

дисперсией<sup>1</sup>. В первой колонке представлены различные типы квантователей. Во второй колонке — отношение сигнал/шум без адаптации. В третьей и четвертой колонках — отношение сигнал/шум при адаптации на основе оценки дисперсии (5.61) с  $M=128$  и  $1024$  соответственно. Для данного речевого сигнала адаптивные квантователи позволяют получить выигрыш в отношении сигнал/шум на 5,6 дБ. Сходных результатов можно ожидать и на других фразах речевого сигнала. Таким образом, адаптивный квантователь обладает преимуществом по сравнению с неадаптивным неравномерным квантователем. Дополнительное преимущество адаптивных квантователей, не отраженное в табл. 5.4, заключается в том, что соответствующим выбором  $\Delta_{min}$  и  $\Delta_{max}$  можно увеличить отношение сигнал/шум при сохранении малых шумов незанятого канала и широкого динамического диапазона. Это справедливо и в общем случае для большинства правильно спроектированных адаптивных схем. Таким образом, адаптивное квантование позволяет достигнуть лучших результатов по сравнению с мгновенным компрессированием и квантованием по минимуму дисперсии ошибки.

#### 5.4.2. Адаптация по выходному сигналу

Квантователь второго типа показан на рис. 5.27 и 5.28, где отмечено, что дисперсия входного сигнала оценивается по выходному квантованному сигналу или по последовательности кодовых слов. Как и в случае адаптации по входному сигналу, в данном случае шаг квантования и коэффициент усиления прямо и обратно пропорциональны дисперсии входного сигнала в соответствии

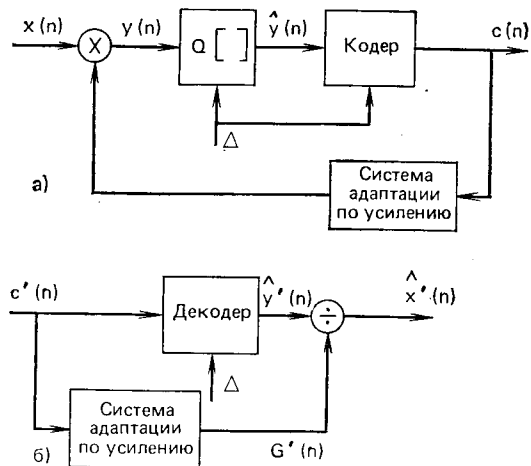


Рис. 5.27. Общая схема адаптации по выходу с переменным усилением: а) кодер; б) декодер

<sup>1</sup> Результаты получены на реальном речевом сигнале.

с (5.57) и (5.58). Такие схемы обладают важным преимуществом, состоящим в том, что шаг квантования или коэффициент усиления не требуется хранить или передавать по каналу связи, поскольку они получены по последовательности кодовых слов. Недостатком подобных квантователей является высокая чувствительность к ошибкам в кодовых словах, ибо эти ошибки приводят не только к неправильной установке уровня квантования, но и к ошибкам в шаге квантования.

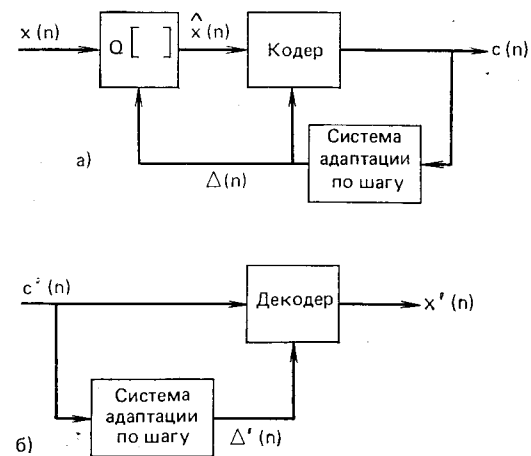


Рис. 5.28. Общая схема адаптации по выходу с переменным шагом: а) кодер; б) декодер

Один из простых подходов состоит в использовании уравнения (5.53) непосредственно для выходного сигнала квантователя:

$$\sigma^2(n) = \sum_{m=-\infty}^{\infty} \hat{x}^2(m) h(n-m). \quad (5.62)$$

В данном случае, однако, нельзя применить буфер для использования нереализуемого фильтра, т. е. оценка дисперсии должна быть основана только на последнем отсчете  $\hat{x}(n)$ , поскольку текущее значение  $\hat{x}(n)$  можно получить только после квантования, которое осуществляется после оценки дисперсии. Можно, например, использовать фильтр с импульсной характеристикой

$$h(n) = \begin{cases} \alpha^{n-1}, & n \geq 1; \\ 0, & \text{в противном случае} \end{cases} \quad (5.63)$$

как в уравнении (5.55). Фильтр может также иметь импульсную характеристику, равную

$$h(n) = \begin{cases} 1/M, & 1 \leq m \leq M; \\ 0, & \text{в противном случае,} \end{cases} \quad (5.64)$$

так что

$$\sigma^2(n) = \frac{1}{M} \sum_{m=n-M}^{n-1} x^2(m). \quad (5.65)$$

Такой квантователь исследован Ноллом [12], который обнаружил, что удовлетворительные результаты по настройке констант  $\Delta_0$  или  $G_0$  в (5.57) и (5.58) (отношение сигнал/шум квантования порядка 12 дБ при трехразрядном квантователе) достигаются при ис-

пользовании окна протяженностью всего в два отсчета. Большие значения  $M$  приводят лишь к незначительному улучшению результатов.

Несколько иной подход (рис. 5.28) подробно исследован Джаянтом [15]. Здесь шаг квантования в адаптивном квантователе изменяется в соответствии с уравнением

$$\Delta(n) = P \Delta(n-1), \quad (5.66)$$

где множитель  $P$  зависит лишь от значения предшествующего кодового слова  $|c(n-1)|$ . Это показано на рис. 5.29 для трехразряд-

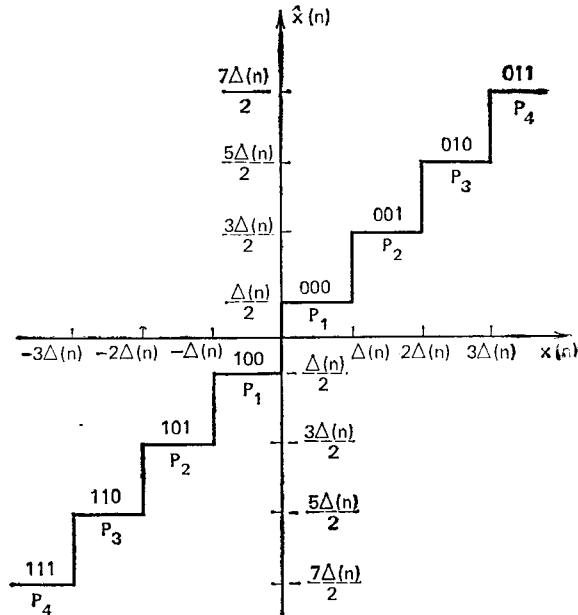


Рис. 5.29. Характеристика трехразрядного адаптивного квантователя

ного равномерного квантователя, где старший разряд — знаковый, а остаток кодового слова — величина отсчета. При этом

$$\hat{x}(n) = \frac{\Delta(n) \text{sign}(c(n))}{2} + \Delta(n) c(n), \quad (5.67)$$

где  $\Delta(n)$  удовлетворяет уравнению (5.66). Поскольку  $\Delta(n)$  зависит от предшествующего значения шага квантования и предшествующего значения кодового слова, последовательность кодовых слов полностью описывает сигнал. На практике необходимо ввести ограничения

$$\Delta_{min} \leq \Delta(n) \leq \Delta_{max}. \quad (5.68)$$

Отношение  $\Delta_{max}/\Delta_{min}$  определяет динамический диапазон шага квантования.

Способ изменения значения  $P$  в уравнении (5.66) на основе кодового слова  $|c(n-1)|$  очевиден. Если предшествующее кодовое слово соответствует наибольшему положительному или наибольшему отрицательному уровню, то естественно предположить, что квантователь переполнился и шаг квантования слишком мал. В этом случае множитель должен быть больше единицы. Наоборот, в том случае, если кодовое слово соответствует наименьшему положительному или наименьшему отрицательному уровню, целесообразно уменьшить шаг квантования путем умножения его на число, меньшее единицы. Разработка подобного квантователя включает выбор множителя, соответствующего каждому из  $2^b$  кодовых слов  $B$ -разрядного квантователя. Джаянт [15] решил эту проблему путем определения набора множителей, которые минимизируют мощность шума квантования. Были получены теоретические результаты для сигнала, распределенного по закону Гаусса, и эмпирические результаты для речевого сигнала на основе итеративных вычислений. Результаты, полученные Джаянтом, приведены на рис. 5.30, где показана приближенная зависимость множителей от величины

$$Q = [1 + 2 |c(n-1)|] / (2^B - 1). \quad (5.69)$$

Заштрихованная область соответствует ожидаемому изменению множителей при изменении статистики входного сигнала или  $B$ . Конкретные значения множителей будут лежать внутри заштрихованной области. Важно подчеркнуть, что множители будут такими, чтобы нарастание происходило более интенсивно, чем убывание. В табл. 5.5 приведены значения множителей для  $B=2, 3, 4, 5$ .

Увеличение отношения сигнал/шум, достигаемое при использовании метода адаптивного квантования, иллюстрирует табл. 5.6. Множители из табл. 5.5 обеспечивают  $\Delta_{max}/\Delta_{min} = 100$ . Как следует из табл. 5.6, достигается выигрыш около 4,7 дБ по отношению к квантованию на основе  $\mu$ -закона. Улучшение на 2—4 дБ достигается также по отношению к неадаптивному случаю. В другом исследовании Нолл [7] установил, что отношения сигнал/шум для трехразрядного квантователя, построенного по  $\mu$ -закону, и адаптивного квантователя составляют 9,4 и 14,1 дБ соответственно. В этом эксперименте, в отличие от эксперимента Джаянта, множители были равны {0,8; 0,8; 1,3; 1,9}. Тот факт, что столь различ-

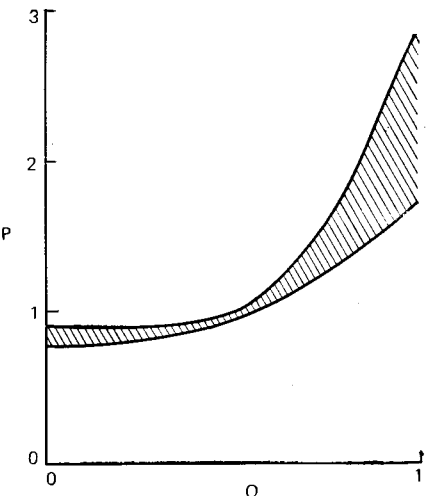


Рис. 5.30. Общая форма функции кратности речевого квантователя для  $B > 2$  [15]

но различны, что столь различ-

Таблица 5.5

Множители изменения шага в методах адаптивного квантования

B	Тип кодера	
	ИКМ	РИКМ
2	0,6; 2,2	0,8; 1,6
3	0,85; 1; 1; 1,5	0,9; 0,9; 1,25; 1,75
4	0,8; 0,8; 0,8; 0,8 1,2; 1,6; 2,0; 2,4	0,9; 0,9; 0,9; 0,9 1,2; 1,6; 2,0; 2,4
5	0,85; 0,85; 0,85; 0,85 0,85; 0,85; 0,85; 0,85 1,2; 1,4; 1,6; 1,8; 2,0; 2,2; 2,4; 2,6	0,9; 0,9; 0,9; 0,9; 0,95; 0,95; 0,95; 0,95; 1,2; 1,5; 1,8; 2,1; 2,4; 2,7; 3,0; 3,3

ные множители в разных экспериментах дают сравнимые результаты, означает, что значения множителей не столь существенны.

Таблица 5.6

Выигрыш в отношении сигнал/шум, дБ, при использовании оптимальных множителей изменения шага для адаптивного квантования [15]

B	Логарифмическая ИКМ с $\mu$ -законом ( $\mu=100$ ) квантования	Адаптивная ИКМ с равномерным квантованием
2	3	9
3	8	15
4	15	19

### 5.4.3. Общие замечания

Как следует из результатов данного параграфа, имеется практически неограниченное множество методов адаптивного квантования. Большинство этих методов приводит к выигрышу в отношении сигнал/шум по сравнению с квантованием по  $\mu$ -закону и обеспечивает такой же динамический диапазон. Выбирая малым  $\Delta_{min}$ , можно уменьшить шум незанятого канала. Таким образом, адаптивное квантование обладает рядом полезных свойств. Однако не следует ожидать, что дальнейшее совершенствование методов адаптации даст существенный выигрыш в скорости передачи, поскольку эти методы используют лишь наши знания о распределении мгновенных значений речевого сигнала. Поэтому в следующем параграфе рассмотрим использование корреляции между соседними отсчетами на основе методов разностного квантования.

### 5.5. Общая теория разностного квантования

Анализ рис. 5.7 позволяет сделать вывод, что между соседними отсчетами сигнала имеется значительная корреляция, которая слабо убывает по мере увеличения интервала между отсчетами.

Это означает, что, вообще говоря, сигнал изменяется медленно и разность между соседними отсчетами будет иметь меньшую дисперсию, чем исходный сигнал, в чем можно легко убедиться (см. задачу 5.10). Такие рассуждения позволяют ввести общие методы разностного квантования, показанные на рис. 5.31 [16, 17]. Здесь на входе квантователя действует сигнал

$$d(n) = x(n) - \tilde{x}(n), \quad (5.70)$$

который представляет собой разность сигнала  $x(n)$  и оценки предсказанного значения входного сигнала, которое обозначено через  $\tilde{x}(n)$ . Предсказанное значение представляет собой выходной сигнал предсказателя  $P$ ; входным сигналом, как будет ясно из дальнейшего, является квантованный входной сигнал  $x(n)$ .

