циклический код CRC. В приемной части интерфейса обнаруженные ошибки маскируются.

При передаче информации по линиям связи сигналов прямоугольной формы из-за неизбежных в канале амплитудно-фазовых искажений возникает *джиттер* (фазовые

качания фронтов импульсов). При демодуляции этот джиттер приводит к возникновению кодовых ошибок, поэтому для уменьшения вероятности этих ошибок применяют частотно-фазовую коррекцию АЧХ линии передачи и специальные схемы приема данных, называемые деджиттер.

*Канальная модуляция* предназначена для преобразования спектра передаваемого цифрового потока с тем, чтобы согласовать его с АЧХ кабеля, и таким образом

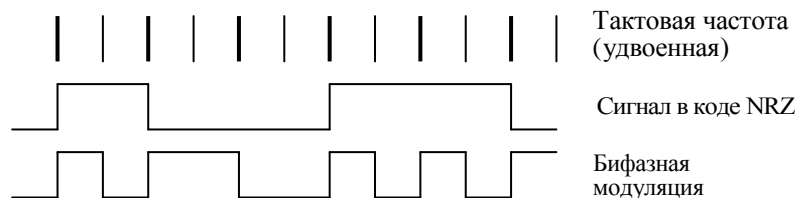


Рис.11.2. Бифазная модуляция

минимизировать частотные и фазовые искажения сигнала в линии связи. Выбор способа канальной модуляции в значительной мере зависит от типа кабеля и его длины. Канальная модуляция позволяет уменьшить джиттер передаваемых дискретных сигналов, снизить влияние низкочастотных наводок и повысить помехоустойчивость.

В звуковых интерфейсах используется только такая канальная модуляция, которая обладает свойством *самосинхронизации*. Под этим термином понимается, что в приемной части – тактовая частота выделяется из принимаемого канального кода, она необходима для демодуляции принимаемого сигнала. В цифровых звуковых интерфейсах используются два способа канальной модуляции, обладающие свойством самосинхронизации, – *бифазная модуляция* и модуляция на основе *группового канального кода 4/5 NRZI*. Все характеристики этих канальных кодов связаны с периодом их тактовой частоты  $T_{ch}$ .

При бифазной модуляции тактовая частота канального кода  $T_{ad}$  – должна вдвое превышать тактовую частоту двоичного сигнала на входе модулятора. Это значит, что требуется выполнение равенства  $T_{ch} = T_{ad} / 2$ , где  $T_{ad}$  – период тактовой частоты двоичного кода. *Бифазная модуляция* двоичного потока осуществляется по логическим «правилам» - логическому «0» в центре кодовой комбинации соответствует перепад от 1 к 0 (рис.11.2). Логическим «1» соответствуют перепады от 0 к 1 (изменение фазы на  $180^\circ$ ). В результате канальной модуляции формируется двухчастотный импульсный сигнал с кодовыми интервалами времени  $T_{max} = T_{ch}$  и  $T_{min} = T_{ch} / 2$ , которые определяют спектр канального кода (рис.11.3).

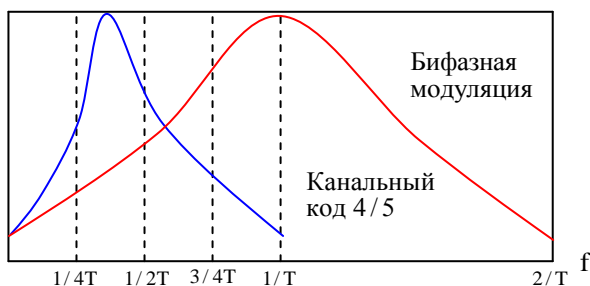


Рис.11.3. Спектральная плотность мощности

данных передается двумя канальными битами. Демодуляция канального кода осуществляется без кодовых ошибок, пока пиковые значения джиттера не превысят значений  $\pm 0,25T_{ad}$  ( $\pm 0,5T_{ch}$ ).

Дискретный спектр сигнала в канальном коде определяется равенствами

$$f_{max} = 1/2T_{min}, \quad f_{min} = 1/2T_{max}$$

В этом спектре отсутствует постоянная составляющая, максимум спектральной плотности мощности на частоте  $1/T_{ad}$  (рис. 11.3.), и спектр простирается до частоты  $2/T_{ad}$ . Бифазная модуляция применяется в интерфейсах AES3 (Audio Engineering Society), AES3-id и S/PDIF.

*Канальный код 4/5 NRZI* имеет 20% избыточность и поэтому он более эффективен, чем бифазная модуляция. Период тактовой частоты кода определяется равенством

$T_{ch} = 4T_{ad} / 5$ . При модуляции исходный поток бит данных разделяется на группы по 4 бита, и, в соответствии с таблицей кодирования, эти группы преобразуются в кодовые комбинации (символы) из 5 бит. В такой кодовой последовательности число последовательных «нулей» не превышает двух, а число последовательных «единиц» не превышает восьми. Код трехчастотный, интервалы времени кодовых комбинаций могут принимать значения  $T_{ch}, 2T_{ch}, 3T_{ch}$ . В этом канальном коде для передачи одного бита данных требуется 1,6 канальных бит. Демодуляция осуществляется без кодовых ошибок, пока пиковые значения джиттера не превысят значения  $\pm 0,4T_{ad}$ . Это значит, что помехозащищенность канального кода 4/5 в 1,6 раза выше помехозащищенности бифазной модуляции. В спектре модулированного сигнала отсутствует постоянная составляющая и максимум спектральной плотности мощности находится на частоте  $0,4/T_{ad}$  (рис.11.3.), поэтому частотные искажения в области высоких частот не играют большой роли. Канальный код 4/5 применяется в интерфейсе AES10 (MADI).

## 11.2. Интерфейсы AES3, AES3id и SPDIF

Цифровые звуковые цифровые интерфейсы AES3, AES3id и SPDIF, предназначенные для передачи данных цифровых стереофонических каналов, имеют одинаковую структуру организации цифрового потока. Основой этой организации является последовательность выборок в двоичном дополнительном коде, с числом разрядов до 24

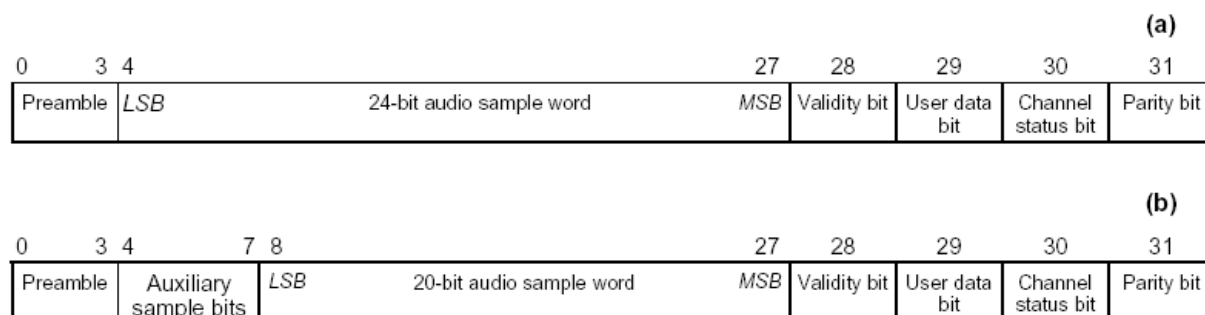


Рис.11.4. Формат субфрейма с выборкой одного звукового канала длиной 24 и 20 бит

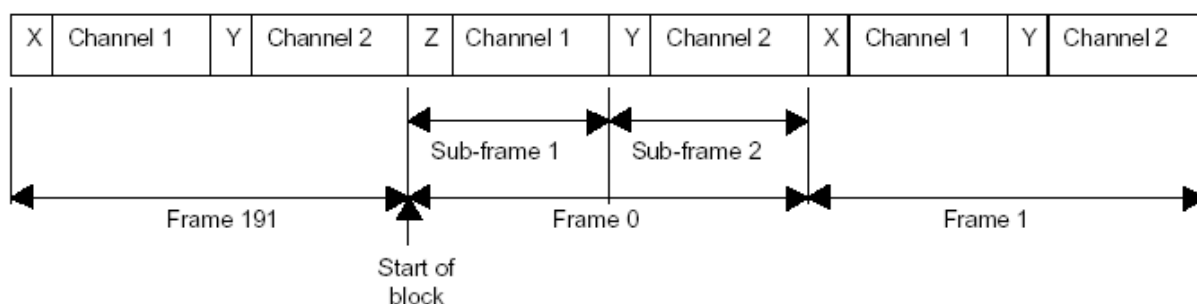


Рис.11.5. Формат фрейма из 192 субфреймов

. Выборки двух звуковых каналов передаются в виде отдельных блоков, называемых суб-фреймами. В эти блоки включается вспомогательная информация, синхросигналы и проверочные биты кода обнаружения ошибок (рис.11.4.).

