УДК 681.3.06.004.239.056(075.8)

Сэедхесамеддин Дабирсиаги

ВЫБОР ЦИФРОВОГО ПРЕДСТАВЛЕНИЯ СИГНАЛА В СЛУХОВЫХ АППАРАТАХ

Сеєдхесамеддін Дабірсіагі. Вибір цифрового представлення сигналу в слухових апаратах. Розглянуті види цифрових представлень сигналу в сучасних слухових апаратах. Показано, що з погляду відношення сигнал/шум, квантування і динамічного діапазону перетворення сигналу якнайкращим видом редставленя є логарифмічне.

Ключові слова: ЦИФРОВИЙ СЛУХОВИЙ АПАРАТ, ЛОГАРИФМІЧНЕ ПЕРЕТВОРЕННЯ СИГНАЛУ, ВІДНОШЕННЯ СИГНАЛ/ШУМ, ДИНАМІЧНИЙ ДІАПАЗОН

Сэедхесамеддин Дабирсиаги. Выбор цифрового представления сигнала в слуховых аппаратах. Рассмотрены виды цифровых представлений сигнала в современных слуховых аппаратах. Показано, что с точки зрения отношения сигнал/шум, квантования и динамического диапазона преобразования сигнала наилучшим видом представления является логарифмическое.

Ключевые слова: ЦИФРОВОЙ СЛУХОВОЙ АППАРАТ, ЛОГАРИФМИЧЕСКОЕ ПРЕОБРАЗОВАНИЕ СИГНАЛА, ОТНОШЕНИЕ СИГНАЛ/ШУМ, ДИНАМИЧЕСКИЙ ДИАПАЗОН

Seyedsameddin Dabirsiadi. A choice of signal digital presentations in hearings devices. The types of signal digital presentations are considered in modern hearing devices. It is shown that from a signal/quantization noise and dynamic range viewpoint the best type of presentation is logarithmic.

Key words: DIGITAL HEARING DEVICE, LOGARITHMIC SIGNAL SHAPING, RELATION is SIGNAL/NOISE, DYNAMIC RANGE

Введение. В настоящей работе рассматриваются три вида цифровых представлений входных сигналов слуховых аппаратов (СА): линейное представление, логарифмическое представление и несколько представлений с плавающей запятой. Каждое цифровое представление отображает нормализованные сигналы в диапазоне значений от -1 до $+1$ и используется во всей модели СА. Как показано на рис. 1, аналого-цифровой преобразователь (АЦП) преобразует входной звуковой сигнал непосредственно в соответствующее цифровое представление и все внутренние сигналы остаются в этом представлении до того, как выходной цифро-аналоговый преобразователь (ЦАП) преобразует это представление обратно в аналоговую форму.

Рис. 1. Упрощенная схема цифрового СА

Любое совершенствование СА включает в себя улучшение цифровой обработки сигнала, что, в свою очередь, приводит к увеличению потребляемой мощности.

Целью работы является выбор цифрового представления сигнала, позволяющий увеличить отношение сигнал/шум квантования и динамический диапазон преобразования сигнала.

Основная часть. Рассмотрим вначале стандартное 16-ти битное ***линейное представление*** сигнала в дополнительном двоичном коде. В нем 0 показывает количество битов, отведенных на целое число, а 15 – количество битов для представления числа после запятой. Входные значения варьируются от -1 до $1-2^{-15}$, с постояным шагом квантования равным 2^{-15}. Это представление широко используется в устройствах цифровой обработки сигналов [1]. Преобразовать двоичное число в действительное число можно с помощью формулы:

$$x = -a_{15} \cdot 2^0 + \sum_{i=0}^{14} a_i \cdot 2^{i-15}. \qquad (1)$$

Каждое преобразование из действительного числа в двоичное число начинается с 16-ти битной последовательности:

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
a_{15}	a_{14}	a_{13}	a_{12}	a_{11}	a_{10}	a_9	a_8	a_7	a_6	a_5	a_4	a_3	a_2	a_1	a_0

где бит a_{14} имеет значение ½ = 2^{-1}, бит a_{13} = ¼ = 2^{-2} и так далее до бита $a_0 = 1/2^{15} = 2^{-15}$.

Предположим, что необходимо преобразовать число $-0{,}6768$. Это отрицательное число, поэтому сначала определяем его абсолютное значение, то есть 0,6768. Далее добавляем $\frac{1}{2}\cdot\frac{1}{2^{15}}$, чтобы учесть любые округления, и получаем 0,676830517578. Это значение больше чем 2^{-1}, поэтому биту a_{14} присваивается значение 1, а 2^{-1} вычитается из входных данных. В результате остается 0,176830517578. Это значение сравнивается с 2^{-2}. Поскольку 2^{-2} больше остатка, биту a_{13} присваивается значение 0. Далее остаток сравнивается с 2^{-3} и оказывается, что его значение больше. Поэтому биту a_{12} присваивается значение 1, а 2^{-3} вычитается из входных данных. Теперь входные данные имеют значение 0,051830517578 и т.д. Результатом расмотренного процесса является двоичное число:

0	1	0	1	0	1	1	0	1	0	1	0	0	0	0	1
a_{15}	a_{14}	a_{13}	a_{12}	a_{11}	a_{10}	a_9	a_8	a_7	a_6	a_5	a_4	a_3	a_2	a_1	a_0

Не следует забывать, что входные данные первоначально имели отрицательное значение. Для того, чтобы это учесть, используется оператор инверсии двоичного числа и добавляется 1. При этом бит a_{15} становится равным 1, что обозначает знак числа:

1	0	1	0	1	0	0	1	0	1	0	1	1	1	1	1
a_{15}	a_{14}	a_{13}	a_{12}	a_{11}	a_{10}	a_9	a_8	a_7	a_6	a_5	a_4	a_3	a_2	a_1	a_0

Таким образом, число $x = -0{,}6768$ в двоичной линейной форме представления выглядит как 1010100101011111, число x=0,0143 как 0000000111010101, число x=0,5574 как 0100011101011001, число x=0,9826 как 0111110111000110 и т.д.