Разностный сигнал можно также назвать погрешностью предсказания, поскольку это — величина, на которую предсказанное значение отличается от точного значения входного сигнала. Оставив пока в стороне вопрос о вычислении предсказанного значения, отметим, что вместо входного сигнала квантованию подвергается разностный сигнал. Квантователь может быть адаптивный или неадаптивный, равномерный или неравномерный, но во всех случаях его параметры должны соответствовать дисперсии погрешности предсказания. Квантованная погрешность может быть представлена в виде

$$\hat{d}(n) = d(n) + e(n), \quad (5.71)$$

где  $e(n)$  — ошибка квантования. В соответствии с рис. 5.31а погрешность с выхода квантователя складывается с предсказанным значением для получения квантованного сигнала:

$$\hat{x}(n) = \tilde{x}(n) + \hat{d}(n). \quad (5.72)$$

Подставляя (5.70) и (5.71) в (5.72), видим, что

$$\hat{x}(n) = x(n) + e(n), \quad (5.73)$$

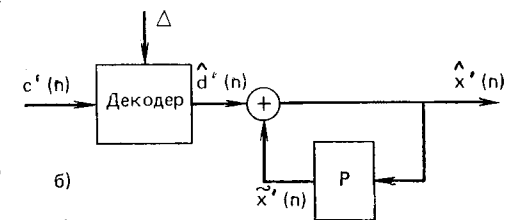
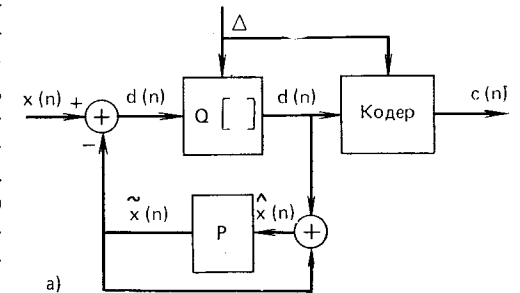


Рис. 5.31. Общая схема разностного квантования:  
а) кодер; б) декодер

Таким образом, независимо от свойств устройства, обозначенного через  $p$ , квантованный речевой сигнал отличается от входного только на величину шума квантования разностного сигнала. Таким образом, если предсказатель хороший, дисперсия  $d(n)$  будет меньше, чем дисперсия  $x(n)$ , так что квантователь с заданным количеством уровней даст меньшую погрешность при квантовании разности, чем при квантовании исходного сигнала.

Следует отметить, что для передачи по каналу связи или занесения в запоминающее устройство используется квантованный разностный сигнал. Схема, восстанавливающая входной сигнал по последовательности кодовых слов, неявно имеется и на рис. 5.31а. В явном виде она показана на рис. 5.31б и содержит декодер, восстанавливающий разностный сигнал, и предсказатель (такой же, как и на передающей стороне). Очевидно, что при совпадении  $c'(n)$  и  $c(n)$  сигнал  $\hat{x}'(n) = \hat{x}(n)$  и отличается от  $x(n)$  лишь ошибкой квантования, вносимой квантованием сигнала  $d(n)$ .

Отношение сигнал/(шум квантования) для системы рис. 5.31 равно

$$SNR = E[x^2(n)]/E[e^2(n)] = \sigma_x^2/\sigma_e^2. \quad (5.74)$$

Выражение (5.74) можно переписать в виде

$$SNR = (\sigma_x^2/\sigma_d^2) (\sigma_d^2/\sigma_e^2) = G_p \cdot SNR_Q, \quad (5.75)$$

где

$$SNR_Q = \sigma_d^2/\sigma_e^2 \quad (5.76)$$

представляет собой отношение сигнал/(шум квантования), а величина

$$G_p = \sigma_x^2/\sigma_d^2 \quad (5.77)$$

определяется как коэффициент усиления, обусловленный разностным кодированием<sup>1</sup>.

Величина  $SNR_Q$  зависит от конкретно используемого квантователя и известных свойств  $d(n)$ ,  $SNR_Q$  и может быть максимизирован на основе методов предыдущего параграфа. Величина  $G_p$  определяет выигрыш в отношении сигнал/шум при использовании разностного представления. Очевидно, что необходимо увеличить  $p$  выбором схемы предсказания. Для данного сигнала  $\sigma_x^2$  — величина фиксированная, следовательно,  $G_p$  можно максимизировать за счет минимизации знаменателя в (5.77), т. е. минимизацией дисперсии погрешности предсказания.

Для решения поставленной задачи следует определить тип предсказателя  $p$ . Один из методов, хорошо согласующийся с предыдущими рассуждениями о моделях сигнала и приводящий к несложным математическим выкладкам, состоит в использовании

<sup>1</sup> Величина  $G_p$  является коэффициентом усиления системы, обратной предсказателю, т. е. авторегрессионной модели сигнала. Этот коэффициент характеризует эффективность предсказания. (Прим. ред.)

линейного предсказания, т. е.  $\tilde{x}(n)$  представляется в виде линейной комбинации предшествующих значений сигнала

$$\tilde{x}(n) = \sum_{k=1}^p \alpha_k \hat{x}(n-k). \quad (5.78)$$

Предсказанное значение является, таким образом, выходным сигналом фильтра с передаточной функцией вида

$$P(z) = \sum_{k=1}^p \alpha_k z^{-k}, \quad (5.79)$$

на вход которого поступает восстановленный сигнал  $\hat{x}(n)$ . Попутно отметим, что восстановленный сигнал представляет собой выходной сигнал фильтра с передаточной функцией

$$H(z) = \frac{1}{1 - \sum_{k=1}^p \alpha_k z^{-k}}, \quad (5.80)$$

на вход которого поступает квантованная погрешность. Дисперсия погрешности предсказания на рис. 5.31

$$\begin{aligned} \sigma_d^2 &= E[d^2(n)] = E[(x(n) - \tilde{x}(n))^2] = E\left[x(n) - \sum_{k=1}^p \alpha_k \hat{x}(n-k)\right]^2 = \\ &= E\left[x(n) - \sum_{k=1}^p \alpha_k x(n-k) - \sum_{k=1}^p \alpha_k e(n-k)\right]^2. \end{aligned} \quad (5.81)$$

Для вычисления множества параметров предсказания, которые минимизируют  $\sigma_d^2$ , продифференцируем  $\sigma_d^2$  по каждому параметру и приравняем производную к нулю. Это приводит к системе уравнений

$$\begin{aligned} \frac{\partial \sigma_d^2}{\partial \alpha_j} &= -2E\left[\left[x(n) - \sum_{k=1}^p \alpha_k (x(n-k) + e(n-k))\right][x(n-j) + \right. \\ &\left. + e(n-j)\right] = 0, \quad 1 \leq j \leq p. \end{aligned} \quad (5.82)$$

Уравнения (5.82) можно переписать в более компактной форме:

$$E[(x(n) - \tilde{x}(n)) \hat{x}(n-j)] = E[d(n) \hat{x}(n-j)] = 0, \quad 1 \leq j \leq p. \quad (5.83)$$

Если коэффициенты предсказания таковы, что  $\sigma_d^2$  минимально, то погрешность предсказания не коррелирована с последними значениями сигнала на входе предсказателя, т. е. ортогональна  $\hat{x}(n-j)$  при  $1 \leq j \leq p$ .

Уравнения (5.82) можно представить в виде

$$E[x(n-j)x(n)] + E[e(n-j)x(n)] = \sum_{k=1}^p \alpha_k E[x(n-j)x(n-k)] +$$

$$+ \sum_{k=1}^p \alpha_k E[e(n-j)x(n-k)] + \sum_{k=1}^p \alpha_k E[x(n-j)e(n-k)] + \\ + \sum_{k=1}^p \alpha_k E[e(n-j)e(n-k)], \quad (5.84)$$

где  $1 \leq j \leq p$ . Предполагая, что квантование достаточно точное, допустим, что  $e(n)$  не коррелирована с  $x(n)$  и  $e(n)$  — стационарная последовательность. Тогда

$$E[x(n-j)e(n-k)] = 0 \text{ для всех } n, j \text{ и } k \quad (5.85)$$

$$E[e(n-j)e(n-k)] = \sigma_e^2 \delta(j-k). \quad (5.86)$$

Используя эти предположения, уравнение (5.84) можно упростить:

$$\varphi(j) = \sum_{k=1}^p \alpha_k [\varphi(j-k) + \sigma_e^2 \delta(j-k)], \quad (5.87) \\ 1 \leq j \leq p,$$

где  $\varphi(j)$  — автокорреляционная функция  $x(n)$ . Если разделить обе части этого уравнения на  $\sigma_x^2$  и обозначить нормированную корреляционную функцию через

$$\rho(j) = \varphi(j)/\sigma_x^2, \quad (5.88)$$

то (5.87) переписывается в матричной форме в виде

$$\rho = \mathbf{C} \alpha, \quad (5.89a)$$

где

$$\mathbf{C} = \begin{vmatrix} \left(1 + \frac{1}{SNR}\right) & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & \left(1 + \frac{1}{SNR}\right) & \dots & \rho(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(p-1) & \rho(p-2) & \dots & \left(1 + \frac{1}{SNR}\right) \end{vmatrix}, \quad (5.896)$$

$$\rho = \begin{vmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{vmatrix}, \quad \alpha = \begin{vmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{vmatrix} \quad (5.89b), (5.89g)$$

и  $SNR = \sigma_x^2/\sigma_e^2$ . Таким образом, вектор оптимальных коэффициентов предсказания получается как решение матричного уравнения (5.89a):

$$\alpha = \mathbf{C}^{-1} \rho. \quad (5.90)$$

В общем случае матрица  $\mathbf{C}^{-1}$  может быть вычислена различными численными методами, в том числе и методом, основанным на том обстоятельстве, что  $\mathbf{C}$  — тридиагональная матрица (см. гл. 8). Однако в общем случае уравнения (5.89a) неразрешимы, поскольку матрица содержит члены, которые зависят от отношения сигнал/шум [см. (5.89b)], а отношение сигнал/шум, в свою очередь, определяется коэффициентами линейного предсказания (5.89a). Один из выходов из этого положения заключается в том, чтобы пренебречь членом  $1/SNR$  в (5.89). При  $p=1$  такое упрощение не требуется, поскольку уравнение (5.90) можно решить непосредственно:

$$\alpha_1 = \frac{\rho(1)}{1 + (1/SNR)}. \quad (5.91)$$

Уравнение (5.91) показывает, что  $\alpha_1 < \rho(1)$ .

Несмотря на трудности прямого решения уравнений относительно коэффициентов предсказания, можно получить выражение для оптимального  $G_p$  через  $\alpha_i$ . Для этого определим  $\sigma_d^2$ , переписав (5.81) в виде

$$\sigma_d^2 = E[(x(n) - \tilde{x}(n))(x(n) - \tilde{x}(n))] = E[(x(n) - \tilde{x}(n))x(n)] - \\ - E[(x(n) - \tilde{x}(n))\tilde{x}(n)]. \quad (5.92)$$

Используя (5.83), можно показать, что для оптимальных коэффициентов предсказания второе слагаемое в последнем уравнении обращается в нуль, т. е. предсказанное значение не коррелировано с погрешностью предсказания (см. задачу 5.12). Таким образом, можно записать

$$\sigma_d^2 = E[(x(n) - \tilde{x}(n))x(n)] = E[x^2(n)] - E\left[\sum_{k=1}^p \alpha_k (x(n-k) + e(n-k))x(n)\right]. \quad (5.93)$$

Используя предположение о некоррелированности сигнала и шума, получаем

$$\sigma_d^2 = \sigma_x^2 - \sum_{k=1}^p \alpha_k \varphi(k) = \sigma_x^2 \left[1 - \sum_{k=1}^p \alpha_k \rho(k)\right]. \quad (5.94)$$

Учитывая (5.77), имеем

$$(G_p)_{opt} = \left[1 - \sum_{k=1}^p \alpha_k \rho(k)\right]^{-1} \quad (5.95)$$

где  $\alpha_k$  удовлетворяют уравнению (5.89a). При  $p=1$  можно оценить эффект влияния субоптимального значения  $\alpha_1$  на величину  $G_p = \sigma_x^2/\sigma_d^2$ . Из (5.95) получим

$$(G_p)_{opt} = [1 - \alpha_1 \rho(1)]^{-1}. \quad (5.96)$$

Выбирая произвольное значение  $\alpha_1$  и повторяя дифференцирование, приводящее к (5.94), получим

$$\sigma_d^2 = \sigma_x^2 [1 - 2\alpha_1 \rho(1) + \alpha_1^2] + \alpha_1^2 \sigma_e^2 \quad (5.97)$$

или

$$(G_p)_{arb} = \left[ 1 - 2\alpha_1 \rho(1) + \alpha_1^2 \left( 1 + \frac{1}{SNR} \right) \right]^{-1}. \quad (5.98)$$

Член  $\alpha_1^2/SNR$  показывает увеличение дисперсии  $d(n)$  из-за шума  $e(n)$  в петле обратной связи. Легко показать (см. задачу 5.13), что (5.98) можно переписать в виде

$$(G_p)_{arb} = \frac{1 - (\alpha_1^2 | SNR_Q)}{1 - 2\alpha_1 \rho(1) + \alpha_1^2} \quad (5.99)$$

при любом  $\alpha_1$ , включая оптимальное. Например, если  $\alpha_1 = \rho(1)$  (что соответствует субоптимальному случаю в соответствии с (5.91)), получаем

$$(G_p)_{subopt} = \frac{1 - (\rho^2(1)/SNR_Q)}{1 - \rho^2(1)} = \left[ \frac{1}{1 - \rho^2(1)} \right] \left[ 1 - \frac{\rho^2(1)}{SNR_Q} \right]. \quad (5.100)$$

Таким образом, коэффициент, рассчитанный при отсутствии квантователя, уменьшается за счет второго сомножителя в (5.100), обусловленного шумом квантования в петле обратной связи.

Для получения оптимального коэффициента усиления можно продифференцировать уравнение (5.99) по  $\alpha_1$ . Тогда получим уравнение<sup>1</sup>

$$\frac{d(G_p)}{d\alpha_1} = 0, \quad (5.101)$$

которое можно решить при оптимальном  $\alpha_1$ .

Предположим, что можно пренебречь слагаемым  $1/SNR$  в (5.89). Тогда для предсказания первого порядка из (5.91) следует, что  $\alpha_1 = \rho(1)$  и коэффициент имеет вид

$$(G_p)_{opt} = [1 - \rho^2(1)]^{-1}. \quad (5.102)$$

Таким образом, пока  $\rho(1) \neq 0$ , отношение сигнал/шум будет увеличиваться за счет предсказания. Ранее приводилась типичная корреляционная функция (см. рис. 5.5) для речевого сигнала на выходе полосового фильтра и фильтра нижних частот при частоте дискретизации 8 кГц [4]. Заштрихованная область показывает изменения  $\rho(n)$  корреляции для четырех различных дикторов, а центральная кривая — среднее значение корреляционной функции по всем четырем дикторам. Из кривых видно, что при частоте дискретизации, соответствующей частоте Найквиста, справедливо неравенство

$$\rho(1) > 0,80. \quad (5.103)$$

<sup>1</sup> Авторы выражают благодарность профессору П. Ноллу за полезные замечания по изложенному анализу.