Цифровые данные стереофонического звукового канала попеременно передаются в виде суб-фреймов с выборками левого и правого каналов, обозначенные на рис.11.5. как Channel 1 и Channel 2. Эти блоки данных отличаются синхросигналами (X, Y и Z) с тем, чтобы их можно было опознавать. Они, соответственно, обозначаются как Sub-frame 1 и Sub-frame 2. Фрейм данных состоит из 192 суб-фреймов, в нем содержится совокупная служебная информация объемом 192 бита, передаваемая отдельными битами в суб-фреймах.

Если по интерфейсу передается одноканальный (моно) звуковой сигнал, то его выборки размещаются в суб-фреймах 1. В суб-фреймах 2 эта информация может дублироваться, или передаваться логические «0». Если по интерфейсу передается 5...7 канальный звуковой сигнал (например, 5.1), то основной канал передается в субфреймах 1, а остальные – в субфреймах 2.

Каждый суб-фрейм включает в себя 32 слота (0...31), определяющих местоположения бит данных.

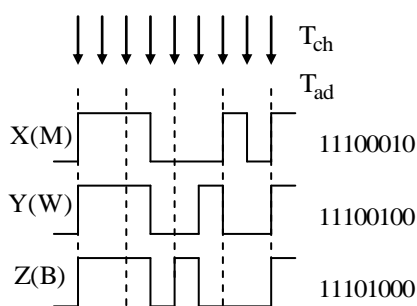


Рис.11.6. Синхрогруппы бифазного кода

**Слоты 0...3 (preamble)** предназначены для размещения сигналов синхронизации, которые занимают это место после осуществления канальной модуляции. При передаче данных используются три синхрогруппы, обозначаемые в стандартах AES и IEC958 по разному, – X(M), Y(W) и Z(B) (рис.11.6). Первая синхрогруппа обозначает начало суб-фрейма 1, вторая – начало суб-фрейма 2 и третья – начало фрейма. Для того чтобы синхрогруппу можно было легко опознать, в её конфигурацию включают не встречающиеся в канальном коде интервалы времени между изменениями уровня. Синхрогруппы распознаются и в инверсной форме.

**Слоты 4...7 (auxiliary sample bits)** предназначены для вспомогательной информации, такой как звуковой канал для переговоров между продюсерами или разговорной связи между студиями. Они могут использоваться для увеличения числа передаваемых бит в выборках до 24 и передачи не стандартных данных.

**Слоты 8...27 (20-bit audio sample word)** предназначены для 20 бит выборок левого или правого звуковых каналов. Выборка начинается с младшего разряда (LSB) заканчивается старшим разрядом (MSB). Если число разрядов меньше 20, то в неиспользуемые слоты записываются логические «0»

**Слот 28 (validity)** предназначен для размещения информации о достоверности выборки. Достоверной является выборка, не являющаяся результатом коррекции или маскирования (интерполяции, приглушения и т.д.). В этом случае бит в этом слоте равен нулю. В противном случае он равен 1.

**Слот 29 (user)** предназначен для размещения пользовательской информации о текущем времени, названии песни, все, что относится к этой информации. В рамках каждого фрейма объем такой информации равен 192 бита. В случае CD формата, в слоте 29 передаются данные каналов Q-W субкода.

**Слот 30 (channel)** предназначен для размещения данных, характеризующих передаваемую цифровую информацию – число звуковых каналов, частота дискретизации, число разрядов в выборках, наличие преэмпфазиса и другие. В совокупности эти данные называются *статусом канала*.

**Слот 31 (parity)** предназначен для размещения проверочного бита кода обнаружения ошибок в выборках. Если ошибок в выборке данного суб-фрейма нет, бит в этом слоте равен нулю.

*Фрейм статуса канала* из 192 бит разделяется на 24 байта. Свое назначение имеют как отдельные байты, так и отдельные биты в байтах. Эта информация разделяется на отдельные группы, такие как :

- управление,
- код категории информации,
- код источника звука,
- число звуковых каналов,
- частота дискретизации,
- число разрядов,
- тактовая частота и др..

Подробное описание назначения каждого байта и бита приводится в стандартах на интерфейсы.

### 11.3. Интерфейсы AES3 и AES/EBU

**Цифровые звуковые интерфейсы по стандартам AES3 и AES/EBU.** Это один и тот же интерфейс более известный под названием AES3, который предназначен для работы с профессиональной аудио техникой. Сначала стандарт на этот интерфейс был принят только AES (Audio Engineering Society), а затем он стал европейским стандартом (EBU -European Broadcasting Union) и изменил свое название, позже стал международным стандартом.

Стандарт AES3 много раз изменялся и дополнялся. Полное название последней версии: AES3-2003. AES standard for digital audio — Digital input-output interfacing — Serial transmission format for two-channel linearly represented digital audio data.

В соответствии с этим стандартом интерфейс AES3 предназначен для последовательной передачи только цифровых аудио данных *стереофонических звуковых каналов* в формате *линейной импульсно-кодовой модуляции*. Число разрядов в выборке может быть от 16 до 24 бит, могут использоваться стандартные частоты дискретизации: 48, 44,1 и 32 кГц. Тактовая частота передачи аудио данных  $f_{ad}$  однозначно определяется частотой дискретизации  $f_s$  и она не зависит от числа разрядов в выборках

$$f_{ad} = 2 \cdot f_s \cdot 32$$

При частоте дискретизации 48 кГц  $f_{ad} = 3,072$  МГц. Стандартом допускается использование не стандартных частот дискретизации: 22,05; 24,0; 88,2; 96,0; 176,4 и 192 кГц.

В интерфейсе AES3 для передачи цифровых данных по каналу связи используется *бифазная модуляция*, при которой тактовая частота канального кода  $f_{ch}$  должна быть вдвое выше частоты  $f_{ad}$  ( $f_{ch} = 2f_{ad}$ ). При  $f_s = 48$  кГц  $f_{ch} = 6,144$  МГц ( $T_{ch} = 0,163$  мкс), а спектр канального кода содержит две основные частоты - 6,144 и 3,072 МГц и занимает полосу от 100 кГц до 6 МГц.

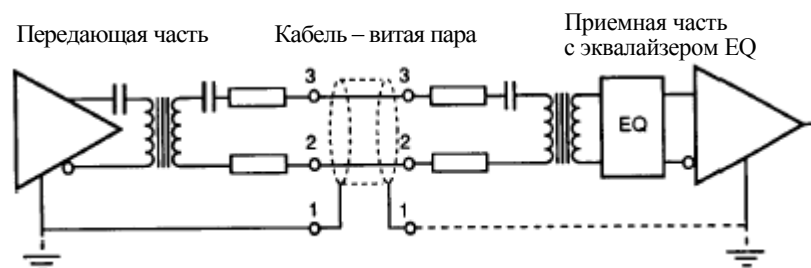


Рис.11.7. Схема интерфейса AES3

По стандарту AES3 передача цифровых данных производится с помощью *витой пары медных проводов в металлической оплетке* (кабель STP) длиной до 100 м без применения частотной коррекции (рис.11.7). При увеличении длины линии до 150 м с помощью эквалайзера производится подъем АЧХ приемника на частоте 10 МГц на 12 дБ. На входе и выходе интерфейса по стандарту должны использоваться трех контактные разъемы XLR-3, у нас они известны под названием СГ-3.

Волновое сопротивление такой линии должно быть равно  $110 \text{ Ом} \pm 10\%$  на частотах от 100 кГц до частоты в 128 раз выше частоты дискретизации (6,144 МГц). В соответствии со стандартом на выходе и входе интерфейса должны использоваться импульсные трансформаторы с симметричным выходом/входом, обеспечивающие гальваническую

развязку сигналов. При использовании трансформатора по двум проводам линии сигналы передаются в противофазе, а шумы и помехи наводятся в фазе.