В качестве примера обратной операции, преобразуем двоичную форму 0000000111010101 в число 0,0143 при помощи (1):

$$x = -0\cdot 2^0 + 0\cdot 2^{14-15} + 0\cdot 2^{13-15} + 0\cdot 2^{12-15} + 0\cdot 2^{11-15} + 0\cdot 2^{10-15} + 0\cdot 2^{9-15} + 1\cdot 2^{8-15} + 1\cdot 2^{7-15} +$$

$$+ 1\cdot 2^{6-15} + 0\cdot 2^{5-15} + 1\cdot 2^{4-15} + 0\cdot 2^{3-15} + 1\cdot 2^{2-15} + 0\cdot 2^{1-15} + 1\cdot 2^{0-15} = 0{,}014312744141.$$

Результат этого преобразования, 0,014312744141, наиболее близок к 0,0143 в этом представлении. Следующее меньшее возможное значение равно 0,014282226563, что больше отличается от 0,0143, чем полученное значение.

Логарифмическое представление основывается на логарифмическом характере восприятия звука человеческим ухом. Как известно [2], 9-ти битное логарифмическое представление имеет основание 0,941. Такое представление имеет формат «знак – величина», с одним знаковым битом (старшим) и 8 последующими битами. В нашей работе логарифмические значения обозначены индексом l, то есть $x_l = \log_{0,941}(x)$. Это представление также позволяет изменять значения сигнала в пределах от -1 до +1. В случае основания 0,941, если $x > y$ в линейном пространстве, то $x_l < y_l$ в логарифмическом пространстве. Например, максимальное значение числа $x = 1$ представляется как

$x_l = 000000000$, а минимальное положительное значение числа $x = 1,8423 \cdot 10^{-7}$ представляется как $x_l = 011111111$. С другой стороны, отрицательное число с наименьшей величиной, $x = -1,8423 \cdot 10^{-7}$ представляется как $x_l = 111111111$, а отрицательное число с наибольшей величиной $x = -1$ представляется как $x_l = 100000000$.

Эти соотношения показаны в виде графиков на рис. 2, 3, а на рис. 4 линейные значения чисел отложены в логарифмическом масштабе.

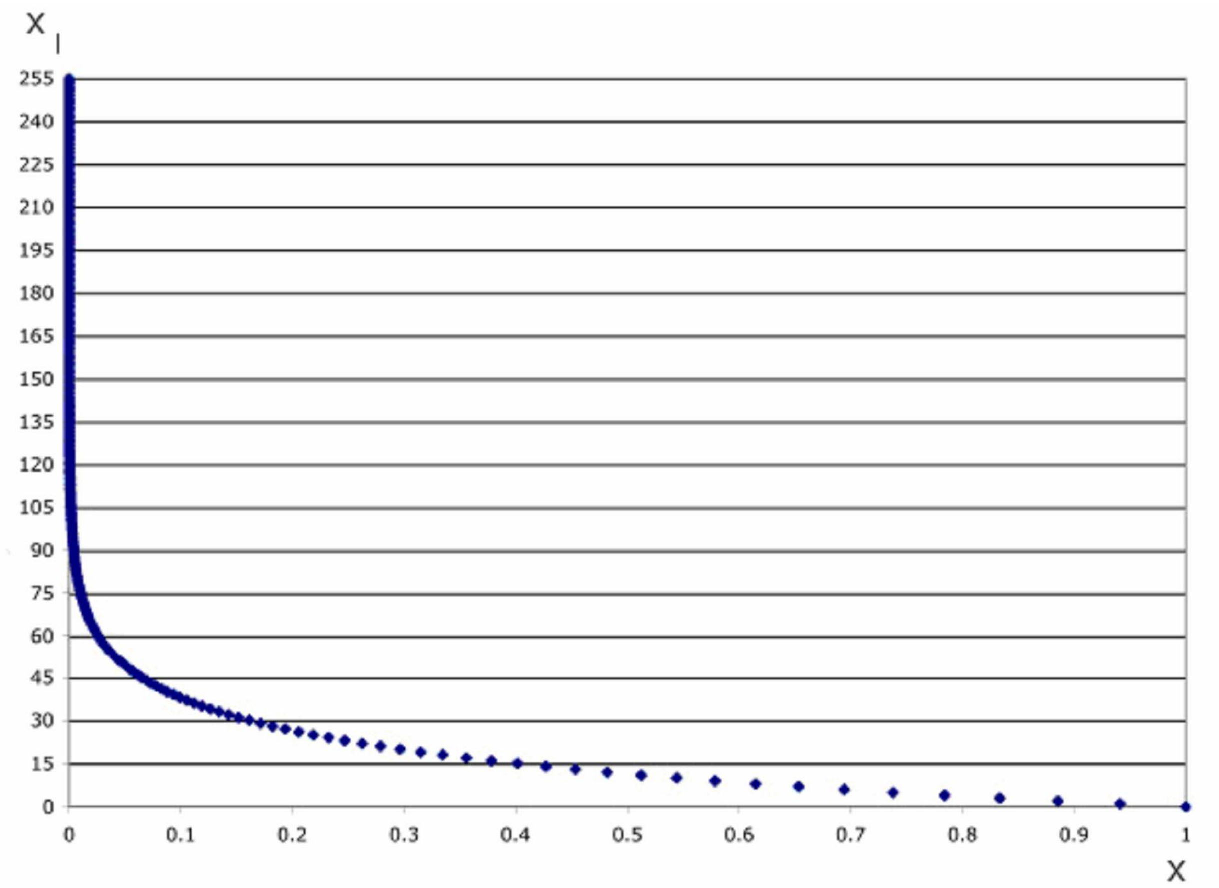

Рис. 2. Прямая зависимость между линейными и логарифмическими значениями чисел

Значения в логарифмическом представлении могут быть вычислены из двоичного представления при помощи уравнения

$$x_l = (-1)^{a_8} \cdot (0,941)^{\sum_{i=1}^{7} a_i \cdot 2^i} . \qquad (2)$$

Каждое число является последовательностью, состоящей из 9 битов:

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_8 & a_7 & a_6 & a_5 & a_4 & a_3 & a_2 & a_1 & a_0 \end{matrix}$$

Самый простой способ получения логарифмического представления – преобразование сначала в целое число, а затем в двоичный код. Это осуществляется путем вычисления $\log_{0,941}$ от входных данных, то есть $\dfrac{\log_{10}(\textit{входные данные})}{\log_{10}(0,941)}$. Вновь обратимся к примеру, в котором входное число имеет значение $-0,6768$. Так как это число отрицательное, то биту a_8 присваивается значение 1. Затем используется абсолютное значение 0,6768. Далее берется логарифм этого значения: $\log_{0,941}(0,6768) = 6,41...$. Это значение округляется до 6 и преобразуется в двоичный код в ходе процесса, который очень похож на процесс в линейном преобразовании. Однако здесь биты имеют значения 1, 2, 4, 8, … , 128 справа налево. Это дает следующее логарифмическое двоичное представление числа 0,6768:

$$\begin{array}{ccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ a_8 & a_7 & a_6 & a_5 & a_4 & a_3 & a_2 & a_1 & a_0 \end{array}$$

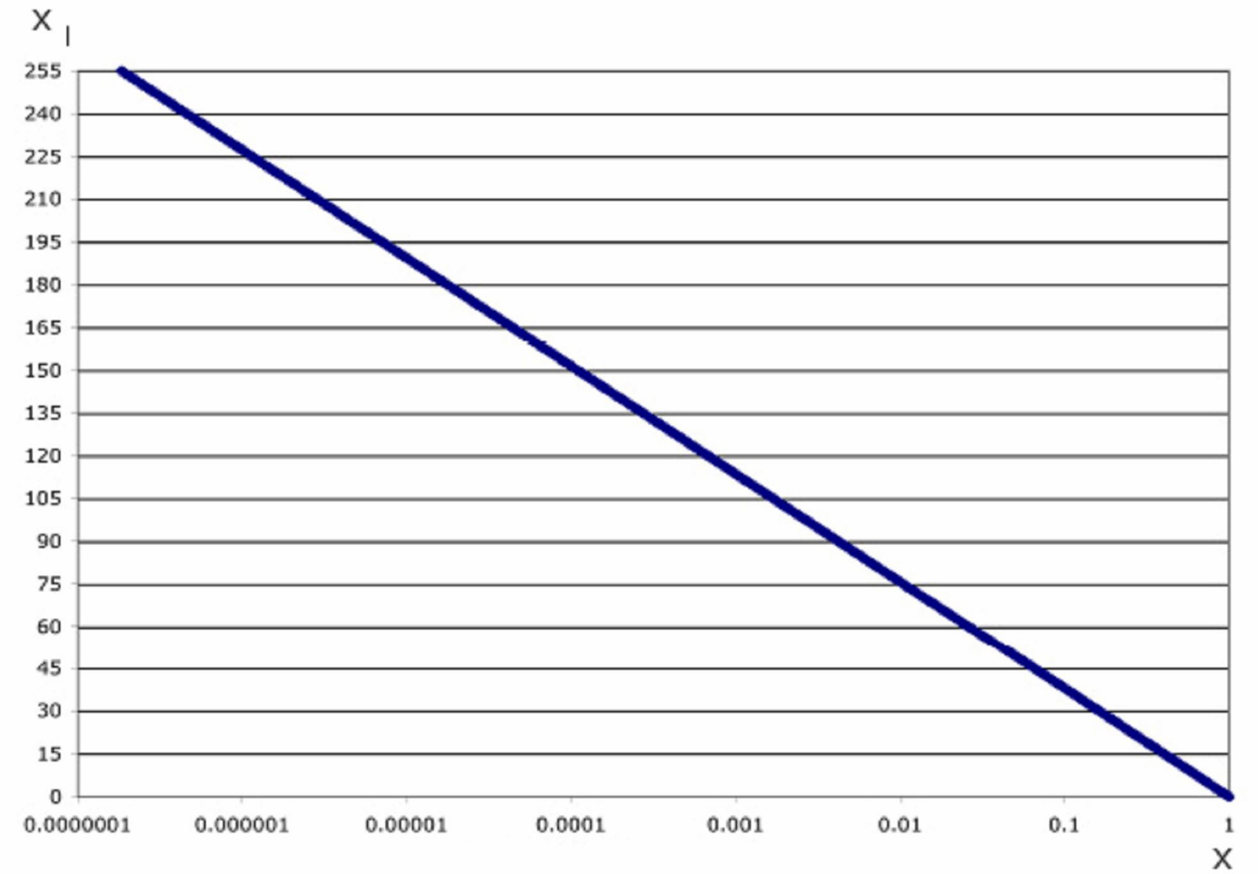

Рис. 3. Логарифмическая зависимость между линейными и логарифмическими значениями чисел

Таким образом, для ранее приведенных чисел их логарифмические эквиваленты имеют вид:

$$\begin{array}{rl} -0,6768 & 100000110 \\ 0,0143 & 001000110 \\ 0,5574 & 000001010 \\ 0,9826 & 000000000 \end{array}$$

В качестве примера обратной операции, преобразуем код 001000110 в число 0,0143 при помощи (2):

$$-1^0 \cdot 0,941^{\left(0\cdot 2^8+0\cdot 2^7+1\cdot 2^6+0\cdot 2^5+0\cdot 2^4+0\cdot 2^3+1\cdot 2^2+1\cdot 2^1+0\cdot 2^0\right)} = 0,014166861232.$$

Результат этого преобразования, 0,014166861232, является ближайшим значением к 0,0143, которое может быть получено при помощи такого представления. При этом с увеличением линейных положительных значений соответствующие значения в двоичном представлении уменьшаются. Фактически, значение 0,9826 настолько близко к единице, что в двоичном представлении оно является нулем.