Это означает, что

$$(G_p)_{opt} > 2,77 \text{ (или 4,43 дБ)}. \quad (5.104)$$

Нолл [4] использовал данные рис. 5.5 для вычисления оптимального значения  $G_p$  как функции от  $p$  для сегментов речи длительностью 55 с, полученных как на выходе фильтра нижних частот, так и на выходе полосового фильтра. Результаты приведены на рис. 5.32<sup>1</sup>. Заштрихованная область показывает разброс результа-

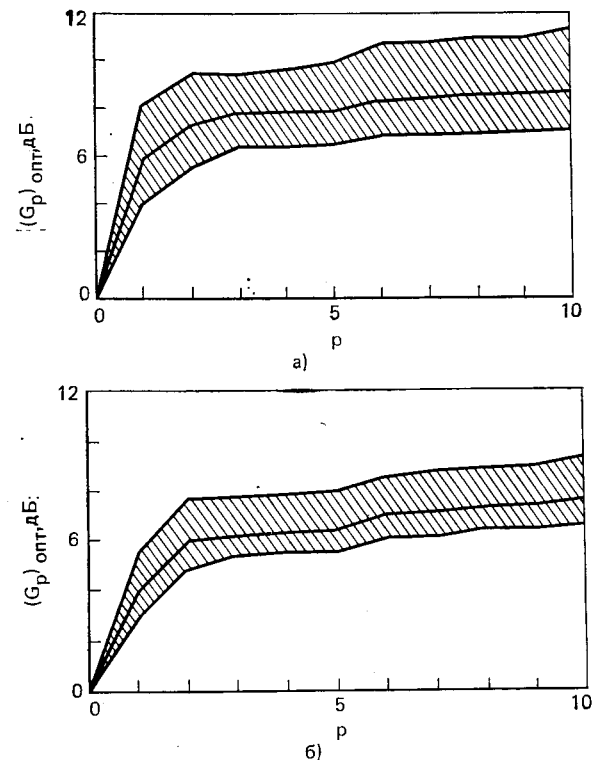


Рис. 5.32. Зависимости оптимального по критерию SNR усиления  $G$  от числа коэффициентов предсказания: а) низкочастотная фильтрация речи; б) высокочастотная фильтрация речи [7]

тов для четырех дикторов, центральная кривая получена усреднением по всем дикторам. Видно, что даже при простом предсказателе можно получить выигрыш 6 дБ. Это эквивалентно добавлению еще одного разряда в квантователь, однако, поскольку на самом деле ничего не добавляется, скорость остается неизменной. Отметим также, что коэффициент не достигает 12 дБ ни при каком  $p$ , т. е. невозможно получить выигрыш, эквивалентный добавлению

<sup>1</sup> Как и ранее, частота дискретизации 8 кГц, т. е.  $n$  надо умножить на 125 мкс.

двух разрядов. С другой стороны, применение разностного квантования позволяет понизить скорость передачи при том же отношении сигнал/шум. Это достигается, конечно, ценой усложнения квантователя.

Несколько основных результатов применения схем разностного квантования вытекают из рис. 5.5 и 5.32. Во-первых, разностное квантование обеспечивает выигрыш по сравнению с непосредственным квантованием. Во-вторых, величина выигрыша зависит от величины корреляции. В-третьих, один и тот же предсказатель не может быть оптимальным для различных дикторов и различного речевого сигнала. Эти обстоятельства стимулируют разработку усовершенствованных схем, в которых сохраняется основная структура, изображенная на рис. 5.31. В них применяются различные адаптивные или неадаптивные квантователи или предсказатели для достижения лучшего качества или меньшей скорости передачи речевого сигнала. Далее рассматриваются примеры, позволяющие судить о возможностях таких систем.

## 5.6. Дельта-модуляция

Примером простого применения разностного квантования является дельта-модуляция (ДМ) [18—24]. В системах такого типа частота дискретизации выбирается во много раз больше, чем частота Найквиста. В результате соседние отсчеты оказываются в большой степени коррелированными. Это следует из результатов § 5.2, где показано, что автокорреляционная функция последовательности отсчетов представляет собой дискретизированную автокорреляционную функцию непрерывного сигнала:

$$\varphi(m) = \varphi_a(mT). \quad (5.105)$$

Используя свойства автокорреляционной функции, естественно предположить, что она возрастает при  $T \rightarrow 0$ . Действительно, можно считать, что, за исключением случая некоррелированного сигнала, имеет место соотношение

$$\varphi(1) \rightarrow \sigma_x^2 \text{ при } T \rightarrow 0. \quad (5.106)$$

Большая корреляция между отсчетами означает, что при уменьшении  $T$  можно более точно предсказать текущий отсчет по предшествующим и, следовательно, уменьшить дисперсию погрешности предсказания. Поэтому более «грубый» квантователь может дать хорошие результаты. В действительности в системе с дельта-модуляцией используется простой одноуровневый (двухуровневый) квантователь. Таким образом, скорость передачи при использовании ДМ численно равна частоте дискретизации.

### 5.6.1. Линейная дельта-модуляция

Схема простейшей системы с дельта-модуляцией приведена на рис. 5.33. В этом случае квантователь имеет только два уровня и шаг квантования фиксирован. Положительный уровень кванто-

вания соответствует  $c(n) = 0$ , а отрицательный  $c(n) = 1$ . Таким образом,

$$\hat{d}(n) = \begin{cases} \Delta, & c(n) = 0; \\ -\Delta, & c(n) = 1. \end{cases} \quad (5.107)$$

На рис. 5.33 показан простой одношаговый предсказатель первого порядка, для которого оптимальный коэффициент усиления

$$(G_p)_{\text{opt}} = [1 - \rho^2(1)]^{-1}. \quad (5.108)$$

Таким образом, если  $\rho(1) \rightarrow 1$ , то  $(G_p)_{\text{opt}} \rightarrow \infty$ . Этот результат носит качественный характер, ибо при использовании «грубого» двухуровневого квантователя предположения, при которых получено выражение для  $(G_p)_{\text{opt}}$ , не справедливы.

Эффект квантования можно увидеть из рис. 5.34а, где показаны аналоговый сигнал  $x_a(t)$ , результирующие отсчеты  $x(n)$ ,  $\hat{x}(n)$  и  $\hat{x}'(n)$  при данном периоде дискретизации в предположении, что  $\alpha$  (множитель в

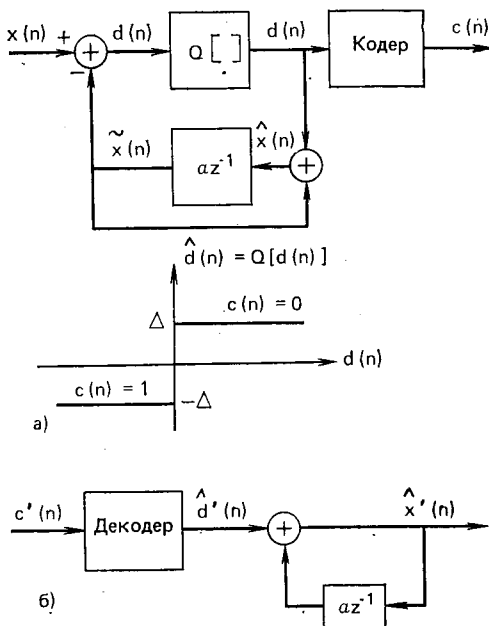


Рис. 5.33. Структурная схема системы с дельта-модуляцией: а) кодер; б) декодер

петле обратной связи) равен единице. Из рис. 5.33а видно, что  $\hat{x}(n)$  в общем случае удовлетворяет уравнению

$$\hat{x}(n) = \alpha \hat{x}(n-1) + \hat{d}(n). \quad (5.109)$$

При  $\alpha \approx 1$  уравнение описывает дискретный аналог интегратора в том смысле, что осуществляется накопление положительных и отрицательных приращений величины  $\Delta$ . Отметим, что входной сигнал квантователя имеет вид

$$d(n) = x(n) - \hat{x}(n-1) = x(n) - x(n-1) - e(n-1). \quad (5.110)$$

Таким образом, пренебрегая ошибкой квантования  $\hat{x}(n-1)$ , значение  $d(n)$  можно представить как разность первого порядка сигнала  $x(n)$ . Разность может рассматриваться как аппроксимация производной входного сигнала, а ее вычисление — как операция, обратная цифровому интегрированию. Если крутизна входного сигнала максимальна, то очевидно, что для того, чтобы последова-

тельность отсчетов  $\{\hat{x}(n)\}$  возрастала так же быстро, как и последовательность  $\{x(n)\}$  в области максимальной крутизны, необходимо потребовать выполнения неравенства

$$\frac{\Delta}{T} \geq \max \left| \frac{dx_a(t)}{dt} \right|. \quad (5.111)$$

Иначе восстановленный сигнал будет «отставать» от исходного, как это показано в левой части рис. 5.34а. Если условие (5.111) выполняется, то такие искажения, называемые *перегрузкой по*

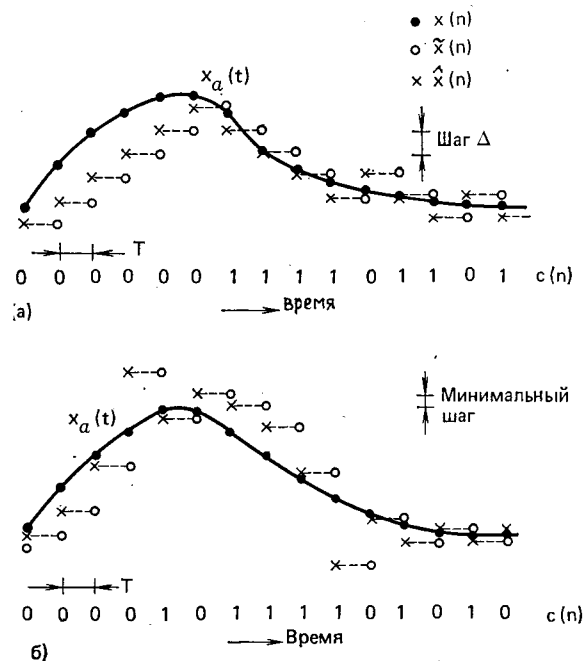


Рис. 5.34. Иллюстрация дельта-модуляции: а) фиксированный шаг; б) адаптивный шаг

*крутизне*, отсутствуют. Отметим, что поскольку максимальная крутизна  $\hat{x}(n)$  ограничивается шагом квантования, то возрастание или убывание последовательности  $\hat{x}(n)$  происходит по соответствующей ступенчатой линии. По этой причине фиксированную (неадаптивную) дельта-модуляцию иногда называют линейной (ЛДМ).

Шаг квантования определяет также и максимальную ошибку, когда крутизна мала. Например, если сигнал на входе равен нулю (канал не занят), сигнал на выходе квантователя представляет собой переменную последовательность нулей и единиц, что приводит к флуктуации восстановленного сигнала вокруг нулевого или иного постоянного уровня с размахом  $\Delta$ . Этот тип ошибок квантования, изображенный в правой части рис. 5.34а, называется *шумом дробления*.

Ранее было показано, что для получения большого динамического диапазона необходимо иметь большой шаг квантования; в то же время для точного описания малых сигналов шаг квантования должен быть малым. В данном случае это относится к динамическому диапазону и амплитуде разностного сигнала (или производной аналогового сигнала). Интуитивно ясно, что выбор шага квантования, минимизирующего среднее квадратическое значение шума квантования, приведет к компромиссу между перегрузкой по крутизне и шумом дробления.

На рис. 5.35, заимствованном из подробного исследования дельта-модуляции, проведенного Абатом [21], изображена зави-

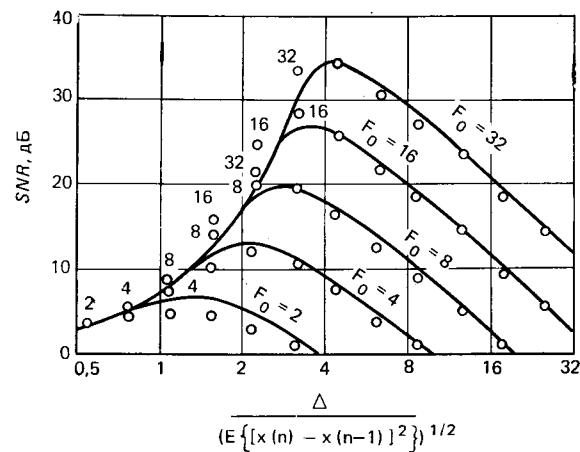


Рис. 5.35. SNR для дельта-модуляторов как функция нормированного шага [21]

симость отношения сигнал/шум от  $\Delta / (E[(x(n) - x(n-1))^2])^{1/2}$  и параметра  $F_0 = F_s / 2F_N$ , где  $F_s$  — частота дискретизации, а  $F_N$  — частота Найквиста. Отметим, что

$$\text{Скорость передачи} = F_s (1 \text{ бит}) = F_s = 2 F_N F_0. \quad (5.112)$$

Таким образом, параметр  $F_0$  играет здесь ту же роль, что и количество двоичных единиц на отсчет в многоуровневом квантователе при дискретизации с удвоенной частотой Найквиста. Кривые получены для сигнала с равномерным спектром в полосе и гауссовским распределением. Для речевого сигнала отношение сигнал/шум несколько больше вследствие большей корреляции, хотя форма кривых такая же. Из рис. 5.35 видно, что для данного значения  $F_0$  отношение сигнал/шум достигает максимума при некотором  $\Delta$ . Значения  $F_0$ , лежащие левее этого максимума, соответствуют перегрузке по крутизне, а правее — шуму дробления. Абат [21] вывел эмпирическую формулу

$$\Delta_{\text{opt}} = \{E[(x(n) - x(n-1))^2]\}^{1/2} \ln(2F_0) \quad (5.113)$$



для оптимального шага квантования, т. е. для расположения максимумов кривой при заданной  $F_0$ . Из рис. 5.35 следует, что оптимальное значение сигнал/шум увеличивается на 9 дБ при удвоении  $F_0$ . Поскольку удвоение эквивалентно удвоению  $F_s$ , можно сказать, что при удвоении скорости передачи отношение сигнал/шум возрастает на 9 дБ. Таким образом, в отличие от ИКМ, где при удвоении количества двоичных единиц на отсчет достигается увеличение отношения сигнал/шум, равное 6 дБ на каждый добавленный бит, здесь увеличение происходит значительно быстрее.

Другая важная особенность кривых рис. 5.35 состоит в том, что они весьма «остроконечны», т. е. отношение сигнал/шум в большой степени зависит от уровня входного сигнала (отметим, что  $E[(x(n) - x(n-1))^2] = 2\sigma_x^2(1 - \rho(1))$ ). Таким образом, для получения отношения сигнал/шум, равного 35 дБ при частоте Найквиста 3 кГц, требуется скорость передачи 200 кбит/с. Однако даже на этой скорости при постоянном шаге квантования требуемое качество может быть достигнуто для весьма узкого диапазона уровней входного сигнала. Для получения хорошего качества восстановленного речевого сигнала, сравнимого с качеством, достигаемым в семиразрядной логарифмической ИКМ, требуется значительно большая скорость.

Основное достоинство ЛДМ состоит в ее простоте. Система может быть реализована на простом аналоговом или цифровом интеграторе и, поскольку используется только одноразрядный код, не требует никакой синхронности по кодовым словам между передатчиком и приемником. Ограничения линейной дельта-модуляции состоят, главным образом, в весьма грубом квантовании погрешности предсказания. С учетом предыдущего обсуждения адаптивных методов квантования естественно предположить, что использование этого подхода в данном случае может существенно улучшить характеристики дельта-модулятора. Наибольший интерес представляют простые адаптивные схемы, которые улучшают характеристики, но не приводят к существенному усложнению системы передачи.