На приемном конце интерфейса с помощью операционного усилителя сигналы суммируются в фазе, а наводки и шумы в значительной степени при этом подавляются. Выходной импеданс трансформаторного выхода усилителя мощности передатчика и входной импеданс трансформаторного входа приемника интерфейса также должны быть равны 110 Ом в диапазоне от 100 кГц до 6 МГц. Максимальный размах выходного сигнала

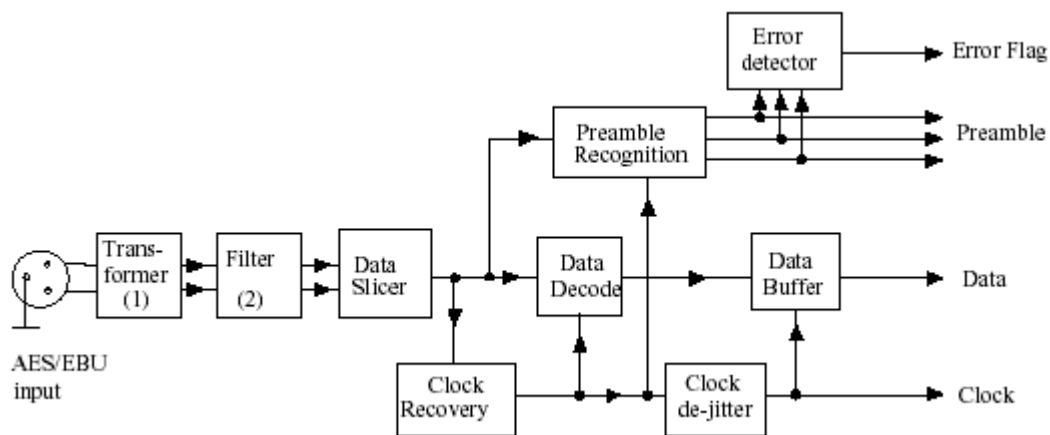


Рис.11.8. Блок-схема приемной части интерфейса AES3

2...7 В, минимальный размах входного сигнала 0,2 В. Фронты импульсов могут быть в пределах от 5 до 30 нс с джиттером не более 4 нс. Функциональная схема приемной части интерфейса приведена на рис.11.8.

*Интерфейс AES3* отличается от других интерфейсов прежде всего особым статусом канала служебной информации. Входящие во фрейм байты статуса канала имеют следующие назначения:

байт 0 – содержит информацию о частоте дискретизации, эмфазисе, о содержании (музыка или данные), о возможности копирования и о коде источника (CD, DAT...)

байт 1 – содержит информацию о формате звукового канала (моно, стерео или многоканальный звук),

байт 2 – содержит информация о числе разрядов в выборках,

байт 3 – используется только при многоканальной передаче,

байт 4 – содержит информация о возможности использования частоты дискретизации в качестве опорной,

байт 5 – резервный,

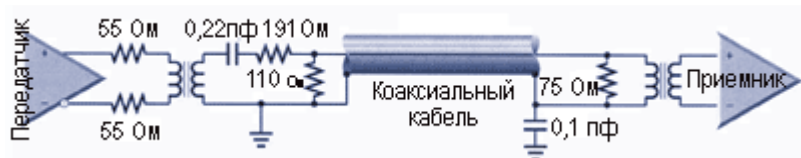


Рис.11.9. Схема интерфейса AES3id с согласующими трансформаторами

байты 6-9 и 10-13 – содержат коды каналов и коды назначения,

байты 14-17 : - содержат 32- разрядные адреса выборок,

байты 18-21 : содержат 32-разрядный тайм-код,

байт 22 – содержит информация о надежности блока данных,

байт 23 – содержит информацию о корректирующем коде CRC, в частности, его порождающий полином -  $G(X) = X^8 + X^4 + X^3 + X^2 + 1$

*Цифровой звуковой интерфейс по стандарту AES3id* также предназначен для работы с профессиональной аудио техникой. Он отличается от интерфейса AES3 только тем, что для увеличения длины линии передачи до 1000 м в нем цифровая информация передается по телевизионному коаксиальному кабелю с волновым сопротивлением в диапазоне частот от 100 кГц до 6 МГц  $75 \text{ Ом} \pm 10\%$ . На входе и выходе используется двух контактный разъем BNC. Из-за применения такого кабеля вход и выход интерфейса построены по несимметричной схеме (рис.11.9). Он допускает передачу данных по

Табл.11.1.	AES3	AES3id	S/PDIF
Interface	Balanced	Unbalanced	Unbalanced
Connector	XLR-3	BNC	RCA
Impedance	110 ohms	75 ohms	75 ohms
Output Level	2-7 Vp-p	1.0 Vp-p	0.5 Vp-p
Max Output	7 Vp-p	1.2 Vp-p	0.6 Vp-p
Max Current	64 mA	1.6 mA	8 mA
Min Input	0.2 V	0.32 V	0.2 V
Cable	STP4	Coax	Coax
Max Distance	100 m	1000 m	10 m

телевизионному коаксиальному кабелю на расстояние до 1 км, размах выходного сигнала 1...1,2 В, минимальный размах входного сигнала  $-0,32 \text{ В}$ . В табл.11.1 приведены сравнительные характеристики передаваемых и принимаемых сигналов трех интерфейсов. Из них видно, затухание сигнала в интерфейсе AES3id с длиной кабеля 1000 м

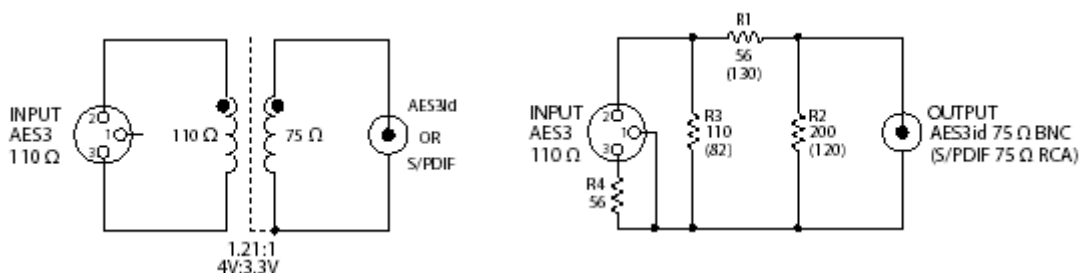


Рис.11.10. Схемы согласования интерфейсов AES3 и AES3id

значительно меньше, чем затухание сигнала в интерфейсе AES3 с длиной кабеля 100 м. Стандартом рекомендуются переходники, позволяющие подключать кабель интерфейса по стандарту AES3id к входу интерфейса по стандарту AES3 и наоборот (рис.11.10).

Вместо коаксиального кабеля этим стандартом рекомендуется использовать волоконно-оптический кабель. Такой интерфейс включает в себя передающий модуль, приемный модуль и оптический кабель. Передача осуществляется с помощью арсенид-

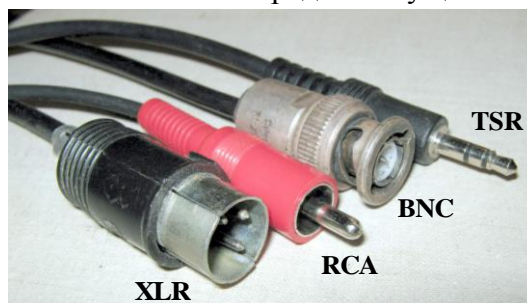


Рис.11.11. Разъемы звуковых интерфейсов

галлиевого лазерного диода с длиной волны излучения  $0,66 \text{ мкм}$ , прием производится кремниевым фотодиодом.

Большинство ведущих фирм используют собственные (нестандартные) цифровые многоканальные интерфейсы, особенно для записи многоканального звука.

#### 11.4. Интерфейсы SDIF-2, SDIF-3 и S/PDIF

Профессиональный **цифровой звуковой интерфейс фирмы Sony SDIF-2** (Sony Digital InterFace) содержит три коаксиальных кабеля с разъемами BNC с волновым сопротивлением 75 Ом. По ним раздельно передаются два звуковых канала и сигнал синхронизации. Передача осуществляется боками по 32 бита с использованием канального кода NRZ, скорость цифрового потока равна 1,53 Мбит/с при частоте дискретизации 48 кГц.