На рис. 4 значение $x = 1$, приведенное на рис. 2, показано в увеличенном масштабе.

Значение 0,9826 находится в пределах области погрешностей 1, следовательно, представлено как 1, что в двоичной логарифмической форме выглядит как 000000000.

На рис. 5 показан противоположный край рис. 2 в увеличенном масштабе.

Здесь каждое логарифмическое значение занимает намного меньший интервал на линейной оси, а наименьшее представимое значение не тождественно нулю.

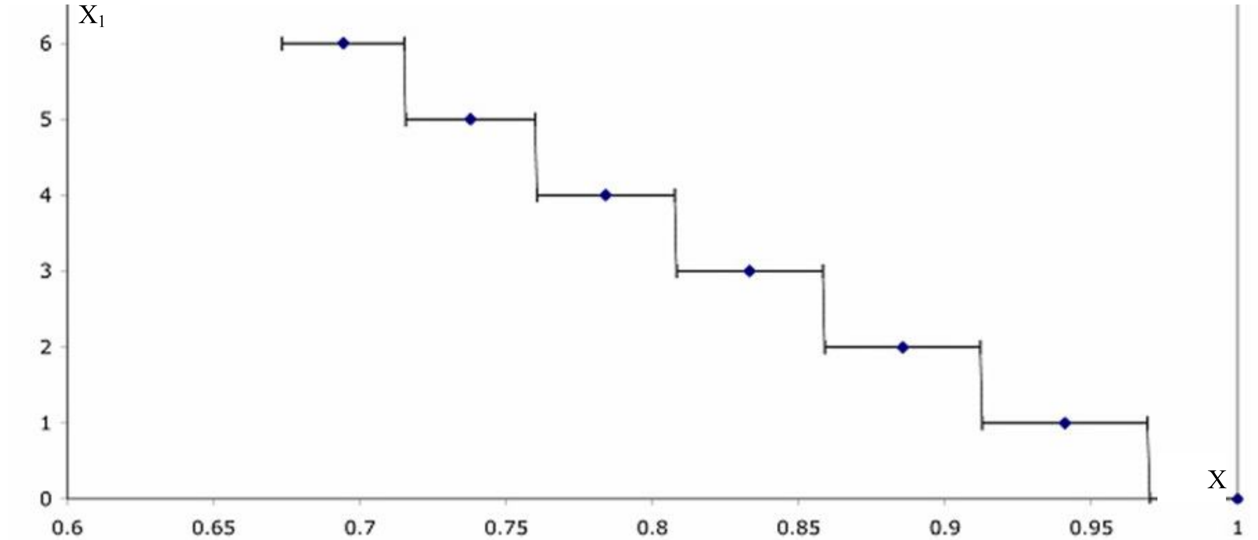

Рис. 4. Зависимость между линейными и логарифмическими значениями чисел в окрестности точки $x = 1$

Числа с плавающей запятой имеют характеристики как линейных, так и логарифмических представлений. При этом мантисса функционирует как линейная часть, а порядок – как логарифмическая часть. Распределение значений этих чисел на вещественной оси не равномерное из-за логарифмического характера порядков и разниц между диапазонами значений мантиссы. Стандартом для чисел с плавающей запятой, предложенным IEEE, является 32-битное число одинарной точности, содержащее один знаковый бит, восемь битов, отведенных для значений порядка, и 23 бита, отведенных под мантиссу числа. Порядок записывается в избыточном представлении со смещением 127, то есть порядок 0 представляется как -127, а порядок 200 представляется как $200-127$=73. Поскольку мантисса нормализована, на данные отводится 24 бита, а первый бит мантиссы, всегда равный единице, подразумевается, но не хранится. Ненормализованные мантиссы, нули и бесконечности выделены в особые случаи.

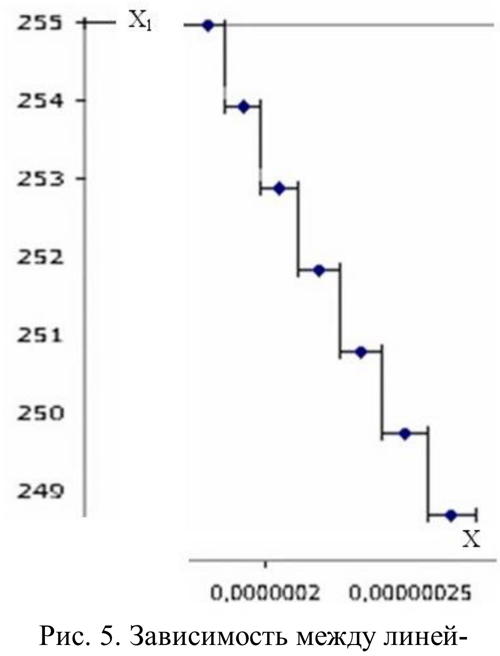

Рис. 5. Зависимость между линейными и логарифмическими значениями чисел в окрестности точки $x = 0$

В настоящей работе приведены результаты анализа 10 различных представлений с плавающей запятой. Девять из них имеют вид $1-E-M$: 1 знаковый бит, E битов порядка, M битов мантиссы. То есть представление 1–4–5 отображает 1 знаковый бит, 4 бита порядка и 5 бит мантиссы. В представлениях рассматривались виды от 1–4–4 до 1–6–6 в различных комбинациях порядков и мантисс. Мантисса представлялась в виде «знак – величина» и связывалась со знаковым битом. Рассматривались только нормализованные мантиссы, с целью упростить конструкцию и сократить число вычислений, исключив сравнения, которые могли бы понадобиться в специальных случаях. Кроме того, сохранялся старший бит мантиссы, поскольку не было необходимости "экономить" область памяти, к тому же этот бит требовался для вычислений. Порядок записывался в избыточном представлении на основании количества бит, выделенных под порядок числа.