### 5.6.2. Адаптивная дельта-модуляция

Известен ряд методов адаптивной дельта-модуляции (АДМ). Большинство этих методов основано на адаптации по выходу, когда шаг квантования перестраивается по выходной последовательности кодовых слов. Общий вид системы показан на рис. 5.36. Подобные схемы обладают тем преимуществом, что не требуют синхронизации по кодовым словам, поскольку при отсутствии ошибок шаг квантования как передатчика, так и приемника перестраивается в одной и той же кодовой последовательности.

В данном подпараграфе дается иллюстрация применения адаптивного квантования в дельта-модуляции на примере двух спе-

циальных алгоритмов адаптации. Имеется и множество других возможностей, которые можно найти в литературе [20—24].

Первая из рассматриваемых систем была подробно исследована Н. С. Джаянтом [22]. Алгоритм Джаянта для адаптивной

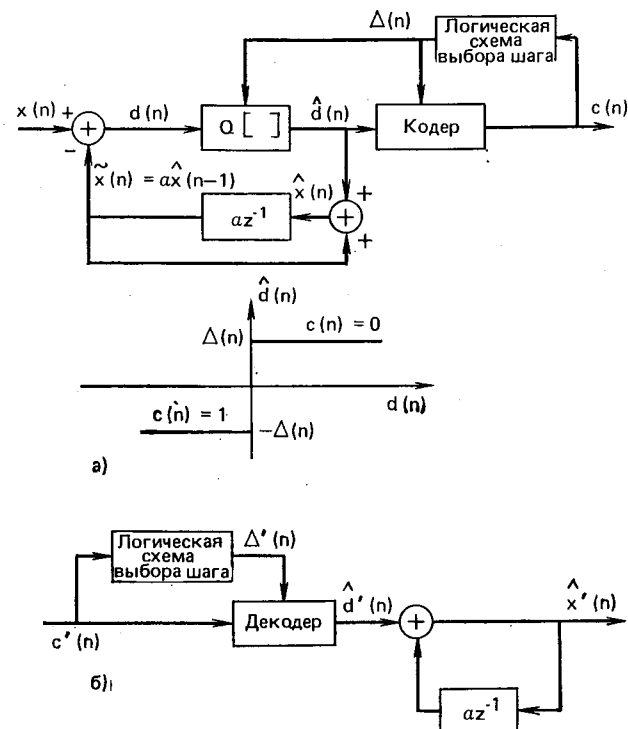


Рис. 5.36. Дельта-модулятор с адаптацией по шагу: а) кодер; б) декодер

дельта-модуляции представляет собой модификацию алгоритма, рассмотренного в 5.4.2. Как и в случае многоуровневого квантователя, шаг квантования подчиняется следующему правилу:

$$\Delta(n) = M \Delta(n-1); \quad (5.114a)$$

$$\Delta_{min} \leq \Delta(n) \leq \Delta_{max}. \quad (5.114b)$$

В том случае множитель является функцией текущего и предыдущего кодовых слов  $c(n)$  и  $c(n-1)$ . Это возможно, поскольку  $c(n)$  зависит только от знака  $d(n)$ , который задается соотношением

$$d(n) = x(n) - \alpha \hat{x}(n-1). \quad (5.115)$$

Таким образом, знак  $d(n)$  определяется перед получением квантованного значения  $\hat{d}(n)$ , которое возникает после вычисления

$\Delta(n)$  в соответствии с (5.114). Алгоритм выбора множителя шага квантования в уравнении (5.114а) имеет вид

$$\begin{cases} M = P > 1, & c(n) = c(n-1); \\ M = Q < 1, & c(n) \neq c(n-1). \end{cases} \quad (5.116)$$

Выбор такого метода адаптации объясняется видом последовательности кодовых слов, наблюдаемой в линейной дельта-модуляции. Например, из рис. 5.34а видно, что период перегрузки по крутизне соответствует отрезкам последовательности, состоящим только из нулей или только из единиц. Период шума дробления соответствует последовательности из чередующихся нулей и единиц вида 010101... На рис. 5.34б показано, как будет квантован сигнал, изображенный на рис. 5.34а с использованием адаптивного дельта-модулятора, описанного соотношениями (5.114) и (5.116).

Для удобства параметры системы в этом случае были приняты следующими:  $P=2$ ,  $Q=1/2$ ,  $a=1$ ; минимальный шаг квантования показан на рисунке. Можно отметить, что начальный участок области большой положительной крутизны порождает последовательность нулей, но в этом случае шаг квантования увеличивается экспоненциально, и это позволяет следить за увеличением крутизны входного сигнала. Области дробления в правой части рисунка вновь соответствует чередующаяся последовательность из нулей и единиц, но в этом случае шаг квантования быстро уменьшается до минимального ( $\Delta_{min}$ ) и остается таковым до тех пор, пока крутизна мала. Поскольку минимальный шаг квантования может быть сделан значительно меньше, чем тот, который необходим для оптимальной работы линейного дельта-модулятора, шум дробления может быть существенно уменьшен. Аналогично максимальный шаг квантования можно сделать большим, чем максимальная крутизна входного сигнала, что приведет к уменьшению шума перегрузки по крутизне.

Параметрами этой системы адаптивной дельта-модуляции являются:  $P$ ,  $Q$ ,  $\Delta_{min}$  и  $\Delta_{max}$ . Границы шага квантования следует выбирать таким образом, чтобы обеспечить необходимый динами-

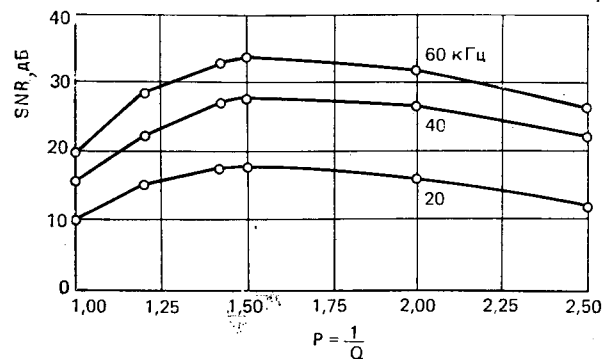


Рис. 5.37. SNR для адаптивного дельта-модулятора как функция  $P$  [22]

ческий диапазон входного сигнала. Отношение  $\Delta_{max}/\Delta_{min}$  должно быть достаточно большим, чтобы обеспечить большую величину отношения сигнал/шум в требуемом диапазоне уровней входного сигнала. Минимальный шаг квантования должен быть достаточно малым, чтобы минимизировать шум незанятого канала. Джаянт [22] показал, что  $P$  и  $Q$  должны удовлетворять соотношению

$$PQ \leq 1 \quad (5.117)$$

для устойчивости системы, т. е. для поддержания шага квантования таким, чтобы он соответствовал уровню входного сигнала. На рис. 5.37 представлены результаты моделирования на речевом сигнале с  $PQ=1$  для трех различных частот дискретизации. Видно, что значение отношения сигнал/шум достигает максимума при  $P=1,5$ , однако во всех трех случаях отношение сигнал/шум мало меняется при изменении  $P$  в пределах

$$1,25 < P < 2. \quad (5.118)$$

На рис. 5.38 для сравнения систем АДМ, ЛДМ и логарифмической ИКМ приведены зависимости отношения сигнал/шум от скорости передачи для всех трех случаев. Представленные на рисунке результаты для ЛДМ соответствуют случаю  $P=1/Q$  при дополнительном условии  $P=1=1/Q$ . Результаты для АДМ получены при  $P=1,5$ . В случае логарифмической ИКМ зависимость отношения сигнал/шум от скорости передачи вычислена в соответствии с соотношением (5.38) в предположении частоты дискретизации Найквиста ( $2F_N=6,6$  кГц) при  $\mu=100$ .

Рисунок 5.38 показывает, что при АДМ отношение сигнал/шум на 8 дБ выше, чем при ЛДМ (скорость передачи 20 кбит/с). Этот выигрыш достигает 14 дБ при скорости 60 кбит/с. С удвоением частоты дискретизации (скорости передачи) отношение сигнал/шум увеличивается на 6 дБ при ЛДМ, и на 10 дБ при АДМ. Сравнивая АДМ и логарифмическую ИКМ, отметим, что при скоростях меньше 40 кбит/с АДМ имеет лучшие характеристики, чем логарифмическая ИКМ. Для больших скоростей логарифмическая ИКМ имеет лучшее отношение сигнал/шум. Например, как следует из рис. 5.38, система с АДМ требует скорости, приблизительно равной 60 кбит/с для достижения того же качества, что и при се-

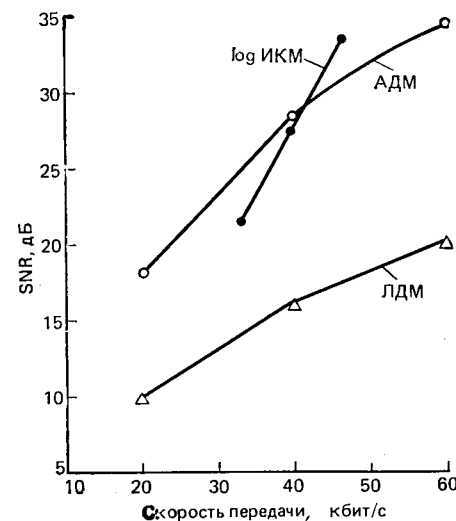


Рис. 5.38. Зависимость SNR от скорости передачи для трех схем кодирования с частотой дискретизации 6,6 кГц

миразрядной логарифмической ИКМ со скоростью передачи около 46 кбит/с.

Улучшение качества системы АДМ достигнуто путем ее незначительного усложнения. Поскольку адаптация осуществляется по выходному потоку двоичных символов, система АДМ сохраняет основное преимущество систем с дельта-модуляцией, т. е. не требует синхронизации по кодовым словам. Таким образом, во многих случаях целесообразно использовать АДМ вместо логарифмической ИКМ даже за счет незначительного увеличения скорости передачи.

Другим примером адаптивного квантования с дельта-модуляцией является дельта-модуляция с изменяющейся крутизной (ИКДМ). Эта система (впервые предложенная Гриффесом [23]) не отличается от системы, изображенной на рис. 5.35, но шаг квантования в этом случае изменяется в соответствии с уравнениями

$$\Delta(n) = \begin{cases} \beta \Delta(n-1) + D_2, & c(n) = c(n-1) = c(n-2); \\ \beta \Delta(n-1) + D_1, & \text{в противном случае,} \end{cases} \quad (5.119a)$$

где  $0 < \beta < 1$  и  $D_2 \gg D_1 > 0$ . В этом случае минимальный и максимальный шаги квантования определяются рекуррентным соотношением  $\Delta(n)$  (см. задачу 5.14).

Суть метода, как и раньше, состоит в том, чтобы увеличить шаг квантования при возникновении последовательности двоичных символов, свидетельствующей о перегрузке по крутизне. В случае возникновения трех последовательных символов единиц или нулей к шагу квантования добавляется приращение  $D_2$ . В отсутствие трех последовательных одинаковых символов шаг квантования уменьшается (так как  $\beta < 1$ ), пока не достигнет  $\Delta_{min}$ . Таким образом, шаг квантования увеличивается при перегрузке по крутизне и уменьшается при ее отсутствии. Величины  $\Delta_{min}$  и  $\Delta_{max}$  вновь выбираются из условия обеспечения требуемого динамического диапазона и малого шума дробления в условиях незагруженного канала. Параметр  $\beta$  определяет скорость адаптации. Если  $\beta$  близко к единице, то скорость нарастания и уменьшения  $\Delta(n)$  мала. С другой стороны, если  $\beta$  намного меньше единицы, то адаптация происходит быстрее. Таким образом, адаптация в данном случае может быть как слоговой, так и мгновенной.

Такая система может быть использована в случае, когда необходима малая чувствительность к ошибкам в канале и понижены требования к качеству речевого сигнала по сравнению с коммерческими каналами связи. При этом используется слоговая адаптация. Кроме того, коэффициент предсказания  $\alpha$  устанавливается значительно меньшим единицы, в результате чего влияние ошибок в канале существенно ослабляется. За нечувствительность к ошибкам в канале приходится расплачиваться понижением качества восприятия речи при их отсутствии. Основное достоинство системы АДМ в данном случае состоит в том, что она обладает достаточной гибкостью, позволяющей осуществлять обмен между качеством передачи и помехоустойчивостью.

### 5.6.3. Предсказание высокого порядка в дельта-модуляции

Для простоты в большинстве систем ЛДМ и АДМ используются предсказатели первого порядка вида

$$\tilde{x}(n) = \alpha \hat{x}(n-1), \quad (5.120)$$

как показано на рис. 5.36. В этом случае восстановленный сигнал удовлетворяет разностному уравнению

$$\hat{x}(n) = \alpha \hat{x}(n-1) + \hat{d}(n), \quad (5.121)$$

которое определяется передаточной функцией

$$H_1(z) = (1 - \alpha z^{-1})^{-1}. \quad (5.122)$$

Ранее отмечалось, что (5.121) соответствует цифровому интегратору (если  $\alpha=1$ ). Когда  $\alpha < 1$ , такое устройство иногда называют квазинтегратором.

Результаты рис. 5.32 показывают<sup>1</sup>, что в системе с дельта-модуляцией можно получить большее отношение сигнал/шум при использовании предсказателя второго порядка, для которого

$$\tilde{x}(n) = \alpha_1 \hat{x}(n-1) + \alpha_2 \hat{x}(n-2). \quad (5.123)$$

В этом случае

$$\hat{x}(n) = \alpha_1 \hat{x}(n-1) + \alpha_2 \hat{x}(n-2) + d(n), \quad (5.124)$$

что соответствует передаточной функции

$$H_2(z) = (1 - \alpha_1 z^{-1} - \alpha_2 z^{-2})^{-1}. \quad (5.125)$$

В [25] показано, что предсказатель второго порядка дает выигрыш по сравнению с предсказателем первого порядка, когда оба полюса  $H_2(z)$  действительны:

$$H_2(z) = \frac{1}{(1 - az^{-1})(1 - bz^{-1})}, \quad 0 < a, b < 1. \quad (5.126)$$

Такую систему часто называют системой с двойным интегрированием. Увеличение отношения сигнал/шум по сравнению с системой с одним интегратором может достигать 4 дБ в зависимости от диктора и речевого сигнала [25].

К сожалению, использование предсказателей более высокого порядка в системах с АДМ не является таким же простым делом, как замена предсказателя первого порядка предсказателем второго порядка, так как алгоритм адаптивного квантования связан с алгоритмом предсказания. Например, случаю незагруженного канала будут соответствовать различные последовательности двоичных символов в зависимости от порядка предсказателя. Для предсказателя второго порядка эта последовательность может быть

<sup>1</sup> Для более точного анализа необходимо знать значения автокорреляционной функции речевого сигнала на задержках, меньших чем 125 мкс (соответствующих высшей частоте дискретизации при дельта-модуляции), для вычисления коэффициента усиления.

010101 ... или 00110011 ... в зависимости от выбора  $\alpha_1$  и  $\alpha_2$  и последнего состояния системы перед тем, как сигнал на входе стал равен нулю. Это требует использовать алгоритм адаптации, основанный более чем на двух последовательных двоичных символах в случае, если шаг квантования достиг своего минимального значения для незанятого канала.