Профессиональный **цифровой интерфейс фирмы Sony SDIF-3** предназначен для передачи цифровых звуковых сигналов формата DSD (Direct Stream Digital), используемого в системе Super Audio CD, каждый канал передается по отдельному коаксиальному 75-омному кабелю с использованием бифазной канальной модуляции. Профессиональный **цифровой интерфейс фирмы Sony MAC –DSD** (Multi-Channel Audio Connection) предназначены для передачи 24 каналов DSD по витой паре проводов со скоростью 100 Мбит/с.

Многоканальный цифровой оптический интерфейс фирмы **Alesis** часто называемый **ADI** или **Lightpipe**, предназначен для передачи 8 звуковых каналов в канальном коде NRZI. Передача осуществляется блоками по 192 бита (24 бита, 8 каналов) с тактовой частотой 12,288 МГц. Длина кабеля небольшая от 33 до 100 футов.

**Цифровой звуковой интерфейс по стандарту S/PDIF** (Sony/Philips Digital Interface Format) предназначен для работы с бытовой аудио техникой, однако его основные технические характеристики, такие как организация цифрового потока и канальная модуляция, точно такие же как в профессиональном интерфейсе **AES3**. Отличия касаются частоты дискретизации, типа используемого кабеля и разъемов и, главное, передаваемой информацией статуса канала.

В соответствии со стандартом **S/PDIF** частота дискретизации передаваемых сигналов может иметь три значения:

44,1 кГц – при работе с CD плеерами,

48 кГц – при работе с бытовыми цифровыми магнитофонами DAT,

32 кГц – при работе с бытовыми цифровыми радиоприемниками DSR.

Передача цифровых аудио данных должна осуществляться по коаксиальному кабелю длиной от 1 до 10 м с волновым сопротивлением 75 Ом с использованием разъема RCA. Предусмотрена возможность использования оптического кабеля. Передача и прием сигналов, передаваемых по кабелю, осуществляются по несимметричной схеме, такой же, как в профессиональном интерфейсе AES3id. Максимальный размах выходного сигнала 0,5...0,6 В, минимальный размах входного сигнала 0,2 В. Сравнительные характеристики электрических сигналов различных интерфейсов приведены в табл.1.

По стандарту **SPDIF** из 24 байт статуса канала используются только первые 5 байт и корректирующее кодирование не производится. В этих байтах передается следующая информация

**Байт 0 .**

- профессиональное/бытовое оборудование,
- звук/данные,
- можно/нельзя копировать,
- есть/нет преэмфазис,
- число звуковых каналов,
- режим- только 00 (во стандарту CD)..

**Байт 1.**



- код категории оборудования (CD, DAT,...),

- код назначения.

#### Байт 2.

- номер источника и номер канала.

#### Байт 3

- частота дискретизации,

- точность тактовой синхронизации.

#### Байт 4.

- максимальное число разрядов (20/24),

- используемое число разрядов.

Различие данных в статусе канала делают интерфейсы AES3 и SPDIF несовместимыми.

### 11.5. Интерфейс AES10

*Цифровой звуковой интерфейс по стандарту AES10* (MADI – Multichannel Digital Audio Interface). Этот профессиональный интерфейс предназначен для последовательной асинхронной передачи многоканального цифрового звука на основе линейной импульсно-кодовой модуляции. *Синхронизация* осуществляется с помощью *отдельного интерфейса*. Наиболее часто интерфейс AES10 используется при многоканальной записи звука, когда магнитофон и пульт звукооператора объединены в единый комплекс (рис.11.12).

*Число передаваемых каналов зависит от частоты дискретизации :*

- 32 кГц – 48 кГц ( $\pm 12\%$ ) – 56 каналов,
- 32 кГц – 48 кГц (точно) – 64 канала,
- 64 кГц – 96 кГц ( $\pm 12\%$ ) – 28 каналов,

число разрядов в каждом канале может быть до 24. Число передаваемых стерео каналов в два раза меньше. Организация цифрового потока в интерфейсе AES10 такая же, как в



Рис.11.12. Интерфейсы AES10 (MADI), AES10 и AES24

интерфейсе AES3 (рис.11.4 и 11.5), только *за один период дискретизации передаются выборки не двух, а всех каналов*.

Применение внешней синхронизации исключает необходимость использования синхрогрупп в канальном коде. Вместо них в слотах 0-3 субфреймов передается информация, касающаяся именно интерфейса **AES10**:

- Слот 0 – синхросигнал «1», указывает начало передачи нулевого звукового канала,
- Слот 1 - указывается состояние канала (актив/не актив),
- Слот 3 - указывается моно/стерео в суб-фреймах А и В,
- Слот 4 - синхросигнал «1» указывает начало фрейма.

Для передачи информации по каналу связи в интерфейсе **AES10** используется канальный код 4/5. *Тактовая частота* этого кода равна 125 МГц, она *не зависит от числа передаваемых каналов и частоты дискретизации*. При такой частоте возможна передача мультиплексированного сигнала со скоростью цифрового потока до 100 Мбит/с. При частоте дискретизации 48 кГц и 64 моно каналах скорость цифрового потока равна 73,7 Мбит/с. Для обеспечения постоянной скорости цифрового потока в линии связи 125 Мбит/с добавляются каналы с нулевой информацией. Данные этих каналов могут произвольно вводиться в любое место фрейма.

В интерфейсе **AES3** передача возможна по коаксиальному кабелю с разъемом BNC и волновым сопротивлением 75 Ом длиной не более 50 м. Волоконно-оптический кабель позволяет осуществлять передачу на расстояние до 2 км.

При использовании интерфейсов AES10 внешняя синхронизация осуществляется с помощью специального **интерфейса по стандарту AES11**. От этого интерфейса часто осуществляется синхронизация работы всего звукового комплекса аппаратуры студии. Для дистанционного управления таким комплексом используется специальный **интерфейс по стандарту AES24**. С его помощью подаются команды управления и осуществляется мониторинг.

### 11.6. Аудио система персонального компьютера

Обязательными атрибутами мультимедийного компьютера являются привод CD-ROM или DVD и звуковая карта (рис.11.13). Все современные звуковые карты по их назначению можно разделить на три группы:



Рис.11.13. Аудио система компьютера

- звуковые, позволяющие только воспроизводить звук или записывать и воспроизводить,
- музыкальные, предназначенные только для синтеза электронной музыки,
- комбинированные, позволяющие записывать и воспроизводить звук, а также синтезировать электронную музыку.

В зависимости от технических характеристик звуковые карты первых двух групп могут использоваться для очень широкого применения (воспроизведения CD/DVD дисков, компьютерные игры, интернет, телефон...), только для профессиональной записи или для профессионального синтеза электронной музыки. По своим характеристикам и возможностям особо выделяются звуковые карты, предназначенные для использования в студиях цифровой звукозаписи.

Большинство звуковых карт имеют интерфейс подключения к компьютеру через шину PCI (Peripheral Component Interconnect). Профессиональные звуковые карты в виде внешних блоков подключаются через интерфейсы USB 1.1 или USB 2.0 (Universal Serial Bus). Звуковые карты PCI имеют 32-разрядную шину для обмена данными, но процедуры цифровой обработки звука и приема/передачи результатов обработки могут быть с разрядностью 64 и более. PCI аудио карты за счет намного более высокой скорости работы шины всегда поддерживают полный дуплекс.

Все функции звуковой карты выполняют, входящие в нее программные модули: кодек, входной и выходной аналоговые микшеры, цифровой микшер и сигнальный процессор, который осуществляет все операции по заданной программе (рис.11.13).

Кодек выполняет функции аналого-цифрового и цифро-аналогового преобразований. Число АЦП и ЦАП в этом кодеке определяется числом входных и выходных звуковых

каналов конкретной карты. В большинстве звуковых картах широкого назначения эти конвертеры 24-разрядные. Они строятся на основе сигма-дельта модуляции, при которой в АЦП и ЦАП используются 5..10- разрядные квантователи с частотно-зависимой отрицательной обратной связью. При этом обеспечивается отношение сигнал/шум не более 100...106 дБ, что эквивалентно всего 17-18 разрядам. Применение 24-разрядной организации цифрового потока связано только с тем, что кодовые комбинации передаются в виде байтов. Все входные и выходные цифровые потоки аудио данных организованы в формате линейной импульсно-кодовой модуляции. Все современные кодеки обеспечивают работу АЦП и ЦАП в режиме дуплекса, при котором запись и воспроизведение могут осуществляться от разных источников одновременно.