Для расчета десятичного значения числа можно воспользоваться выражением:

$$x_{пл} = \frac{1}{k}(-1)^{зн} \cdot \frac{мант}{2^M} \cdot 2^{пор\left(-2^{E-1}\right)}, \tag{3}$$

где k масштабирует входные данные в пределах от -1 до 1, то есть

$$k = \frac{2^M - 1}{2^M} \cdot 2^{\left(2^E - 1\right) - 2^{E-1}}. \tag{4}$$

Обозначения *мант* и *пор* указывают на целые значения соответствующих бит:

$$мант = \sum_{i=0}^{M-1} m_i \cdot 2^i, \tag{5}$$

$$пор = \sum_{i=0}^{E-1} e_i \cdot 2^i, \tag{6}$$

а E и M – на количество бит в порядке и мантиссе. Поскольку масштаб не влияет на вычисления масштабный коэффициент k опущен при создании совокупности значений входных данных. Полярность чисел отображается знаковым битом 0 для неотрицательных чисел, и битом 1 для отрицательных чисел.

В качестве примера представления с плавающей запятой используем вид 1–4–5 (в зависимости от избранного вида представления, могут использоваться от 9 до 13 бит):

Знаковый бит	Порядок				Мантисса				
s	e_3	e_2	e_1	e_0	m_4	m_3	m_2	m_1	m_0
0	0	0	0	0	0	0	0	0	0

Биту s присваивается значение 1, поскольку входное число, $-0{,}6768$, имеет отрицательное значение. После этого преобразуется абсолютное значение, то есть:

$$0,6768 = \frac{\frac{мант}{2^M} \cdot 2^{пор\left(-2^{E-1}\right)}}{k}.$$

Далее вычисляется масштабный коэффициент k (4). Для вида 1–4–5 $k = \left(2^5 - 1\right) \cdot 2^{-5} \cdot 2^{\left(2^4 - 1\right) - 2^{4-1}} = 124$. Теперь входное положительное число умножается на коэффициент k, что дает: $0,6768 \cdot 124 = 83,9232 = \frac{мант}{2^5} \cdot 2^{пор\left(-2^{4-1}\right)}$.

Для вычисления мантиссы необходимо сначала определить наибольшее значение порядка, которое меньше масштабированного входного значения: $\frac{83,9232}{2^{пор(-8)}} < 1$.

Значением порядка может быть целое число в пределах от 0 до 15 включительно. В данном случае значение порядка должно равняться 15, поэтому битам с e_3 по e_0 присваиваются значения 1. Таким образом, $83,9232 = \frac{мант}{32} \cdot 2^{15-8}$. Откуда $мант = 20,9808$. Но мантисса также должна быть выражена целым числом, поэтому полученный результат округляется до значения 21. Теперь можно присвоить двоичные значения оставшимся битам для числа 21, то есть битам m_0, m_2 и m_4 присваиваются значения 1.

Знаковый бит	Порядок				Мантисса				
s	e_3	e_2	e_1	e_0	m_4	m_3	m_2	m_1	m_0
1	1	1	1	1	1	0	1	0	1

Таким образом, рассмотренные ранее числа для представления с плавающей запятой в виде 1–4–5 выглядят так: $-0,6768 \to 1111110101$; $0,0143 \to 0100111100$; $0,5574 \to 0111110001$; $0,9826 \to 0111111110$.

В качестве примера обратной операции, преобразуем 0100111100 в число 0,0143, используя уравнения (3), (5) и (6) для $k = 124$:

$$мант = 1\cdot 2^4 + 1\cdot 2^3 + 1\cdot 2^2 + 0\cdot 2^1 + 0\cdot 2^0 = 28,$$

$$пор = 1\cdot 2^3 + 0\cdot 2^2 + 0\cdot 2^1 + 1\cdot 2^0 = 9,$$

$$x_{пл} = \frac{(-1)^0 \cdot \frac{28}{2^5} \cdot 2^{9-2^{4-1}}}{124} = 0,014112903226.$$

Результат этого преобразования, 0,014112903226, является наиболее близким значением к 0,0143. Следующим ближайшим значением, большим 0,0143 в этом представлении будет число 0,014616935484.

Для реализации возможностей увеличения отношения сигнал/шум квантования (ОСШК) и динамического диапазона преобразования сигнала было разработано ***модифицированное представление с плавающей запятой***. В нем значения мантиссы и порядка представлены в дополнительном коде. Мантисса объединена со знаковым битом, из-за чего вид 1–4–5 преобразуется в вид 4–6.

Новые значения мантиссы и порядка вычисляются следующим образом:

$$пор = -e_3 \cdot 2^3 + \sum_{i=0}^{2} e_i \cdot 2^i, \tag{7}$$

$$мант = -m_5 \cdot 2^5 + \sum_{i=0}^{4} m_i \cdot 2^i. \tag{8}$$

Значение числа с плавающей запятой определяется равенством:

$$x_{пл} = \frac{\frac{мант}{2^{M-1}} \cdot 2^{пор}}{j}, \tag{9}$$

где j – коэффициент, который незначительно корректируется по сравнению с коэффициентом k:

$$j = \frac{2^{M-1}-1}{2^{M-1}} \cdot 2^{E} - 1. \qquad (10)$$

Для примера представим число в виде 4–6.