Построение систем с АДМ с предсказателем высокого порядка в настоящее время подробно не исследовано. Вопрос о целесообразности усложнения предсказателя и квантователя зависит от величины выигрыша в качестве передачи, которое может быть при этом получено. Использование многоуровневых квантователей, подобных рассмотренным в § 5.4, до некоторой степени упрощает разработку систем, но предполагает разделение двоичного потока на кодовые слова. Ниже рассматриваются методы разностного квантования с использованием многоуровневых квантователей.

### 5.7. Разностная ИКМ

Системы, аналогичные изображенной на рис. 5.31, будут далее называться системами с разностной ИКМ (РИКМ). Дельта-модулятор также можно называть одноразрядной системой с ИКМ. В общем случае, однако, термин «разностная ИКМ» применяется по отношению к системам, в которых квантователь имеет более двух уровней квантования.

Как следует из рис. 5.32, системы с РИКМ обеспечивают выигрыш от 4 до 11 дБ по сравнению с прямым квантованием (ИКМ). Наибольший выигрыш достигается при переходе от системы без предсказания к предсказателю первого порядка, несколько меньший — при увеличении порядка от одного до 4—5, после чего выигрыш незначителен. Как указывалось в § 5.5, это увеличение отношения сигнал/шум означает, что системы с РИКМ могут обеспечивать данное отношение сигнал/шум при разрядности, меньшей на единицу, чем это требовалось бы при прямом квантовании речевого сигнала. Следовательно, методы, изложенные в § 5.3 и 5.4, могут быть использованы для оценки качества, которое может быть достигнуто при применении обычного квантователя в разностной схеме. Например, для системы с разностной ИКМ и равномерным неадаптивным квантователем отношение сигнал/шум будет приблизительно на 6 дБ больше, чем для такого же квантователя при прямом квантовании входного сигнала. Разностная ИКМ будет обладать теми же свойствами, что и обычная ИКМ, т. е. отношение сигнал/шум будет увеличиваться на 6 дБ для каждого дополнительного разряда кодового слова, и будет зависеть также от уровня входного сигнала. Аналогично использование квантователя по  $\mu$ -закону в разностной схеме увеличит отношение сигнал/шум на 6 дБ, и в то же время ее характеристики будут нечувствительны к уровню входного сигнала.

На рис. 5.32 показаны изменения коэффициента усиления в зависимости от диктора и ширины полосы частот сигнала. Зна-

чительный разброс возникает при обработке различных фраз речевого сигнала, что является следствием нестационарности речи. Не существует единственного множества коэффициентов предсказателя, которые были бы оптимальными для различного речевого материала и различных дикторов.

Этот разброс совместно с изменениями уровня сигнала, которые характерны для систем связи, приводит к необходимости использования адаптивных предсказателей и адаптивных квантователей для получения наилучших характеристик при различных дикторах и в различных условиях. Такие системы называются системами адаптивной разностной ИКМ (АРИКМ).

#### 5.7.1. АРИКМ с адаптивным квантованием

Результаты анализа адаптивного квантования, изложенные в § 5.4, можно применить и в случае РИКМ. Имеется два основных способа управления адаптивным квантованием.

На рис. 5.39 показано применение квантователей с адаптацией по входу в системе АРИКМ [7]. Здесь шаг квантования пропорционален дисперсии сигнала на его входе. Однако, поскольку раз-

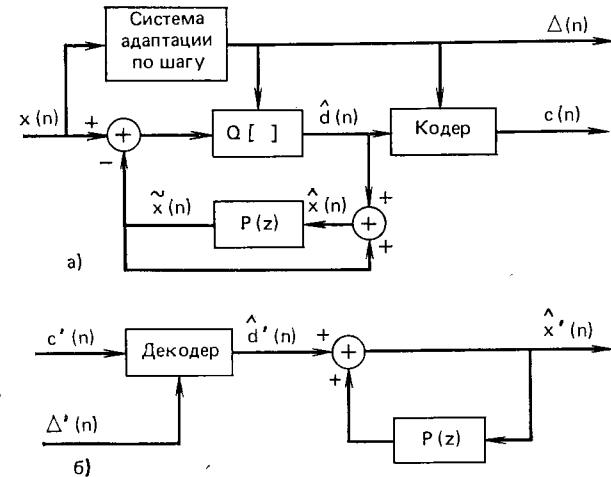


Рис. 5.39. Система АРИКМ с адаптивным по входу квантователем: а) кодер; б) декодер

ностный сигнал  $d(n)$  пропорционален входному сигналу, целесообразно управлять шагом квантования или по  $d(n)$ , или, как это показано на рис. 5.39, по входному сигналу  $x(n)$ . В 5.4.1 дано несколько алгоритмов управления шагом квантования. Как следует из результатов этого раздела, адаптивное квантование может обеспечить выигрыш около 5 дБ по сравнению со стандартной не-

адаптивной ИКМ  $\mu$ -законом квантования. Этот выигрыш совместно с дополнительным увеличением отношения сигнал/шум 6 дБ, которое можно получить при применении разностной схемы с неадаптивным квантованием, означает, что АРИКМ с адаптацией по входу позволит получить отношение сигнал/шум на 10—11 дБ больше, чем при использовании неадаптивного квантователя с тем же числом уровней.

На рис. 5.40 показано использование квантователя с адаптацией по выходу в системе АРИКМ [26]. Если, например, адаптация осуществляется в соответствии с уравнениями (5.66)—(5.68),

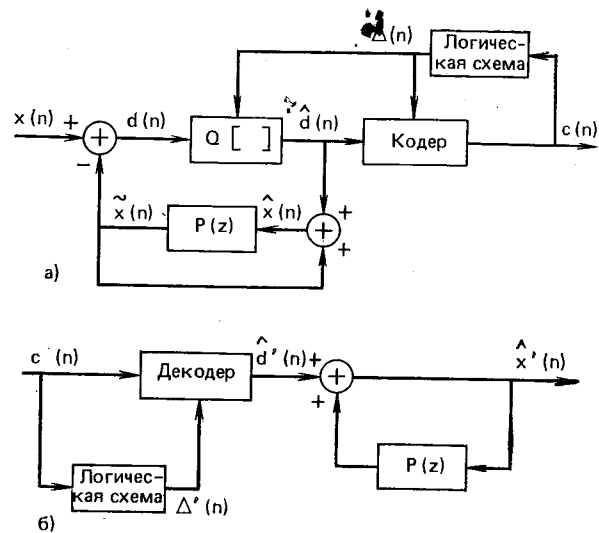


Рис. 5.40. Система АРИКМ с адаптивным по выходу квантователем: а) кодер; б) декодер

можно ожидать выигрыш 4—6 дБ по сравнению с неадаптивным квантованием по  $\mu$ -закону с тем же числом уровней. Таким образом, адаптация как по выходу, так и по входу позволит достигнуть выигрыша, равного 10—12 дБ по сравнению с неадаптивным квантованием с тем же числом уровней. Кроме того, адаптивный квантователь позволяет расширить динамический диапазон. Дополнительным преимуществом адаптации по выходу является то, что не требуется передавать дополнительную информацию о шаге квантования. Это, однако, делает восстановленный сигнал более чувствительным к ошибкам в канале связи. При адаптации по входу кодовые слова и шаг квантования представляют собой описание сигнала. Хотя это увеличивает сложность представления, однако появляется возможность передачи шага квантования с защитой его от ошибок, что позволяет существенно улучшить качество восстановленного сигнала [27, 28].

### 5.7.2. АРИКМ с адаптивным предсказанием

Выше рассматривались системы с неадаптивным предсказателем и было выяснено, что даже при использовании предсказателей высокого порядка можно ожидать, что разностное квантование при благоприятных условиях даст выигрыш около 10—12 дБ. Кроме того, величина выигрыша зависит от диктора и речевого материала. Учитывая нестандартность речевого сигнала, естественно рассмотреть адаптивный предсказатель, который, как и адаптивный квантователь, следит за текущими изменениями в речевом сигнале [29]. Система АРИКМ, содержащая адаптивный квантователь и адаптивный предсказатель, изображена на рис. 5.41.

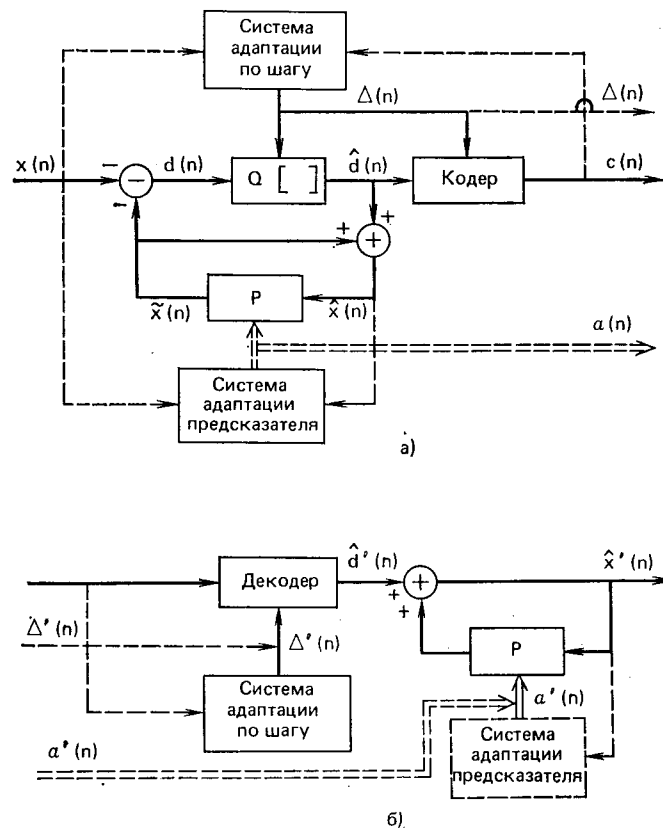


Рис. 5.41. Система АРИКМ с адаптивным квантователем и адаптивным предсказанием: а) кодер; б) декодер

Пунктирные линии показывают, что адаптация предсказателя и квантователя может осуществляться по входному и выходному сигналам. В первом случае к последовательности кодовых слов  $c(n)$  для полного описания речевого сигнала необходимо доба-

вить  $\Delta(n)$  или коэффициенты предсказания  $a(n) = \{a_k(n)\}$ , (или и то и другое).

Предполагается, что коэффициенты предсказания зависят от времени так, что предсказанное значение имеет вид

$$\tilde{x}(n) = \sum_{k=1}^p \alpha_k(n) \hat{x}(n-k). \quad (5.127)$$

При адаптации коэффициентов предсказателя  $a(n)$  обычно предполагается, что свойства речевого сигнала не меняются в течение короткого интервала времени. Коэффициенты предсказания выбираются, следовательно, так, чтобы минимизировать средний квадрат погрешности предсказания на коротком интервале времени. При адаптации по входу предсказатель адаптируется по измерениям входного сигнала (это справедливо, если в соотношениях § 5.5 пренебречь членом  $1/SNR$ ). Применяя те же выкладки, которые использовались при выводе (5.87) и (5.89), и пренебрегая влиянием ошибок квантования, можно показать, что оптимальные коэффициенты предсказателя удовлетворяют уравнениям

$$R_n(j) = \sum_{k=1}^p \alpha_k(n) R_n(j-k), \quad j=1, 2, \dots, p, \quad (5.128)$$

где  $R_n(j)$  — кратковременная автокорреляционная функция [соотношение (4.24)]

$$R_n(j) = \sum_{m=-\infty}^{\infty} x(m) w(n-m) x(j+m) w(n-m-j), \quad 0 \leq j \leq p \quad (5.129)$$

и  $w(n-m)$  — взвешивающая функция (временное окно). Можно использовать прямоугольное окно или другие подходящие окна (например, окно Хемминга длиной  $N$ ). Поскольку параметры речи меняются относительно медленно, целесообразно подстраивать параметры предсказателя  $a(n)$  периодически. Например, новую оценку можно вычислять через каждые 10—20 мс, полагая, что на этих интервалах она остается постоянной. Длительность окна может быть равна этому интервалу или быть несколько больше. В последнем случае соседние сегменты речевого сигнала будут пересекаться. Как следует из (5.129), при вычислении оценок корреляционной функции в (5.128) предполагается, что перед вычислением  $R_n(j)$  необходимо записать  $N$  отсчетов  $x(n)$  в буфер. Множество параметров  $a(n)$ , удовлетворяющих (5.128), используется в схеме рис. 5.41а для того, чтобы квантовать входной сигнал в течение интервала  $N$  отсчетов, начиная с  $n$ . Таким образом, для восстановления входного сигнала по последовательности кодовых слов необходимо знать коэффициенты предсказания (и возможно шаг квантования), как это показано на рис. 5.41б. Особенности вычисления изменяющихся во времени параметров предсказания изучаются в гл. 8.

Для оценки преимуществ адаптивного предсказания Нолл [7] исследовал зависимость коэффициента качества предсказания  $G_p$  от порядка предсказателя для фиксированного и адаптивного слушателей. На рис. 5.42 приведена зависимость величины<sup>1</sup>

$$10 \log_{10} [G_p] = 10 \log_{10} \left[ \frac{E[x^2(n)]}{E[d^2(n)]} \right] \quad (5.130)$$

от порядка предсказателя  $p$  для случаев адаптивного и неадаптивного предсказания. Нижняя кривая, полученная вычислением автокорреляционной функции с усреднением по всей фразе с после-

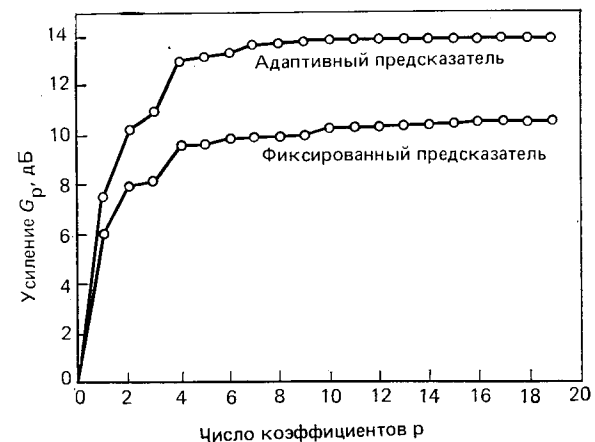


Рис. 5.42. Зависимость показателя качества предсказания от числа коэффициентов предсказания для женского голоса (полоса частот 0—32 000 Гц) [7]

дующим решением системы уравнений (5.89), показывает, что максимальное значение коэффициента качества предсказания составляет около 10,5 дБ. Верхняя кривая получена путем определения такой длины окна  $L$  и таких значений коэффициентов предсказания  $a(n)$ , которые максимизируют  $G_p$  на всей входной фразе при данном значении  $p$ . Эта максимальная величина изображена для каждого  $p$ . Максимальное значение коэффициента качества предсказания при этом составляет около 14 дБ. Таким образом, Нолл [7] предположил, что приемлемая верхняя граница выигрыша АРИКМ систем с неадаптивным и адаптивным предсказателями составляет 10,5 и 14 дБ соответственно. Кривые рис. 5.42, однако, не отражают того обстоятельства, что неадаптивный оптимальный предсказатель весьма чувствителен как к диктору, так и к речевому материалу, в отличие от адаптивного предсказателя, чувствительность которого заметно меньше.