Частота дискретизации входного и выходного цифровых сигналов может быть от 8 до 192 кГц, она должна быть кратной стандартным значениям 44,1 и 48 кГц. В большинстве современных звуковых карт она равна 96 кГц, выбор высокой частоты дискретизации связан прежде всего с тем, что при этом проще реализуется сигма-дельта модуляция. При разработке программного обеспечения кодека (аудио кодек AC'97) были рекомендованы частоты дискретизации: 8,000, 11,025, 16,000, 22,050, 32,000, 44,100 и 48,000 кГц. Основной частотой принята частота 48 кГц, а все остальные сигналы приводятся к ней по

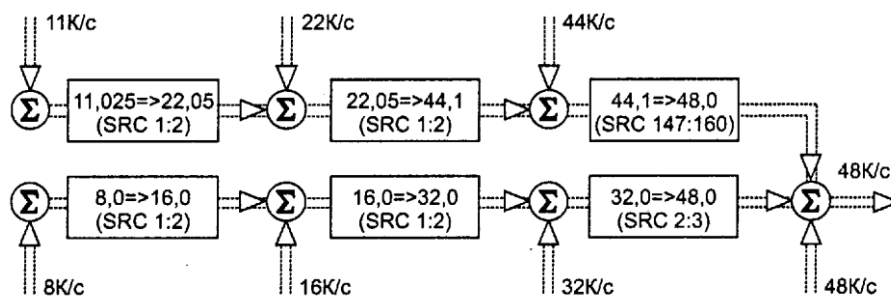


Рис.11.14. Схема преобразования частот дискретизации

схеме, приведенной на рис.11.14. Запись цифрового звука на жесткий диск или в оперативную память производится с частотой дискретизации всегда 48 кГц в виде файлов с расширением WAV (сокращенно от Wave — волна). Для перевода выходных файлов в другие форматы, например, MP3 требуется отдельное программное обеспечение.

Для иллюстрации технических возможностей в табл.11.2 приведены сравнительные характеристики профессиональных звуковых карт, прошедших тестовые испытания.

Студийную звуковую карту Pro Tools звуковой картой можно только условно, так как она является сложным программно-аппаратным комплексом для записи, редактирования, обработки и сведения многоканального звука и может представлять собой совокупность

Параметры		Audiophile	WT2496	DSP24Value	Echo MIA
Неравномерность (от 40 Гц до 15 кГц), дБ:	АЧХ	-0.09, +0.00	-0.01, +0.06	-3.44, +0.00	-0.05, +0.03
Уровень шума, дБ (А):		-99.6	-98.9	-86.2	-106.8
Динамический диапазон, дБ (А):		97.4	96.4	83.7	104.6
Нелин. искажения, %:		0.003	0.003	0.002	0.004
Интермод. искажения, %:		0.004	0.004	0.017	0.003
Взаимопроникновение каналов, дБ:		-91.1	-96.6	-86.6	-107.7

десятка PCI-карт и еще большего количества внешних приборов. Также довольно условна цена: в зависимости от сложности системы она может быть от 20 до 200 тыс. долларов.

Основа Pro Tools — это PCI-карта MIX Core Card, которая несет на борту шесть высокопроизводительных сигнальных процессоров от Motorola (используются DSP 56301 Onyx) и разъем для подключения внешнего интерфейса. Интерфейсы бывают разные (такие как Arogee, Prism Sound и Mytek), самым распространенным является интерфейс

888/24. Расшифровывается этот набор цифр следующим образом: восемь аналоговых входов, восемь аналоговых выходов, восемь цифровых входов/выходов, и все это работает с разрядностью 24 бита. С помощью одного или нескольких таких интерфейсов и осуществляется ввод/вывод звука в компьютер.

К аналоговым входам звуковых карт широкого назначения подключаются внешние аудиоустройства: микрофоны, магнитофоны, электрогитары и т. д. (рис.11.16). Число входных каналов может быть от 2 до 12. Для их подключения обычно используются разъемы StereoJack, но в профессиональных звуковых картах могут применяться разъемы XLR и другие. Line in и Mic In — это линейный и микрофонный входы. Они обычно выполнены на разъемах типа «мини-джек» (рис.11.17). Отдельный вход Mic In предусмотрен из-за того, что у микрофонов сигнал имеет низкий уровень и его нужно усилить, перед тем как направлять на преобразователь. Во многих профессиональных картах микрофонного входа нет, так как усилитель интегрируется с микрофоном.



Рис.11.15. Профессиональная звуковая карта EMU-0404

На некоторых типах звуковых плат установлен дополнительный вход Aux In. Сигнал с этого входа минует кодек звуковой платы и поступает на выходной аналоговый микшер, а оттуда — сразу на выход. Этот вход позволяет упростить коммутацию внешних устройств и использовать внутренний микшер звуковой платы для смешивания сигналов с внешнего и внутренних источников. Вход Aux In обычно делается на разъеме типа «мини-джек». Вход проигрывателя компакт-дисков CD In, как правило, расположен прямо на звуковой

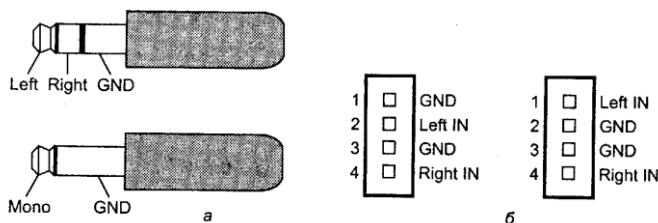


Рис.11.17. Разъемы для подключения аудио сигналов: а – мини джек, б – внутренний разъем подключения CD-ROM

плате.

Аналоговые входной и выходной микшеры с программным управлением обеспечивает регулировку уровня входных и выходных сигналов. Микшер позволяет смешивать

входные сигналы от нескольких источников (микрофона, CD, внешнего входа и синтезатора). В звуковой карте каждый входной звуковой канал имеет отдельные регуляторы уровня. Для монофонических источников (например, микрофон) кроме регулятора уровня имеется регулятор панорамы, позволяющий балансировать уровни сигналов, посылаемых от данного источника в левый и правый стереоканалы. Звуковая карта обычно имеет регулировку тембра по низким и высоким частотам или даже эквалайзер — многополосный регулятор тембра.

Синтезатор звуковой карты обеспечивает имитацию звучания музыкальных инструментов и воспроизведение различных звуков. Из множества методов синтеза в звуковых картах в основном используют два — частотный и волновой.

FM Music Synthesizer — синтезатор с частотной модуляцией (Frequency Modulated), обеспечивающий невысокое качество синтеза.

WT Music Synthesizer — синтезаторы с табличным синтезом (Wave Table). Большинство WT-синтезаторов содержат встроенные ПЗУ со стандартным набором инструментов General MIDI (GM), это 128 мелодических и 37 ударных инструментов. Некоторые синтезаторы имеют ОЗУ для загрузки дополнительных звуков.

Звуковой тракт универсальных карт рассчитан прежде на возможность воспроизведения DVD дисков с многоканальным звуковым сопровождением. Для воспроизведения записи по системе Digital Dolby число выходных каналов должно быть 6 в формате 5.1. Для воспроизведения цифровой звукозаписи по системе DTS (Digital Theatre System) число каналов увеличивается до 8 в формате 7.1. В соответствии с этими форматами определяется число звуковых колонок и число выходных разъемов. Встроенный усилитель имеет мощность до 4 Вт на канал, но на некоторых звуковых картах выходная мощность, достаточна только для наушников.

Звуковые колонки для PC несколько отличаются от обычных бытовых акустических систем. Они, как правило, малогабаритные, поскольку предназначены для установки на столе, по бокам от монитора. Малые габариты, конечно же, отражаются на качестве и выходной мощности. Активные колонки имеют встроенный усилитель, требующий внешнего питания. Они могут иметь регуляторы громкости и тембра. Пассивные колонки встроенного усилителя не имеют, их мощность невелика. Для высококачественного воспроизведения лучше использовать внешний усилитель с собственными акустическими системами или стереонаушники.