Порядок				Мантисса					
e_3	e_2	e_1	e_0	m_5	m_4	m_3	m_2	m_1	m_0
0	0	0	0	0	0	0	0	0	0

Важно учесть знак преобразуемого числа, $-0{,}6768$, потому что последующие действия выполняются с его абсолютным значением. В данном случае масштабный коэффициент $j = \frac{2^{6-1}-1}{2^{6-1}} \cdot 2^{4-1} - 1 = 6{,}75$, поэтому 0,6768 · 6,75 = 4,5684. После этого следует снова определить наибольшее целое значение порядка, которое меньше входного значения, следовательно, $4{,}5684 / 2^{пор} < 1$. Для этого необходимо, чтобы значение *пор* равнялось 3. Теперь, $4{,}5684 = \frac{мант}{2^{6-1}} \cdot 2^{3}$ и, следовательно, $мант = 18{,}2736$ и округляется до 18.

Так как первоначальное значение было отрицательным, то мантиссу необходимо скорректировать. В результате получим

Порядок				Мантисса					
e_3	e_2	e_1	e_0	m_5	m_4	m_3	m_2	m_1	m_0
0	0	1	1	1	0	1	1	1	0

Таким образом, 0011101110 = $-0{,}663594470046$. Следующим большим значением будет 0011101111 = -0,700460829493. Рассмотренные ранее числа в виде 4–6 выглядят так: $-0{,}6768 \rightarrow$ 0011101110; 0,0143 $\rightarrow$ 1101011001; 0,5574 $\rightarrow$ 0010011110; 0,9826 $\rightarrow$ 0011011011.

На рис. 6…10 показаны графики ОСШК для линейного, логарифмического и представлений с плавающей запятой и порядком четыре.

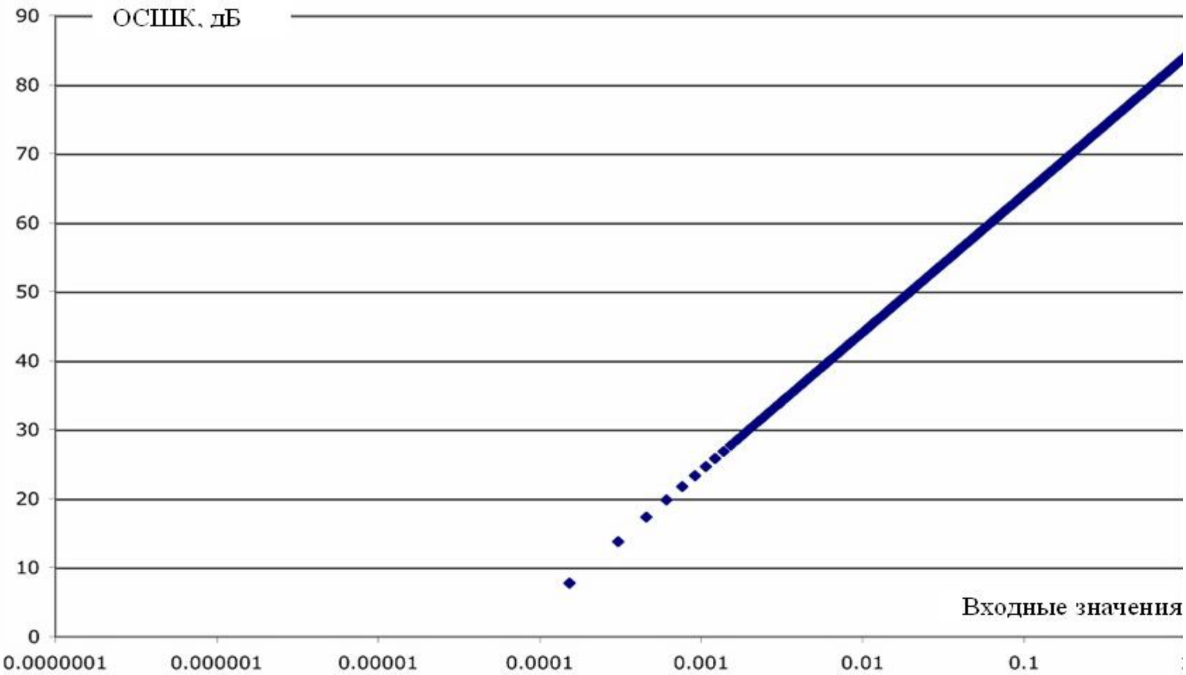

Рис. 6. ОСШК для линейного представления

Для всех этих представлений, за исключением логарифмического, ОСШК рассчитывалось при условии синусоидального входного сигнала с равномерным распределением шума квантования между уровнями U_i и U_{i+1}:

$$\text{ОСШК} = 20\log_{10}\left(\frac{\frac{U_i}{\sqrt{2}}}{\frac{U_{i+1}-U_i}{\sqrt{12}}}\right). \quad (11)$$