При адаптивном предсказании устраняется избыточность речевого сигнала. Если возможно качественное предсказание, то погрешность предсказания  $d(n)$  будет совершенно некоррелированной (белым шумом). Можно сказать, что избыточность устраняет-

<sup>1</sup> На данном рисунке представлены результаты, полученные для одного диктора. Кроме того, ошибки в петле обратной связи не учитывались.

ся в строгом соответствии с моделью речевого сигнала. Как следует из рис. 5.42, коэффициент усиления незначительно увеличивается при увеличении порядка предсказания от 4 до 5. Такое предсказание, однако, не учитывает важного источника избыточности в речевом сигнале, а именно корреляцию вследствие квазипериодического характера вокализованной речи. Один из подходов к учету этой корреляции предложен Аталом и Шредером [29], которые использовали предсказатель вида

$$\tilde{x}(n) = \beta \hat{x}(n-M) + \sum_{k=1}^p \alpha_k [\hat{x}(n-k) - \beta \hat{x}(n-k-M)], \quad (5.131)$$

где параметры предсказания  $\beta$ ,  $M$  и  $\{\alpha_k\}$  адаптируются на интервале, равном  $N$  отсчетов. Пренебрегая эффектами погрешности квантования в  $\hat{x}(n)$ , можем записать ошибку предсказания в виде

$$d(n) = x(n) - \tilde{x}(n) = x(n) - \beta x(n-M) - \sum_{k=1}^p \alpha_k [x(n-k) - \beta x(n-k-M)], \quad (5.132)$$

которая может быть выражена как

$$d(n) = v(n) - \sum_{k=1}^p \alpha_k v(n-k). \quad (5.133)$$

Здесь

$$v(n) = x(n) - \beta x(n-M). \quad (5.134)$$

Непосредственное вычисление значений  $\beta$ ,  $M$  и  $\{\alpha_k\}$ , которые минимизируют  $d(n)$ , затруднительно. Поэтому Атал и Шредер [29] предложили субоптимальное решение, в котором сначала минимизируется дисперсия  $v(n)$ , а затем — дисперсия  $d(n)$  при заданных  $\beta$  и  $M$ . Таким образом, в данном случае автокорреляционная функция вычисляется так же, как и раньше, но для задержек, попадающих в область периода основного тона. Коэффициент предсказания  $\beta$  выбирается совпадающим со значением пика нормированной автокорреляционной функции, а  $M$  представляет собой положение этого пика в  $R_n(j)$ . Таким образом, значение  $\beta$  учитывает изменение амплитуды между последовательными периодами, в то время как  $M$  — это период основного тона (в числе отсчетов). По известным  $M$  и  $\beta$  вычисляется последовательность  $v(n)$  и определяется ее корреляционная функция для  $j = 0, 1, \dots, p$ , по которой можно определить коэффициенты предсказания из уравнения (5.128), где  $R_n(j)$  является кратковременной автокорреляционной функцией последовательности  $v(n)$ .

Для представления речевого сигнала на основе подобного метода необходимо передавать или хранить квантованный разност-

ный сигнал, шаг квантования (если осуществляется адаптация по входу) и коэффициенты предсказания (квантованные). В работе Атала и Шредера использовался одноразрядный квантователь для разностного сигнала и шаг квантования изменялся каждые 5 мс (33 отсчета на частоте 6,67 кГц) для минимизации ошибки квантования. Кроме того, через каждые 5 мс оценивались и параметры предсказателя. Хотя в работе не приводятся данные по отношению сигнал/шум, предполагается, что можно получить высокое качество восстановленного сигнала на скоростях порядка 10 кбит/с. Используя соответствующее квантование параметров и коэффициентов, Джаянт [8] утверждает, что можно получить выигрыш, равный 20 дБ, по сравнению с ИКМ.

К сожалению, до настоящего времени не проведены исследования потенциальных возможностей адаптивного предсказания, включая параметры основного тона. Однако, несомненно, что методы, подобные приведенным выше являются наиболее сложными среди цифровых методов кодирования речевого колебания. Антиподом АРИКМ может служить линейная дельта-модуляция с ее простым процессом квантования и однородным потоком однозарядных двоичных кодовых слов. Выбор схемы квантования зависит от различных факторов, в том числе от требуемой скорости передачи, качества передачи, сложности кодера и цифрового представления. В следующем параграфе собраны результаты нескольких сравнительных исследований, позволяющие классифицировать различные методы квантования. Но перед этим кратко рассмотрим вопросы управления по выходу в адаптивном предсказании.

Один из подходов основан на вычислении корреляционной функции квантованного сигнала. Таким образом, в (5.128)  $R_n(j)$  заменяется на

$$\sum_{m=-\infty}^{\infty} \hat{x}(m) \hat{w}(n-m) \hat{x}(m+j) \hat{w}(n-m-j), \quad 0 \leq j \leq p. \quad (5.135)$$

В этом случае окно должно иметь вид

$$w(m) = \begin{cases} 1, & 0 \leq m \leq N-1; \\ 0, & \text{в противном случае,} \end{cases} \quad (5.136)$$

т. е. оценки коэффициентов предсказания должны основываться на прошлых квантованных отсчетах, а не на будущих, которые не могут быть получены до коэффициентов предсказания. Как и в случае адаптации квантователя, в данном случае по каналу связи передаются только последовательность кодовых слов. Но предсказание с управлением по выходу не получило широкого распространения из-за присущей ему чувствительности к ошибкам и худших характеристик, обусловленных использованием для управления искаженного входного сигнала. Интересный подход к управлению по выходу рассмотрен Стрехом [30], исследовавшим градиентные методы для подстройки коэффициентов предсказателя.

## 5.8. Сравнение систем

При сравнении цифровых систем кодирования речевого колебания в качестве критерия достаточно использовать отношение сигнал/шум. Однако в конечном счете качество систем передачи речевого сигнала необходимо оценивать по слуховому восприятию. Вопрос о том, насколько хорошо звучит кодированный сигнал по сравнению с исходным, является вопросом большой важности. К сожалению, субъективное качество восприятия, как правило, не поддается количественным оценкам и законченные результаты в этом направлении отсутствуют. Поэтому в данном параграфе содержится обзор и сравнение ряда систем кодирования речевого сигнала по отношению сигнал/шум квантования, а также приводится ряд субъективных оценок, позволяющих сопоставить полученные результаты.

Нолл [7] приводит подробное сравнительное исследование схем кодирования речи. Он рассматривал следующие системы:

1) неадаптивную ИКМ с логарифмическим компандированием и с  $\mu=100$ ,  $X_{max}=8\sigma_x$  (ИКМ);

2) адаптивную ИКМ (оптимальный гауссовский квантователь с управлением по входному сигналу (ИКМ-АК<sub>вх</sub>);

3) разностную ИКМ с предсказателем первого порядка (неадаптивным) и адаптивным «гауссовским квантователем» с управлением по выходному сигналу (РИКМ-АК<sub>вых</sub>);

4) адаптивную разностную ИКМ с адаптивным предсказателем первого порядка и адаптивным гауссовским квантователем, управляемыми по входному сигналу (АРИКМ1-АК<sub>вх</sub>) (протяженность окна 32);

5) адаптивную разностную ИКМ с предсказателем четвертого порядка и адаптивным «квантователем Лапласа», адаптирующимися по входному сигналу (АРИКМ4-АК<sub>вых</sub>) (протяженность окна 128);

6) адаптивную разностную ИКМ с адаптивным предсказателем 12-го порядка и адаптивным «гамма-квантователем» с адаптацией по входному сигналу (АРИКМ12-АК<sub>вх</sub>) (протяженность окна 256).

Во всех системах частота дискретизации составляла 8 кГц, а длина кодового слова изменялась от 2 до 5 бит на отсчет, т. е. скорость изменялась от 16 до 40 кбит/с. Отношение сигнал/(шум квантования) для всех систем приведено на рис. 5.43. Представленные кривые позволяют отметить ряд интересных особенностей. Во-первых, нижняя кривая соответствует квантователю с кодовыми словами длиной 2 бит. По мере увеличения длины кодового слова кривые поднимаются вверх и отстоят одна от другой на 6 дБ. Заметим также, что добавление как неадаптивного предсказателя, так и адаптивного квантователя позволяет получить выигрыш в отношении сигнал/шум, но адаптация простого предсказателя практически не приводит к выигрышу. Однако видно, что

адаптивные предсказатели более высокого порядка приводят к существенному выигрышу в отношении сигнал/шум.

При телефонной передаче обычно предполагают, что приемлемое качество речи обеспечивается квантователем по  $\mu$ -закону при 6–7 бит на отсчет. Из (5.38) видно, что семиразрядный квантова-

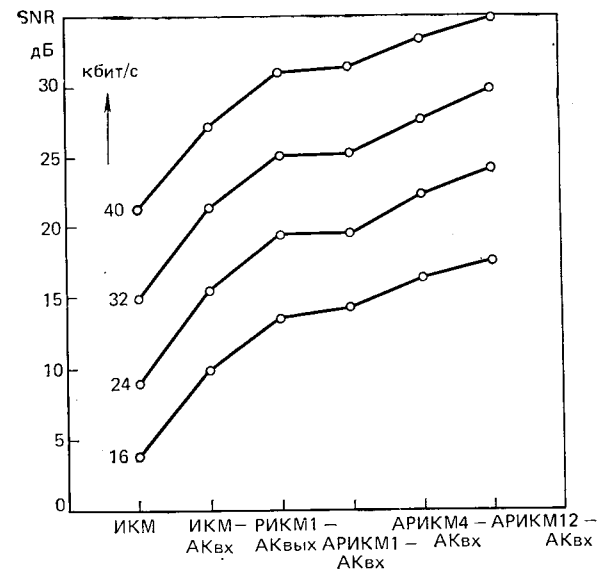


Рис. 5.43. Отношение сигнал/шум для квантования от 2 до 5 бит на отсчет (16–40 кбит/с): АК<sub>вх</sub> — адаптивный квантователь с адаптацией по входу; АК<sub>вых</sub> — адаптивный квантователь с адаптацией по выходу; АРИКМ<sub>r</sub> — система АРИКМ с порядком предсказателя r [7]

тель при  $\mu=100$  дает отношение сигнал/шум около 33 дБ. На основе рис. 5.43 можно сделать вывод, что такое же качество обеспечивает пятиразрядный адаптивный квантователь и адаптивный предсказатель. На практике субъективное качество речевого сигнала в системах с АРИКМ оказывается лучше, чем с ИКМ при том же отношении сигнал/шум. Исследуя системы с АРИКМ и адаптацией квантователя по выходу и неадаптивным предсказателем, в [26] установлено, что слушатели отдают предпочтение АРИКМ сигналу перед сигналом логарифмической ИКМ с большим отношением сигнал/шум. Результаты теста представлены в табл. 5.7, где ИКМ — это система 1 в исследовании Нолла, а АРИКМ — это система 3 (см. выше). Из представленных результатов следует, что четырехразрядная АРИКМ оказывается лучше шестиразрядной ИКМ. Этот результат не будет казаться удивительным, если вспомнить, что выигрыш в отношении сигнал/шум при использовании АРИКМ с неадаптивным предсказателем и адаптивным квантователем составляет 10–12 дБ или, грубо говоря, двухразрядной, но фактически четырехразрядной АРИКМ отдают предпочтение перед шестиразрядной ИКМ, несмотря на несколько меньшее отношение сигнал/шум.

Исследуя адаптивное предсказание, Атал и Шредер [29] установили, что их система АРИКМ с двухуровневым адаптивным



Таблица 5.7

## Сравнение объективного и субъективного качества системы АРИКМ и лог-ИКМ

Объективная оценка (по SNR)	Субъективная оценка (по предпочтению)
7 бит ИКМ	7 бит ИКМ (высокое качество)
6 бит ИКМ	4 бит АРИКМ
4 бит АРИКМ	6 бит ИКМ
5 бит ИКМ	3 бит АРИКМ
3 бит АРИКМ	5 бит ИКМ
4 бит ИКМ	4 бит ИКМ (низкое качество)

квантователем и сложным адаптивным предсказателем позволяет получить сигнал, не намного худший, чем при шестизрядной логарифмической адаптивной ИКМ. Оцененная скорость передачи в этой системе составляла 10 кбит/с, а для ИКМ в тех же условиях, т. е. при частоте дискретизации 6,67 кГц, требуется скорость 40 кбит/с. В этом случае особенно явно сказывается различие между субъективным качеством и тем, которое можно ожидать из анализа величины отношения сигнал/шум.

Дать точное объяснение этому эффекту затруднительно, но можно предположить, что такое различие возникает вследствие влияния двух факторов: лучших характеристик незанятого канала для адаптивного квантователя и большей корреляции между шумом квантования и сигналом [7].

### 5.9. Преобразования способов кодирования

Из материала данной главы достаточно ясно, что существует множество возможностей для квантования речевого сигнала. Эти методы различаются по своей сложности — от чрезвычайно простой в технической реализации линейной дельта-модуляции, требующей больших скоростей передачи, до разнообразных методов адаптивной разностной ИКМ, обеспечивающих хорошее качество на малых скоростях передачи, но при большей сложности алгоритмов обработки. В результате значительный интерес представляет разработка методов прямого преобразования одного цифрового представления в другое, минуя аналоговое представление сигнала. Эта задача важна по следующим причинам.

1. В больших системах связи весьма вероятно возникновение ситуаций, в которых на районных сетях более важно иметь дешевое оконечное оборудование, а не низкую скорость передачи информации. В других случаях, например при междугородной связи или при хранении речевого сигнала в цифровом блоке памяти, важнее получить сокращенное описание сигнала. Для объединения различных частей системы связи с различными скоростями чрезвычайно важно располагать возможностью прямого преобразования цифровых представлений, используемых в каждой из

подсистем, что позволит избежать дополнительных потерь качества.

2. Реализация низкоскоростных представлений сигнала существенно упрощается при применении простых методов цифрового преобразования. Например, может оказаться полезным использовать линейную дельта-модуляцию, скажем, для аналого-цифрового преобразования, чтобы далее представить полученный сигнал в более компактной форме, такой, как ИКМ или АРИКМ.

3. При обработке речевого сигнала требуется представить его в цифровой форме с помощью линейной ИКМ для того, чтобы далее, например при цифровой фильтрации, можно было использовать отдельные отсчеты сигнала.