Наушники или усилитель можно подключать и к аудио выходу CD-ROM, что позволит прослушивать аудио-CD независимо от наличия звуковой карты. Регулятор уровня этого выхода расположен на лицевой панели CD-ROM. Там же в ряде моделей приводов имеются кнопки воспроизведения и выбора трека, позволяющие управлять проигрыванием без привлечения каких-либо программных средств.

Недавно появились звуковые колонки и микрофоны с цифровым интерфейсом, они к звуковым картам подключаются через интерфейс SPDIF. Доставка аудиоданных к ним и от них может осуществляться программно через контроллер шины USB.

Для подключения электромusикальных инструментов звуковые карты имеют порт MIDI (Musical Instrument Device Interface). Устройством ввода могут служить специальные MIDI-клавиатуры (как на клавишных музыкальных инструментах). Устройством вывода может быть синтезатор звуковой карты или внешний синтезатор, подключаемый к порту MIDI. Компьютер в такой системе используется как мощное средство создания, редактирования и хранения музыкальных произведений. MIDI-интерфейс имеют многие профессиональные и полупрофессиональные клавишные синтезаторы.

## 11.7 Интерфейс MIDI

Цифровой интерфейс музыкальных инструментов *MIDI* (Musical Instrument Digital Interface) является последовательным асинхронным интерфейсом с частотой передачи 31,25

Кбит/с. Этот интерфейс является стандартом для сопряжения компьютеров, синтезаторов, записывающих и воспроизводящих устройств, микшеров, устройств специальных эффектов и другой электромusыкальной техники. В настоящее время интерфейс MIDI имеют и дорогие синтезаторы, и дешевые музыкальные клавиатуры, пригодные в качестве устройств ввода компьютера. Сообщения MIDI широко используются и для передачи музыкальных записей (на дисках и по сети), а также как выходной аудио интерфейс для игр и прочих «звучащих» приложений. Описание музыкальных фрагментов в формате MIDI очень компактно: минута MIDI (файлы с расширением .MID) может занимать менее 10 Кбайт, в то время как минута оцифрованного звука (файлы с расширением .WAV) с качеством аудио-CD занимает около 10 Мбайт. Однако формат MIDI позволяет воспроизводить лишь звуки, на которые способен синтезатор на исполняющей стороне, в то время как цифровое аудио воспроизводит любые звуки. В отличие от оцифрованного звука при воспроизведении данных MIDI пользователь может независимо изменять тональность (транспонировать музыкальный текст) и темп исполнения, причем без искажения тембра и характера звучания инструментов. Современные процессоры и методы обработки сигналов позволяют транспонировать оцифрованный звук даже в реальном времени, но все-таки с рядом ограничений.

В *физическом интерфейсе* применяется *токовая петля 5 мА* (возможно, до 10 мА) с гальванической (оптронной) развязкой входной цепи. Логическому нулю соответствует наличие тока, логической единице (и покою) — отсутствие тока (в «классической» токовой петле телекоммуникаций — наоборот).

Интерфейс определяет три типа портов: *MIDI-In*, *MIDI-Out* и *MIDI-Thru* (рис. 11.18, 5.19).

*Входной порт MIDI-In* представляет собой вход интерфейса «токовая петля», гальванически развязанного от приемника оптроном с быстродействием не хуже 2 мкс. Устройство отслеживает информационный поток на этом входе и реагирует на адресованные ему команды и данные.

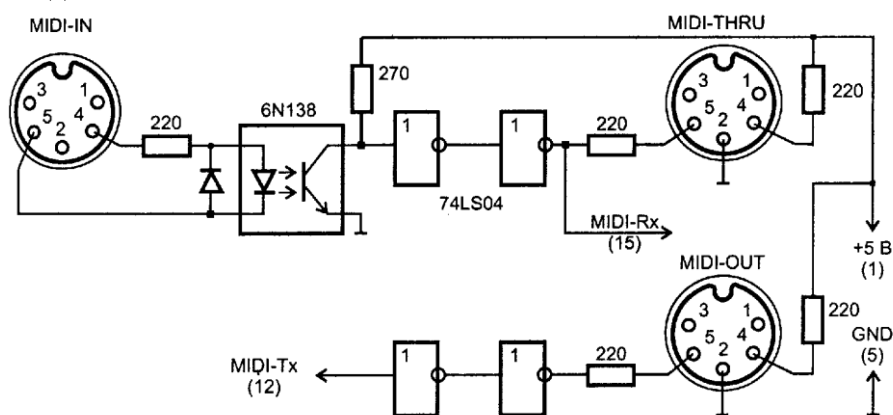


Рис.11.18. Схема кабеля-адаптера MIDI

*Выходной порт MIDI-Out* представляет собой выход источника тока, гальванически связанного со схемой устройства. Ограничительные резисторы предохраняют выходные цепи от повреждения при замыкании на землю или источник 5 В. На выход подается информационный поток от данного устройства. При специальной настройке устройства в этом потоке может содержаться и транслированный входной поток, но это нетипично.

*Транзитный порт MIDI-Thru* служит только для ретрансляции входного потока, по электрическим свойствам он аналогичен выходному. Его наличие не является обязательным для всех устройств.

В качестве разъемов применяются 5-контактные разъемы DIN, распространенные в бытовой звуковой аппаратуре.

Поток данных MIDI исходит от MIDI-контроллера или от секвенсора. *MIDI-контроллер* представляет собой устройство, на котором играют как на музыкальном инструменте (чаще всего это клавиатура). MIDI-контроллер формирует поток сообщений, отражающий события

исполнения (нажатие и отпускание клавиш) в реальном времени. *MIDI-секвенсор* (sequencer) — устройство, позволяющее перехватывать, хранить и редактировать (включая комбинирование, наложение и генерацию) поток MIDI-сообщений и воспроизводить его в заданном темпе. Исходящий поток MIDI-контроллера или секвенсора поступает на его разъем *MIDI-Out*. Для использования в качестве секвенсора компьютер должен иметь порт MIDI и соответствующее прикладное ПО, а также иметь производительность, достаточную для работы в реальном времени.

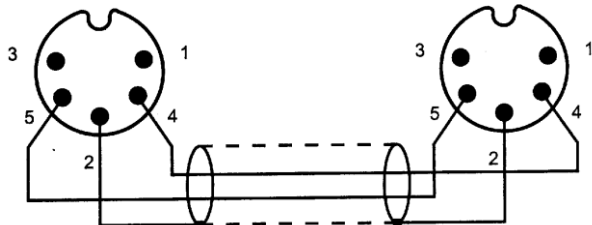


Рис. 11.19. Соединительный кабель MIDI

Конечным приемником потока MIDI-сообщений является *звуковой модуль* (sound module) — синтезатор, формирующий аудиосигнал на основе входящих к нему команд. Применительно к PC синтезатор может входить в состав звуковой карты, а может быть и внешним. Каждому звуковому модулю назначается свой номер

*логического канала* (MIDI Channel) в диапазоне 1-16.

Наличие порта *MIDI-Thru* позволяет соединять устройства в цепочки и более сложные структуры. Возможные варианты топологии должны подчиняться главному правилу: вход *MIDI-In* одного устройства должен подключаться к выходу *MIDI-Out* или *MIDI-Thru* другого устройства. При планировании MIDI-сети необходимо руководствоваться знаниями информационных потоков и связей устройств. Управляющие устройства — клавиатуры, секвенсоры (в режиме воспроизведения), источники синхронизации — должны находиться, естественно, перед управляемыми. Возможно применение и специальных устройств-мультиплекторов, позволяющих логически коммутировать множество входных потоков в один выходной. Сообщения MIDI разделяются на каналные (Channel Messages) и системные (System Messages).

*Канальные сообщения* подразделяются на *голосовые* и *управляющие* (Channel Mode Messages). *Голосовые сообщения* несут основную исполнительную информацию. Исполнение ноты начинается по приему сообщения *Note On*, которое содержит номер ноты и скорость нажатия клавиши. Это сообщение инициирует фазы атаки, спада и удержания. Скорость нажатия задает уровень (громкость) исполнения, качественные синтезаторы в зависимости от скорости могут корректировать и характер исполнения (спектр, форму огибающей, выбирать нужный образец). Нота снимается (отрабатывается фаза затухания) по приему с ее же номером сообщения *Note Off* (или *Note On* с нулевой скоростью). Голосовые сообщения управляют и характером исполнения (громкость, вибрато, смещение строя и т. п.). Каждому параметру синтезатора соответствует *номер контроллера* (в данном контексте — номер управляемого параметра), с помощью которого можно им управлять. Эти контроллеры не следует путать с MIDI-контроллерами — устройствами, с помощью которых исполнитель генерирует MIDI-сообщения. *Управляющие сообщения* определяют, как именно звуковой модуль должен обрабатывать голосовые сообщения.