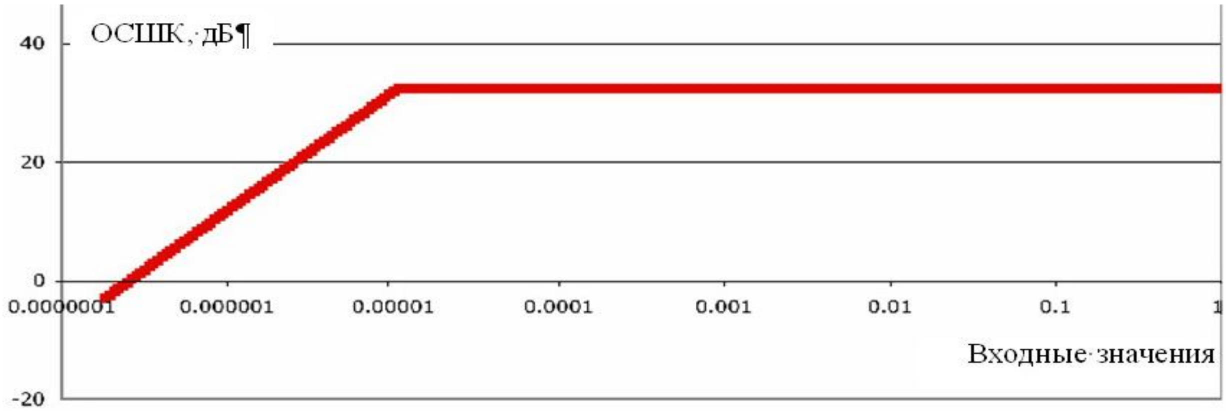

Рис. 7. ОСШК для логарифмического представления

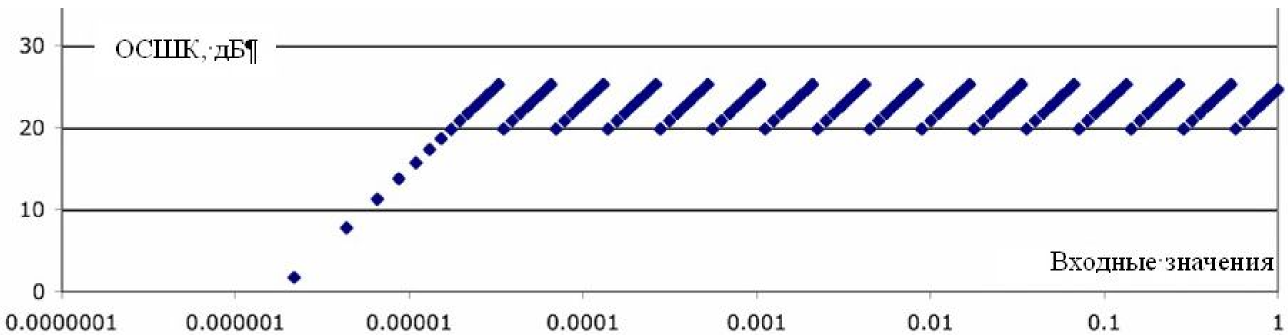

Рис. 8. ОСШК для представления с плавающей запятой вида 1–4–4

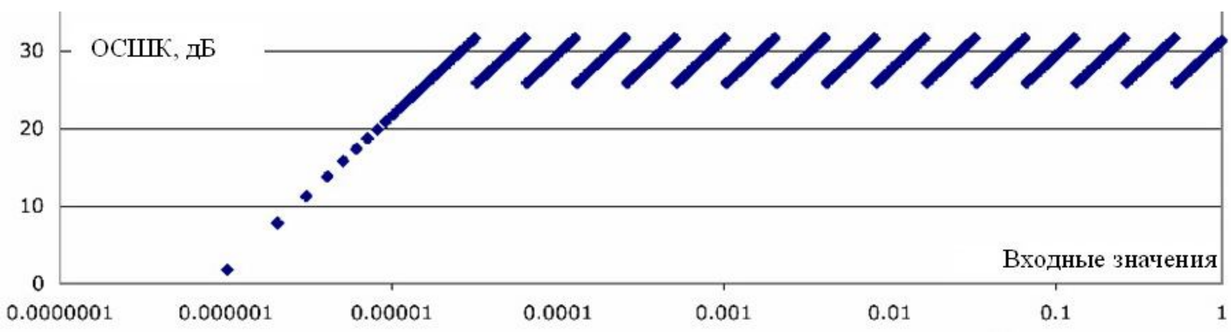

Рис. 9. ОСШК для представления с плавающей запятой вида 1–4–5

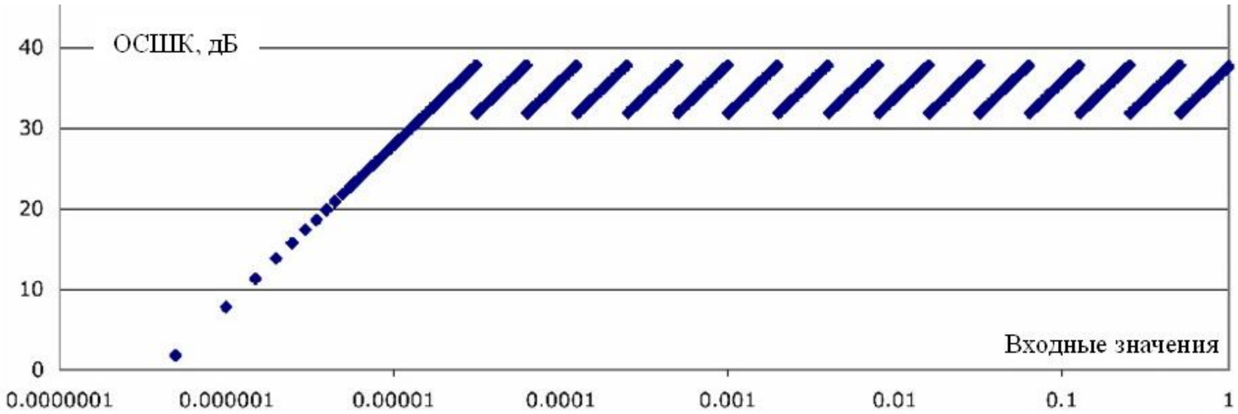

Рис. 10. ОСШК для представления с плавающей запятой вида 1–4–6

Для логарифмического представления, учитывая зону нечувствительности между наименьшим представимым значением и нулем, ОСШК рассчитывалось по формуле:

$$\text{ОСШК} = 20\log_{10}\left[\frac{\dfrac{U_i}{\sqrt{2}}}{\max\left(\dfrac{U_{i+1}-U_i}{\sqrt{12}},\, u_{\min}\right)}\right]. \qquad (12)$$

Стандартное 16-битное линейное представление имеет переменное ОСШК, изменяющееся в пределах от 0 до 96 дБ в динамическом диапазоне преобразования сигнала (рис. 6).