#### 5.9.1. Преобразование ЛДМ в ИКМ

Для получения высококачественного цифрового представления с помощью линейной дельта-модуляции требуется очень большая частота дискретизации и простой двухуровневый квантователь. Такие системы могут быть просто реализованы на основе совместного использования аналоговых и цифровых компонент. Действительно, весь кодер ЛДМ можно легко реализовать как простую интегрирующую цепь [31]. На рис. 5.44 показаны обычный анало-

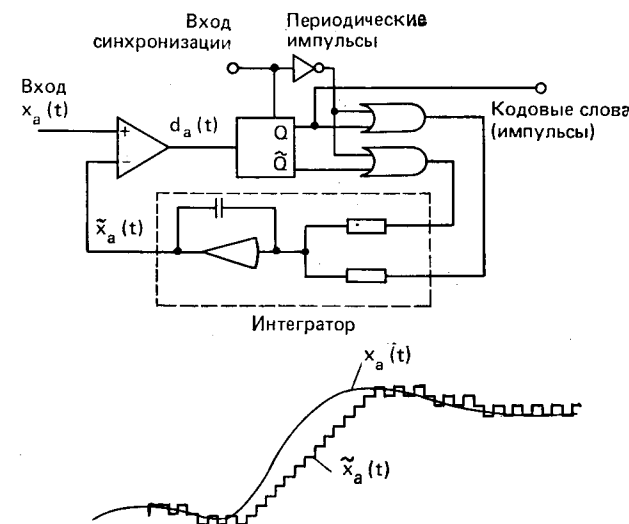


Рис. 5.44. Схема линейного дельта-модулятора [31]

говый компаратор, формирующий разностный сигнал, триггер, определяющий знак этого сигнала, и интегратор, восстанавливающий предсказанное значение для сравнения его с входным сигналом. Это простое объединение аналоговых и цифровых цепей — все, что необходимо для реализации линейного дельта-модулято-

ра. Данное устройство реализовано в виде интегральной схемы, работающей на частотах до 17 МГц [31].

Простота схемы совместно с возможностью работы на весьма высоких скоростях делает это устройство весьма удобным для использования при аналого-цифровых преобразованиях. Простота достигается, конечно, за счет чрезвычайно высокой скорости, которая необходима для получения хорошего качества передачи.

Скорость передачи, однако, можно снизить за счет использования преобразования линейной дельта-модуляции в более эффективные виды цифрового представления, такие, как ИКМ или АРИКМ. Наиболее важным является преобразование ЛДМ в линейную ИКМ, поскольку ИКМ представление требуется при любой цифровой обработке отсчетов аналогового сигнала.

Процесс преобразования ЛДМ в ИКМ включает в себя, во-первых, получение ИКМ представления и, во-вторых, снижение частоты дискретизации до частоты Найквиста. Первый шаг заключается в декодировании последовательности нулей и единиц в последовательность величин  $\pm \Delta$ , а затем в цифровом интегрировании положительных и отрицательных приращений для получения квантованных отсчетов  $x_a(t)$  с частотой дискретизации, соответствующей ЛДМ. Полученная последовательность содержит шумы квантования в полосе  $|\Omega| \leq \pi/T$ , где  $T$  — период дискретизации в ЛДМ, хотя спектр входного сигнала ограничен по полосе значительно более низкой частотой. Таким образом, перед уменьшением частоты дискретизации необходимо устранить шум квантования в полосе от максимальной частоты спектра сигнала до половины частоты дискретизации ЛДМ. Как обсуждалось в 2.4.2, это можно сделать весьма эффективно при использовании КИХ-фильтра нижних частот с частотой среза, равной частоте Найквиста обрабатываемого сигнала [32]. Выходной сигнал фильтра вычисляется для каждых  $M$  отсчетов, где  $M$  — отношение частоты ЛДМ дискретизации к частоте ИКМ дискретизации. Таким образом, ЛДМ-ИКМ-преобразователь содержит интегратор, выход которого фильтру-

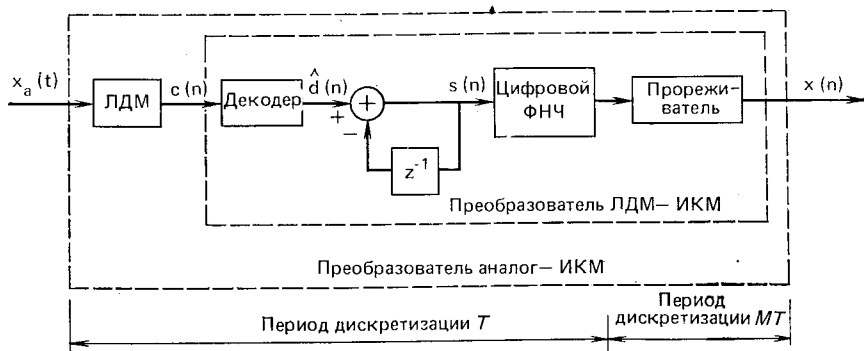


Рис. 5.45. Преобразователь аналог — ИКМ в системе ЛДМ и преобразователь ЛДМ — ИКМ

ется и подвергается дискретизации. Как следует из рис. 5.45, преобразователь ЛДМ-ИКМ совместно с кодером ЛДМ представляет собой устройство аналого-цифрового преобразования, почти полностью выполненное в цифровой форме.

### 5.9.2. Преобразование ИКМ — АРИКМ

Другой пример преобразования — это преобразование линейной ИКМ в АРИКМ [33]. Интерес к этому виду преобразования определяется желанием получить более эффективное представление сигнала, чем то, которое дает линейная ИКМ. На рис. 5.46 показана структурная схема данного преобразования. Очевидно,

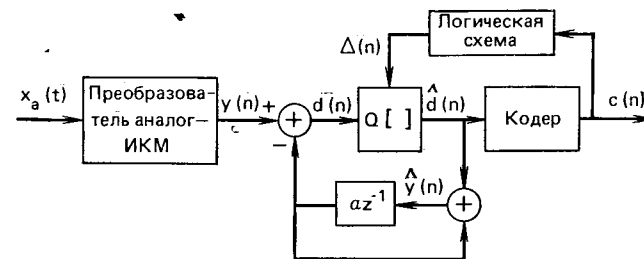


Рис. 5.46. Преобразователь ИКМ — АРИКМ

что основной принцип состоит в реализации обычного алгоритма АРИКМ, применяемого к последовательности на выходе кодера линейной ИКМ. Это достигается путем реализации операций, изображенных на рис. 5.46. Важный вопрос, возникающий при любых цифровых преобразованиях, состоит в ухудшении качества, обусловленном последующей цифровой обработкой. Очевидно, что входной ИКМ сигнал будет содержать погрешность, которую можно охарактеризовать отношением сигнал/шум  $SNR_1$ . Алгоритм АРИКМ внесет дополнительные искажения. Если предположить, что эти искажения не зависят от обрабатываемого сигнала и ошибок, вносимых ИКМ, то можно считать, что общие искажения приближенно равны (см. задачу 5.17)

$$SNR = SNR_1 / (1 + SNR_1 / SNR_2), \quad (5.137)$$

где  $SNR_2$  — отношение сигнал/шум системы с АРИКМ. Из данного соотношения видно, что полное  $SNR$  не может быть больше, чем  $SNR_1$ . Однако если  $SNR_2$  порядка  $SNR_1$ , то потери малы. Действительно, если  $SNR_1 = SNR_2$ , они составляют около 3 дБ.

Кодопреобразование можно рассматривать как точный и эффективный метод применения обычных алгоритмов кодирования речи. Так изначально получают линейную ИКМ, например так, как это описано в 5.9.1, а затем осуществляют цифровую обработку с использованием различных методов кодирования сигнала.

## 5.10. Заключение

В данной главе подробно изложены методы цифрового представления речевого колебания. Показано, что существуют несколько подходов к решению этой задачи. Главное внимание уделено основным принципам, а не описанию этих систем такого типа, предложенных к настоящему времени. Дальнейшее изучение этой темы можно провести по обзорной статье Джаянта [8] и по сборнику статей [34] под его редакцией.

### Задачи

5.1. Функция плотности вероятности равномерного закона распределения

$$p(x) = \begin{cases} 1/\Delta, & |x| < \Delta/2; \\ 0, & \text{в противном случае.} \end{cases}$$

Определить среднее значение и дисперсию равномерного распределения.

5.2. Рассмотреть функцию плотности вероятности распределения Лапласа

$$p(x) = \frac{1}{\sqrt{2}\sigma_x} e^{-\sqrt{2}|x|/\sigma_x} \text{ и определить вероятность того, что } |x| > 4\sigma_x.$$

5.3. Пусть  $x(n)$  — сигнал на входе линейной системы, инвариантной к сдвигу, — представляет собой стационарный белый шум с нулевым средним значением и единичной дисперсией. Показать, что автокорреляционная функция процесса на выходе имеет вид  $\varphi(m) = \sigma_x^2 \sum_{k=-\infty}^{\infty} h(k)h(k+m)$ , где  $\sigma_x^2$  — дисперсия входного сигнала и  $h(n)$  — импульсная характеристика линейной системы.

5.4. Рассмотреть вопросы разработки высококачественной цифровой акустической системы. Отношение сигнал/шум, равное 60 дБ, необходимо обеспечить для пиковых уровней сигнала в диапазоне от 1 до 100. Полезная полоса частот сигнала должна быть не менее 8 кГц.

Требуется:

а) Изобразить основные узлы аналого-цифрового и цифро-аналогового преобразователей.

б) Определить количество разрядов при аналого-цифровом преобразовании.

в) Что является главным условием при выборе частоты дискретизации? Какого типа фильтры необходимо использовать перед аналого-цифровым и после цифро-аналогового преобразований? Оценить наименьшую частоту дискретизации, которая возможна в практическом случае.

г) Как изменяются предъявляемые к системе требования и выбранные параметры, если необходимо обеспечить лишь «телефонное» качество речевого сигнала?

5.5. Речевой сигнал ограничен по полосе идеальным фильтром нижних частот, дискретизирован с частотой Найквиста, квантован в  $B$ -разрядном квантователе и затем преобразован обратно в аналоговую форму в идеальном цифро-аналоговом преобразователе (рис. 3.5.1а). Определим  $y(n) = x(n) + e_1(n)$ , где  $e_1(n)$  — погрешность квантования. Предположим, что шаг квантования  $\Delta = 8\sigma_x/2^B$  и  $B$  достаточно велики, чтобы выполнялись следующие предположения:  $e_1(n)$  — стационарная последовательность;  $e_1(n)$  — не коррелирована с  $x(n)$ ;  $e_1(n)$  — равномерно распределенная последовательность белого шума. Как было показано, отношение сигнал/шум (шум квантования) в этих условиях равно  $SNR_1 = \sigma_x^2/\sigma_{e_1}^2 = (12/64) \cdot 2^{2B}$ . Предположим, что аналоговый сигнал дискретизирован с частотой Найквиста и квантован в  $B$ -разрядном квантователе (рис. 3.5.1б) (предположим также, что  $0 < \epsilon < T$ , т. е. устройства работают не абсолютно синхронно во времени). При этом  $w(n) = y'(n) + e_2(n)$ , где  $e_2(n)$  обладает теми же свойствами, что и  $e_1(n)$ .

а) Показать, что полное отношение сигнал/шум равно  $SNR_2 = SNR_1/2$ .

б) Обобщить результат п.а) для  $N$ -кратного аналого-цифрового преобразования и обратно.

5.6. Хотя обычно предполагают, что погрешность квантования не зависит от сигнала, это предположение нарушается при малом числе уровней квантования.

а) Показать, что  $e(n) = x(n) - \hat{x}(n)$  не является статистически независимым от  $x(n)$  ( $\hat{x}(n)$  — квантованный сигнал). Указание: представить  $\hat{x}(n)$  как  $\hat{x}(n) = [x(n)/\Delta]\Delta + (\Delta/2)$ , где  $[\cdot]$  означает целую часть, т. е. наибольшее целое число.

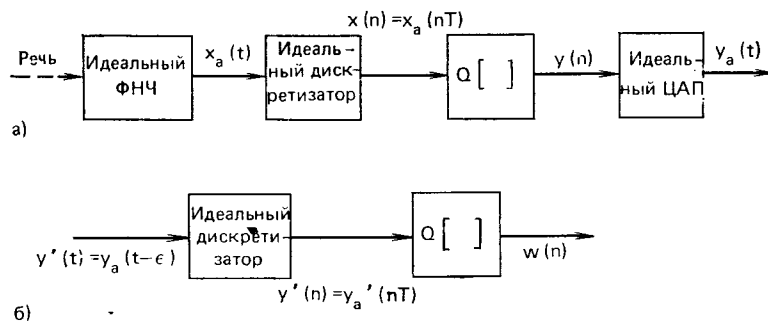


Рис. 3.5.1

меньше или равное величине, указанной в скобках. Процесс представить в виде  $x(n) = [x(n)/\Delta]\Delta + x_f(n) = x_i(n) + x_f(n)$ , где  $x_i(n)$  — целая часть  $x(n)$ , а  $x_f(n)$  остаток от  $x(n)$ . Таким образом,  $e(n)$  можно представить как функцию от  $x(n)$ . Показать, что они не могут быть статистически независимыми.

б) При каких условиях справедливо приближение, что  $x(n)$  и  $e(n)$  независимы?

в) На рис. 3.5.2 представлен метод, позволяющий сделать  $e(n)$  и  $x(n)$  статистически независимыми даже при малом числе уровней квантования. В этом случае  $z(n)$  представляет собой псевдослучайную реализацию типа белого шума с равномерным распределением и функцией плотности вероятности, равной  $p(z) = 1/\Delta$ ,  $-\Delta/2 \leq z \leq \Delta/2$ . Показать, что при этом погрешность квантования статистически не зависит от сигнала при любых значениях  $B$  (последовательность шума называют возмущающей). Указание: рассмотреть последовательность  $e(n)$  для  $y(n)$ .

г) Показать, что дисперсия ошибки при возмущении возрастает и  $\sigma_{e_1}^2 > \sigma_{e_2}^2$ , где  $e_1(n) = x(n) - \hat{y}(n)$  и  $e_2(n) = x(n) - \hat{x}(n)$ .

д) Показать, что путем простого вычитания возмущающего шума из выходной последовательности квантователя можно получить такую же ошибку квантования  $e_2(n) = x(n) - (\hat{y}(n) - z(n))$ , как и при отсутствии возмущений  $\sigma_{e_2}^2 = \sigma_{e_1}^2$ .

5.7. Обычно дисперсию сигнала оценивают, полагая, что она пропорциональна кратковременной энергии, равной  $\sigma^2(n) = \sum_{m=-\infty}^{\infty} x^2(m)h(n-m)$ .

а) Показать, что если  $x(n)$  — стационарный случайный процесс с нулевым средним, то  $E[\sigma^2(n)]$  пропорциональна  $\sigma_x^2$ .

б) Для

$$h(n) = \begin{cases} \alpha^n, & n \geq 0, \quad (|\alpha| < 1); \\ 0, & n < 0 \end{cases}$$

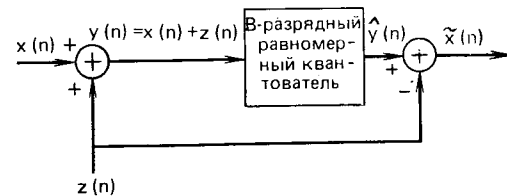


Рис. 3.5.2

и для

$$E[x^2(m)x^2(l)] = \begin{cases} B, & m=l; \\ 0, & m \neq l \end{cases}$$

определить дисперсию  $\sigma^2(n)$  как функцию  $B$  и  $\alpha$ .