*Системные сообщения* служат для синхронизации с видеооборудованием и ленточными записывающими устройствами, выбора произведения и позиции воспроизведения, настройки аналоговых синтезаторов. Сообщения реального времени образуют систему синхронизации *MIDI Sync*, используемую секвенсорами, ритм-машинами и другими тактируемыми устройствами MIDI. Есть и специальные сообщения для обмена различными данными, например для загрузки патчей.

Для обеспечения совместимости различных звуковых модулей был принят стандарт *General MIDI* (GM). Он определяет минимальные требования к модулям и оставляет место для расширений, среди которых широко распространены стандарты GS и XG. Стандарт GM поддерживают все звуковые карты, «понимающие» MIDI. Расширение GS поддерживается довольно большим числом производителей; XG поддерживают лишь синтезаторы (и

звуковые карты) фирм Yamaha и Korg. Для унификации методов хранения и передачи информации MIDI были приняты и стандарты на форматы файлов.

*Хранение и передача данных MIDI* имеют свою специфику. Когда источником сообщений MIDI является MIDI-контроллер и генерируемый им поток исполняется сразу в реальном времени, никакой маркировки сообщений обычно не требуется. Для хранения данных MIDI в виде файлов требуются специальные меры для отметки времени.

Секвенсоры оперируют *треками* — строками нот, исполняемыми параллельно. Трек примерно соответствует партии (или ее части) одного инструмента. Такое представление естественно для композиторов и обеспечивает богатые возможности по редактированию. Программный секвенсор может сохранять партитуры как в своем собственном (ни с кем не совместимым), так и в стандартизованном формате.

При сопряжении устройств MIDI с «инородными» устройствами (видеосистемы, аудиорекордеры) возникает задача их синхронизации. В MIDI используется система MIDI Sync с однобайтными сообщениями SRT. Кроме этой системы в музыкальных инструментах применяется и аппаратная синхронизация Sync24, известная под названием «DIN Sync», однако, несмотря на однотипность разъемов, никакой совместимости интерфейсов MIDI и Sync24 нет. Выделить из потока MIDI сообщения SRT может несложный микроконтроллер (или же аппаратная логика).

В мультимедийных приложениях возникает необходимость синхронизации MIDI-звука с движущимся видеоизображением. Существует несколько систем синхронизации SMPTE, принятых организацией Society of Motion Picture and Television Engineers (Общество инженеров движущихся изображений и телевидения).

Для передачи данных счетчиков SMPTE по интерфейсу MIDI разработана система синхронизации *MTC {MIDI Time Code}*. Она является мостом, соединяющим систему синхронизации MIDI, построенную на отметках долей тактов (beat), с данными в форматах SMPTE. Отметим, что синхронизация MIDI Sync является темпозависимой — частота посылки синхросообщений пропорциональна темпу исполнения, который может меняться. Система SMPTE, а следовательно, и MTC привязана к абсолютному времени — ее частота сообщений определяется только частотой кадров.

Для *синтезаторов с аналоговым управлением* интерфейс MIDI непосредственно не подходит. Для их подключения существуют специальные микроконтроллеры с ЦАП, формирующие управляющее напряжение (линейное или экспоненциальное) и разрешающий сигнал на основе потока сообщений MIDI, адресованных к выбранному каналу. Встроенное программное обеспечение такого контроллера может реализовать все команды, связанные с изменением частоты тона: вибрато, глиссандо, портаменто и т. п.

Наиболее распространенные MIDI-контроллеры представляют собой 4-6-октавные клавиатуры с полноразмерными (как у рояля) или уменьшенными клавишами. Каждая клавиша имеет датчик нажатия: простейший дискретный (кнопочный), фиксирующий лишь факт нажатия-отпускания, или же динамический, измеряющий скорость (силу удара) и передающий это значение в качестве параметра сообщения. Простейшие клавиатуры имеют подпружиненные клавиши, более качественные «взвешенные» клавиатуры для исполнителя по ощущениям близки к настоящим рояльным. Клавиатура может иметь датчик давления на клавишу после удара, «колеса» (wheels) для управления строем канала, а также подачи сообщений управления контроллерами (например, глубины модуляции). К клавиатуре могут подключаться педаль удержания нот (sustain), органы включения различных эффектов, переключения номера канала. На клавиатуре может быть установлен и небольшой дисплей или отдельные индикаторы режима работы.

Интерфейс MIDI в значительной степени ориентирован на клавишные музыкальные инструменты. Однако даже самая хорошая «взвешенная» клавиатура с датчиками скорости и давления не позволяет передать всех нюансов исполнения, например, духовых или струнных (щипковых или смычковых) инструментов. В настоящее время существуют MIDI-контроллеры с интерфейсами, отличными от клавишных. Существуют MIDI-



контроллеры в виде гитар, где датчики определяют место прижима струн и силу удара (щипка). Есть и MIDI-контроллеры с «духовым интерфейсом» — датчики устанавливаются на клапанах, а специальный мундштук измеряет расход проходящего воздуха. Существуют также и преобразователи аналогового сигнала в поток MIDI-инструкций. Они имеют АЦП, оцифровывающий входной сигнал, снятый микрофоном с реального инструмента, и внутренний сигнальный процессор. Для гитар выпускаются специальные звукосниматели, сопряженные с сигнальными процессорами. Задача разложения звука на голосовые (не спектральные!) составляющие достаточно сложна и легче решается для одноголосных инструментов. Если преобразователь использовать как устройство ввода для секвенсора, то ошибки преобразования можно исправлять средствами редактирования музыкального текста в секвенсоре.

Самым примитивным MIDI-контроллером может являться обыкновенная компьютерная клавиатура — она способна генерировать сообщения по фактам нажатия и отпускания клавиш. Ценным свойством клавиатуры является способность реагирования на одновременные нажатия-отпускания клавиш, в том числе и во время удержания нескольких клавиш нажатыми. Конечно, она не воспринимает динамику удара и расположение клавиш отнюдь не фортепьянное, но при отсутствии настоящей клавиатуры поиграть можно (даже аккордами). Возможность ввода MIDI-команд с клавиатуры имеется во многих программах, работающих с MIDI-синтезатором.

Поток сообщений MIDI проигрывается в компьютере с помощью синтезаторов, аппаратных или программных. Чисто аппаратная реализация MIDI выполняется на звуковых картах с собственным процессором. Этот процессор интерпретирует каждое сообщение MIDI в команды управления синтезатором (FM или WT). Простые карты имеют только сами синтезаторы, а управление ими для интерпретации сообщений MIDI осуществляет программный драйвер карты, исполняемый центральным процессором. Современные процессоры способны поддерживать чисто программную интерпретацию MIDI (Soft MIDI), выполняя и собственно синтез звука в цифровой форме. Правда, при этом занимают ресурсы (процессорное время и память).

Направление потоков MIDI задается программно (в среде Windows через Панель управления, значок Multimedia, вкладка MIDI). В простейшем варианте все сообщения MIDI можно посылать на одно из присутствующих в системе устройств, которые включают аппаратные (FM, WT и иные) синтезаторы установленных звуковых карт, их порты, к которым подключены дочерние карты, внешние порты MIDI, а также программные синтезаторы. Более сложную конфигурацию можно задать, указав для каждого из 16 каналов MIDI свой синтезатор (порт назначения).

Внешний порт MIDI (с сигналами TTL) обычно выводится на разъем джойстика, установленный на звуковой карте (или системной плате с интегрированным звуком), и для связи с внешними устройствами (с разъемами DIN и токовым интерфейсом) используются специальные переходники. Программно порт обычно совместим с UART MPU-401. На некоторых системных платах применяются БИС контроллеров интерфейсов, в которых UART, используемый для COM-порта, конфигурированием через CMOS Setup может быть переведен в режим MIDI-порта.