Слабые входные сигналы легко теряются в шуме с такой характеристикой ОСШК. Логарифмическое представление имеет значение ОСШК ≈ 33 дБ практически по всему динамическому диапазону преобразования, что очень близко соответствует условиям, требуемым для различения человеческой речи (рис. 7). Это обстоятельство и послужило причиной выбора значения 0,941 в качестве основания логарифма [2]. На рис. 7 зона нечувствительности около нуля не показана.

Для представлений с плавающей запятой и порядками 5 и 6 характерны ОСШК, очень похожие на ОСШК представлений с порядком 4, поскольку ОСШК преимущественно основывается на значении мантиссы (рис. 8–10). Аналогично, модифицированное представление 4–6 имеет такое же ОСШК, как и представление 1–4–5 (рис. 9). Выделение под значения порядка большего количества бит только расширяет динамический диапазон представления. Как можно увидеть из рисунков, ОСШК для представлений с плавающей запятой варьируется в пределах центрального значения. Это объясняется тем, что мантисса линейна, а порядок носит логарифмический характер. Как только мантисса достигает максимального значения, происходит прибавление порядка, после чего мантисса снова увеличивается.

Динамический диапазон линейного представления составляет 96 дБ, что практически равно требуемым для СА 100 дБ. Динамические диапазоны других представлений показаны на рис. 11.

Для представлений с плавающей запятой динамический диапазон преимущественно определяется значениями порядка. Увеличение мантиссы вызывает незначительное увеличение динамического диапазона, а увеличение порядка вызывает значительные скачки между видами представлений. Модифицированное представление с плавающей запятой вида 4–6 имеет такие же свойства, как и представление вида 1–4–5 и может его заменить.

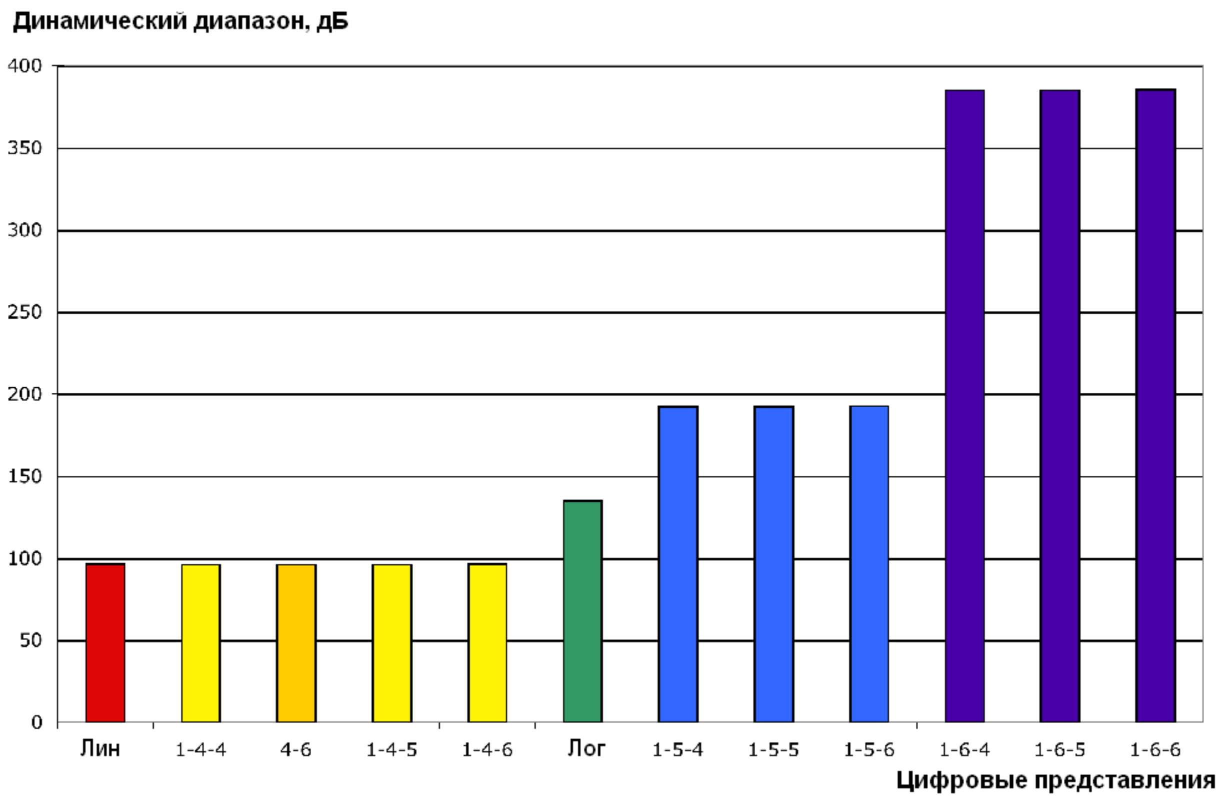

Рис. 11. Динамический диапазон различных представлений

Как видно на рис. 11, динамический диапазон всех представлений, включая логарифмическое, такой же или лучше, чем динамический диапазон линейного представления.

Выводы

1. Сравнение цифровых представлений звуковых сигналов по отношению сигнал/шум квантования и динамическому диапазону показало преимущество 9-битного логарифмического представления с основаниеми 0,9412.

2. Для представлений с плавающей запятой “отношение сигнал/шум квантования” варьируется в пределах центрального значения, а динамический диапазон преобразования преимущественно определяется значениями порядка числа.

Литература

1. Айфичер Э. С. Цифровая обработка сигналов: практический подход / Айфичер Э. С., Джервис Б. У.; пер. с англ. – [2-е изд.]. – М.: Издательский дом «Вильямс», 2004. – 992 с.

2. Engel George. Design and Analysis of Logarithmic Digital Quantizers with Applications to Low-Power Data Interfaces for Speech Processing. – Doctoral Dissertation, Dept. of Electrical Engineering, Washington University. – 1990.