в) Объяснить поведение дисперсии в п.б), если  $\alpha$  изменяется от 0 до 1.

5.8. Рассмотреть адаптивный квантователь, показанный на рис. 3.5.3а. Характеристика двухразрядного квантователя и соответствие кодовых слов представлена на рис. 3.5.3б. Предположим, что шаг квантования адаптируется в со-

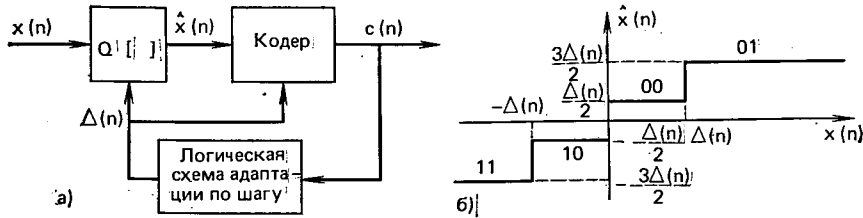


Рис. 3.5.3

ответствии с правилом  $\Delta(n) = M\Delta(n-1)$ , где  $M$  зависит от предшествующего кодового слова  $c(n-1)$  и  $\Delta_{min} \leq \Delta(n) \leq \Delta_{max}$ . Далее предположим, что

$$M = \begin{cases} P, & c(n-1) = 01 \text{ или } 11; \\ 1/P, & c(n-1) = 00 \text{ или } 10. \end{cases}$$

а) Изобразить структурную схему блока адаптации шага квантования.

б) Пусть

$$x(n) = \begin{cases} 0, & n < 5; \\ 20, & 5 \leq n \leq 13; \\ 0, & 13 < n. \end{cases}$$

Предположим, что  $\Delta_{min} = 2$  и  $\Delta_{max} = 30$ , и  $P = 2$ . Рассчитать таблицу значений  $x(n)$ ,  $\Delta(n)$ ,  $c(n)$  и  $\hat{x}(n)$  для  $0 \leq n \leq 25$  (пусть  $n=0$ ,  $\Delta(n) = \Delta_{min} = 2$  и  $c(n) = 00$ ).

г) Построить в одном масштабе последовательности  $x(n)$  и  $\hat{x}(n)$ .

5.9. Рассмотрим систему двухразрядного адаптивного квантования из задачи 5.8. Алгоритм адаптации шага квантования зададим выражением

$$\Delta(n) = \begin{cases} \beta \Delta(n-1) + D, & \text{если } \sum_{k=1}^M LSB[c(n-k)] \geq 2, \\ \beta \Delta(n-1) - \text{в противном случае,} \end{cases}$$

где  $LSB[c(n-k)]$  означает последний значащий разряд в кодовом слове  $c(n-k)$ .

а) Изобразить структурную схему устройства адаптации шага.

б) Определить максимально возможный шаг квантования и выразить его через  $\beta$  и  $D$ . (Указание: рассмотреть отклик на функцию скачка первого уравнения задачи.)

в) Предположим, что

$$x(n) = \begin{cases} 0, & n < 5; \\ 20, & 5 \leq n \leq 13; \\ 0, & 13 < n. \end{cases}$$

Предположим также, что  $M = 1$ ,  $\beta = 0.8$  и  $D = 6$ . Рассчитать таблицу значений

$x(n)$ ,  $\Delta(n)$ ,  $c(n)$  и  $\hat{x}(n)$  для  $0 \leq n \leq 25$  (при этом пусть  $n=0$ ,  $\Delta(n) = 0$  и  $c(n) = 00$ ). Изобразить  $x(n)$  и  $\hat{x}(n)$  в одной координатной системе.

г) Определить значение  $\beta$ , при котором постоянная времени шага адаптации составляет 10 мс.

5.10. Рассмотрим предсказатель первого порядка  $\hat{x}(n) = \alpha x(n-1)$ , где  $x(n)$  — стационарный случайный процесс с нулевым средним.

а) Показать, что погрешность предсказания имеет дисперсию  $\sigma_d^2 = \sigma_x^2(1 + \alpha^2 - 2\alpha\varphi(1)/\sigma_x^2)$ .

б) Показать, что  $\sigma_d^2$  минимальна при  $\alpha = \varphi(1)/\sigma_x^2 = \rho(1)$ .

в) Показать, что минимальная дисперсия погрешности равна  $\sigma_d^2 = \sigma_x^2(1 - \rho^2(1))$ .

г) При каких условиях справедливо соотношение  $\sigma_d^2 < \sigma_x^2$ ?

5.11. Дана последовательность  $x(n)$  с автокорреляционной функцией вида  $\varphi(m)$ . Показать, что разностный сигнал  $d(n) = x(n) - x(n-n_0)$  имеет тем меньшую дисперсию, по сравнению с исходным, чем больше корреляция между  $x(n)$  и  $x(n-n_0)$ . (Предположить, что среднее значение  $x(n)$  равно нулю.)

а) Определить условие на  $\varphi(n_0)$ , при котором  $\sigma_d^2 \leq \sigma_x^2$ .

б) Пусть  $d(n)$  формируется в виде  $d(n) = x(n) - \alpha x(n-n_0)$ , где  $\alpha = \varphi(n_0)/\varphi(0)$ .

Установите ограничения на  $\varphi(n_0)$ , при которых справедливо неравенство  $\sigma_d^2 \leq \sigma_x^2$ .

5.12. Используя соотношения (5.78) и (5.83), доказать, что для оптимальных коэффициентов предсказания имеет место равенство  $E[(x(n) - \hat{x}(n))\hat{x}(n)] = -E[d(n)\hat{x}(n)] = 0$ , т. е. оптимальная погрешность предсказания не коррелирована с сигналом.

5.13. Рассмотрим разностный сигнал  $d(n) = x(n) - \alpha_1 \hat{x}(n-1)$ , где  $\hat{x}(n)$  — квантованный сигнал в разностном коде.

а) Показать, что  $\sigma_d^2 = \sigma_x^2[1 - \alpha_1 \rho(1) + \alpha_1^2] + \alpha_1^2 \sigma_e^2$ .

б) Используя результат а), показать, что  $G_p = \frac{\sigma_x^2}{\sigma_d^2} = \frac{1 - \alpha_1 / (SNR_Q)}{1 - 2\alpha_1 \rho(1) + \alpha_1^2}$ , где  $SNR_Q = \sigma_x^2 / \sigma_e^2$ .

5.14. Для системы с дельта-модуляцией с переменной крутизной алгоритм адаптации шага квантования имеет вид

$$\Delta(n) = \begin{cases} \beta \Delta(n-1) + D_2, & c(n) = c(n-1) = c(n-2); \\ \beta \Delta(n-1) + D_1, & \end{cases}$$

где  $0 < \beta < 1$  и  $0 < D_1 \leq D_2$ .

а) Максимальный шаг квантования возникает при поступлении на вход фильтра сигнала в течение длительного времени  $D_2$ , что соответствует длительному периоду перегрузки по крутизне. Определить  $\Delta_{max}$  через  $D_2$  и  $\beta$ .

б) Минимальный шаг достигается, если в течение длительного времени не возникает последовательности:  $c(n) = c(n-1) = c(n-2)$ .

Определить  $\Delta_{min}$  через  $D_1$  и  $\beta$ .

5.15. Рассмотрим адаптивный дельта-модулятор (рис. 3.5.4а). Двухуровневый квантователь представлен на рис. 3.5.4б. Шаг квантования адаптируется по правилу  $\Delta(n) = M\Delta(n-1)$ , где  $\Delta_{min} \leq \Delta(n) \leq \Delta_{max}$  и множители шага равны

$$M = \begin{cases} P, & \text{если } c(n) = c(n-1); \\ 1/P, & \text{если } c(n) \neq c(n-1). \end{cases}$$

а) Изобразить структурную схему алгоритма адаптации.

б) Пусть

$$x(n) = \begin{cases} 0, & n < 5; \\ 20, & 5 \leq n \leq 13; \\ 0, & 13 < n. \end{cases}$$

Предположим, что  $\Delta_{min} = 1$ ,  $\Delta_{max} = 15$ ,  $\alpha = 1$ , и  $P = 2$ . Рассчитать таблицу значений  $x(n)$ ,  $\hat{x}(n)$ ,  $d(n)$ ,  $\Delta(n)$ ,  $\hat{d}(n)$  и  $\hat{x}(n)$  для  $0 \leq n \leq 25$ .

Предположим, что при  $n=0$ ,  $x(0)=0$ ,  $\hat{x}(0)=1$ ,  $d(0)=1$ ,  $\Delta(0)=\Delta_{\min}=1$  и  $\hat{d}(0)=1$ . Изобразить  $x(n)$  и  $\hat{x}(n)$  для  $0 \leq n \leq 25$ .

5.16. Рассмотрим два кодера, представленных на рис. 3.5.5а и б. В каждом кодере используется двухразрядный квантователь с характеристикой рис. 3.5.5в.

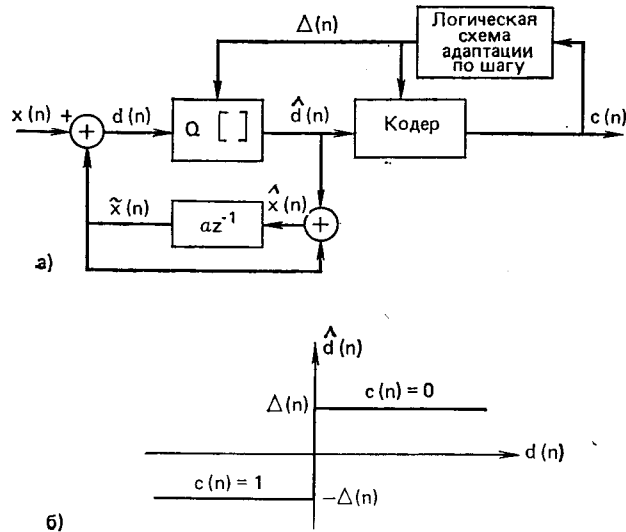


Рис. 3.5.4

Рассмотрим случай свободного канала, т. е. случай, когда есть шум малого уровня. Для простоты положим  $x(n)=0,1 \cos(\pi n/4)$ .

- Для  $0 \leq n \leq 20$  построить  $\hat{x}(n)$  для обоих кодеров.
- Для какого кодера шум незанятого канала будет более неприятным и почему?

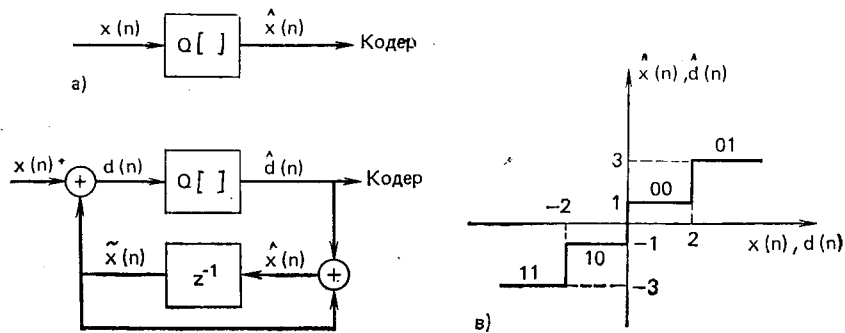


Рис. 3.5.5

5.17. Рассмотрим преобразование ИКМ-АРИКМ (см. рис. 5.46). Сигнал ИКМ  $y(n)$  можно представить в виде  $y(n)=x(n)+e_1(n)$ , где  $x(n)=x_a(nT)$  и  $e_1(n)$  — погрешность квантования при ИКМ представлении. Квантованный АРИКМ сигнал можно представить в виде  $\hat{y}(n)=y(n)+e_2(n)$ , где  $e_2(n)$  — погрешность квантования АРИКМ.

- Полагая погрешности квантования некоррелированными, показать, что полное отношение сигнал/шум  $SNR=\sigma_x^2/(\sigma_{e_1}^2+\sigma_{e_2}^2)$ .
- Показать, что отношение сигнал/шум можно записать в форме  $SNR=SNR_1/[1+(1+SNR_1)/SNR_2]$ , где  $SNR_1=\sigma_x^2/\sigma_{e_1}^2$  и  $SNR_2=\sigma_y^2/\sigma_{e_2}^2$ .

## 6

# Кратковременный анализ Фурье

## 6.0. Введение

Представление сигналов или других функций с помощью сумм синусоид и комплексных экспонент часто приводит к удобным решениям задач науки и техники и помогает понять физику явления глубже, чем это возможно другими методами. Такое представление, часто называемое Фурье-представлением, полезно при обработке сигналов по двум причинам. Во-первых, в линейных системах легко определить отклик на суперпозицию синусоид или комплексных экспонент. Во-вторых, Фурье-представление часто выявляет такие свойства сигнала, которые в первоначальном виде скрыты или по крайней мере не очевидны.

Исследования и техника передачи речи представляют собой области, в которых по традиции преобразование Фурье играло ведущую роль. Чтобы понять, почему это так, полезно вспомнить, что модель образования стационарного речевого сигнала состоит просто из линейной системы, возбуждаемой либо периодически, либо случайно. Спектр на выходе такой модели равен произведению частотной характеристики голосового тракта и спектра возбуждения. Следует ожидать поэтому, что спектр на выходе должен отражать свойства как возбуждения, так и частотной характеристики голосового тракта. Однако речевой сигнал гораздо сложнее, чем просто продолжительный гласный или фрикативный звук. Поэтому стандартное Фурье-представление, вполне пригодное для периодических, импульсных или стационарных случайных сигналов, неприменимо к речевому сигналу, характеристики которого значительно меняются во времени. Однако мы уже имели возможность убедиться в том, что принцип кратковременного анализа оказывается полезным при обработке речи. Например, такие изменяющиеся во времени характеристики, как энергия, переходы через нуль и корреляция, можно считать постоянными на интервалах времени около 10—30 мс. В этой главе будет показано, что аналогичным образом и спектральные характеристики речи можно считать изменяющимися во времени сравнительно медленно.

Для изучения спектральных характеристик речевого сигнала удобно ввести формализм, связанный с понятием кратковременного преобразования Фурье. Введем кратковременное преобразование Фурье и процедуру синтеза, основанную на нем. При этом удобно рассматривать анализ Фурье, как преобразование сигнала в гребенке фильтров. Такой подход позволяет лучше понять как теоретические, так и практические (вычислительные) аспекты кратковременного анализа. Будут рассмотрены также и другие вычислительные методы, основанные на быстром алгоритме вычисления дискретного преобразования Фурье (алгоритмы БПФ). И, наконец, подробно изучив теорию и способ вычисления кратковременного преобразования Фурье, рассмотрим его применение в задачах анализа — синтеза речи (вокодеры), визуального отображения спектра и в таких важных задачах обработки речи, как формантный анализ и выделение основного тона.

<sup>1</sup> Перевод данной главы выполнен при участии канд. физ.-мат. наук А. Ю. Шевердяева. (Прим. ред.)