В последнее время в аудиотехнике намечается тенденция к использованию шины USB, которая пригодна как для обмена данными в традиционном виде, так и для изохронной передачи (с равномерной скоростью поступления данных) аудиоданных в цифровом виде. Интерфейс MIDI имеет ограничение на число каналов (16), и его невысокая скорость передачи ограничивает полифонические возможности (большое число нот в аккорде не может звучать строго одновременно). В то же время производительности современных PC хватает на то, чтобы справляться и с более мощными потоками данных. Для подключения устройств MIDI к компьютеру через USB фирма Roland выпускает 64-канальный процессорный блок S-MPU64, который кроме шины USB имеет 4 входных и 4 выходных

порта MIDI. Программное обеспечение допускает объединение до 4 блоков на одной шине USB, что увеличивает число каналов до 256.

## СПИСОК ЛИТЕРАТУРЫ

1. Pohlman K.C. Principles of Digital Audio, 5rd Ed. McGraw-Hill, 2005.-860 с.
2. Цифровая звукозапись. Под редакцией Дж. Мааса и М. Веркамена. Пер. с англ.- М.: Мир, 2004.-352 с., (Мультимедиа для профи).
3. Watkinson J.R. The Art of Digital Audio. 2<sup>nd</sup> Ed.. Boston, MA: Focal Press, 1994.
4. Watkinson J.R. Coding for Digital Recording. Boston, MA: Focal Press, 1990.
6. Ковалгин Ю.А., Вологдин Э.И. Цифровое кодирование звуковых сигналов.- СПб,: Корона Принт., 2004.- 240 с.
7. Ковалгин Ю.А., Вологдин Э.И., Кацнельсон Л.С. Стереофоническое радиовещание и звукозапись. - М.: Горячая линия - Телеком, 2007.
8. Электроакустика и звуковое вещание: Учебное пособие для вузов / И.А. Алдошина, Э.И.Вологдин и др.; Под ред. Ю.А.Ковалгина.-М.: Горячая линия-Телеком, Радио и связь, 2007.- 872 с.
10. Оптические дисковые системы: Пер. с англ. /Г.Боухьюз, Дж.Браат, А.Хейсер и др.- М.: Радио и связь, 1991.-280 с.
11. Рыжков В.А. и др. Внешние запоминающие устройства на магнитном носителе. - М.: Энергия. 1978.-224 с.
12. Никамин В.А. Форматы цифровой звукозаписи. – СПб: Элби, 1998, -263 с.
13. Никамин В.А. Цифровая звукозапись. Технология и стандарты..- СПб: Наука и техника, 2002.-256 с.
14. Никамин В.А. Аналого-цифровые и цифро-аналоговые преобразователи. Справочник.-СПб: Корона Принт.-М.:”Альтекс-А”, 2003.-224 с.
15. Гитлиц М.В. Магнитная запись сигналов: Учеб. пособие для вузов. - М.: Радио и связь, 1990.
16. Щербина В.И. Цифровая звукозапись. М.: Радио и связь, 1989.
17. Золотухин И.П. Цифровые звуковые магнитофоны. Томск.: Радио и связь, 1990.
18. Х.Накадзима, Х.Огава. Цифровые грампластинки. Пер. с японского. М.:- Радио и связь, 1988, 163 с.
19. Вологдин Э.И. Аналоговая и цифровая звукозапись. Учебное пособие.-СПб ГУТ, СПб, 2004.
20. Колесников В.М. Лазерная звукозапись и цифровое радиовещание.1991.
23. Standard on Compact Disc Audio System, IEC Techn. Com. N 60: Recording, Subcommittee 60A: Sound Recording, EL 15-14-200/82,1982.
24. Standard ESMA –267. 120 mm DVD – Read – Only Disk. 2001 .
25. Standard ESMA –279. 120 mm DVD Recordable Disk (DVD-R). 1999.
30. Digital Audio Recording and Reproducing System on Magnetic Tape for Domestic Use. Dat Cassette System. Draft IEC Publication. – Doc.60A (CO) 115-1,11/1988.
30. Вологдин Э.И., Викторенков А.В., Шпигель Л.М. Автоматическая фокусировка в лазерных системах записи и воспроизведения .- Техника средств связи, сер. ТРПА, вып. 1,1978.
31. Вологдин Э.И., Коган А.М., Козмина Т.А. К расчету суперстратегии декодирования в системе «Компакт Диск».- Техника средств связи, сер. ТРПА, вып.3, 1987.
32. Вологдин Э.И. Детектирование ошибок в цифровых декодерах лазерных проигрывателей. - Техника средств связи, сер. ТРПА, вып.1, 1985.

33. Вологдин Э.И. Исправление ошибок и стираний в декодере лазерного звукопроигрывателя.- Техника средств связи, сер. ТРПА, вып.3,1987.
34. Вологдин Э.И. Динамические методы автоматической фокусировки.- Техника средств связи, сер. ТРПА, вып.2, 1979.
35. Вологдин Э.И. Как возникают и звучат ошибки квантования. - Звукорежиссер, N 6 и 7, 2007.
36. Lipshiz S.P., Vanderkooy J. Pulse-Code Modulation - An Overview. J. Audio Eng. Soc. v. 52, N3, 2004, march, pp.200...215.
37. Craven P. G. Anialis Filters and System Transient Response at High Sample Rates. J. Audio Eng. Soc. v. 52, N3, 2004, march, pp.216...242.
38. Stuart J.R. Coding for High-Resolution Audio Systems. J. Audio Eng. Soc. v. 52, N3, 2004, march, pp.117-144.
39. Reefman D., Janssen E. One-bit Audio: an Overview. J. Audio Eng. Soc. v. 52, N3, 2004, march, pp.166...189.
42. M.W. Hauser. Principles of Oversampling A/D conversion. J. Audio Eng. Soc. v. 39, N1/2, 1991, January/February, pp.3-26.
43. Blesser B. Digitization of Audio: A Comprehensive Examination of Theory, Implementation, Current Practice. J. Audio Eng. Soc. v.26, N10, 1978, october, pp.739-771.
44. Doi T.T. Channel Coding for Digital Audio Recording. J. Audio Eng. Soc. v. 31, N4, april, 1983, pp.224...238.
45. Van De Plassche R.J. Dijkman E.C. A monolithic 16-bit D/A Conversion System for Digital Audio. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 54-60.
46. James U.Lemke. The State of the Art in High-Density Magnetic Recoding. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 54-60.
47. Hiroshi Ogava, Kees A. Schouhamer Immink. EFM – The Modulation Method for the Compact Disc Digital Audio System. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 117-127.
48. E.R.Berlekamp. Error-Correcting Codes for Digital Audio. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 127-139.
49. Andrew J.Viterbi. Coding and Interleaving for Correcting Burst and Random Errors in Recording Media. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 139-147
50. Toshi T. Doi. Error Correction for Digital Audio Recordings. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 147-178.
51. L.B.Vries, K.Odaka. CIRC – The Error-Correcting Code for the Compact Disc Digital Audio System. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 178-189.
52. W.Verkaik. Compact Disc (CD) Mastering – An Industrial Process. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 189-196.
53. S.Miyaoka. Manufacturing Technology of the Compact Disc. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 196-205.
54. B.A.Blesser, Advanced Analog-to-Digital Conversion and Filtering: Data Conversion. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 37-53.
55. Toshi T.Doi. Recent Progress in Digital Audio Technology. DIGITAL AUDIO AES (Collected Papers from the AES Premiere Conference, Rye, N.Y. 1982, june 3-6), pp. 23-33.
56. Buckley K.M. A Fundamental Introduction to Compact Disc Player. 1994, pp.12. University of Minnesota. Web: <http://www.meridian-audio.com/ara>

61. Error Detection and Correction, UK Patent Application, GB No 076569 , published 2 Dec. 1981.
66. Устинов.В. Киноплёнка, магнитная лента, DVD... (Обзор. Современные носители для записи видео цифровых форматов). Научно-технический журнал 625. N 4, 2005. с. 5-16.
67. Чирков Л. Носители записи (Обзор). Звукорежиссер. N 6, 2001, с. 3-9.
68. Меерзон Б. Многодорожечные рекордеры. Звукорежиссер. N 7, 2000, с. 5-7.
69. Меерзон Б. Стереофонические рекордеры. Звукорежиссер. N10, 2000, с. 3-9.
70. Батов С. Аналого-цифровые и цифро-аналоговые преобразователи (Обзор). Звукорежиссер. N 8, 2000, с. 3-9.
71. Вейценхольд А. Звуковые рабочие станции. Звукорежиссер. N 3, 2001, с. 3-6.