

где нагрузочное сопротивление излучения через губы приближенно равно

$$Z_L(\Omega) = \frac{i\Omega L_r R_r}{R_r + i\Omega L_r} \quad (3.296)$$

Электрический аналог этого нагрузочного сопротивления есть параллельное соединение сопротивления излучения R_r и индуктивности излучения L_r . Значения R_r и L_r , которые хорошо соответствуют предположению о плоской бесконечной поверхности отражения, равны [2]

$$R_r = 128/9 \pi^2; L_r = 8 a/3 \pi c, \quad (3.30a); (3.30b)$$

где a — радиус отверстия и c — скорость звука.

Характер нагрузочного сопротивления излучения влияет на структуру волны в голосовом тракте через граничные условия (3.29). Легко видеть из (3.296), что на низких частотах $Z_L(\Omega) \approx \approx 0$, т. е. на очень низких частотах сопротивление излучения примерно соответствует короткому замыканию, которое и предполагалось ранее. Аналогично из (3.29) видно, что для диапазона средних частот (когда $\Omega L_r \ll R_r$) $Z_L(\Omega) \approx i\Omega L_r$. На высоких частотах ($\Omega L_r \gg R_r$) $Z_L(\Omega) \approx R_r$. Это отчетливо видно на рис. 3.20, где показаны действительная и мнимая части $Z_L(\Omega)$ как функции от Ω для типичных значений параметров. Энергия излучения пропорциональна реальной части сопротивления излучения. Таким образом, для полной системы речеобразования (голосовой тракт, рас-

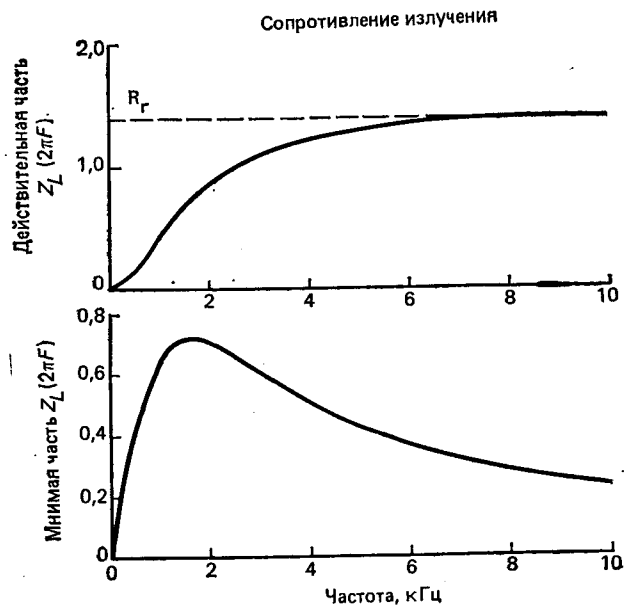


Рис. 3.20. Действительная и мнимая части сопротивления излучения

считываемый совместно с излучением) потери на излучение на высоких частотах значительны. Для количественной оценки этого эффекта уравнения (3.25), (3.26в), (3.29) можно решить совместно для случая однородной инвариантной к сдвигу трубы с мягкими стенками при наличии потерь на трение, теплопроводность и излучение с плоской и бесконечно протяженной поверхностью отражения. На рис. 3.21 показана результирующая частотная характеристика

$$V_a(i\Omega) = U(l, \Omega)/U_G(\Omega) \quad (3.31)$$

для входного сигнала $U(0, t) = U_G(\Omega) e^{i\Omega t}$. Сравнение рис. 3.21, 3.17 и 3.18 показывает, что основным эффектом наличия потерь является расширение резонансных областей, т. е. увеличение затухания

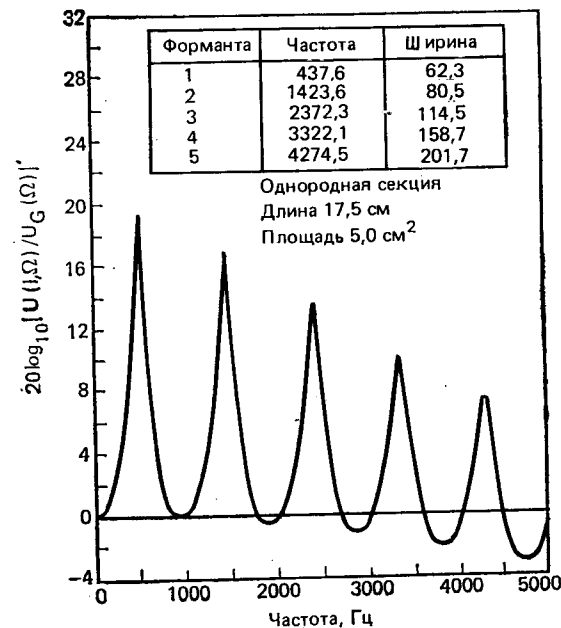


Рис. 3.21. Частотная характеристика однородной трубы с мягкими стенками, потерями на трение и теплопроводностью [18]

хания, и снижение резонансных (формантных) частот. Как и следовало ожидать, расширение резонансных областей происходит в основном на высоких частотах. Ширина первой форманты изменяется за счет потерь на стенках голосового тракта, ширина высокочастотных формант — за счет потерь на излучение, ширина второй и третьей формант формируется в результате влияния всех потерь в голосовом тракте. Частотная характеристика, показанная на рис. 3.21, определена для скорости потока у губ и скорости входного потока у губ. Представляет интерес взаимосвязь между

звуковым давлением у губ и скоростью потока у голосовой щели, особенно в том случае, когда для преобразования акустической волны в электрическое колебание используется микрофон, чувствительный к звуковому давлению. Так как $P(l, \Omega)$ и $U(l, \Omega)$ описываются выражением (3.29а), передаточная функция для звукового давления имеет простой вид:

$$H_a(\Omega) = \frac{P(l, \Omega)}{U_G(\Omega)} = \frac{P(l, \Omega)}{U(l, \Omega)} \frac{U(l, \Omega)}{U_G(\Omega)} = Z_L(\Omega) V_a(\Omega). \quad (3.32)$$

Из рис. 3.21 видно, что влияние потерь приводит к подъему характеристик на высоких частотах и к их спаду до нуля при $\Omega = 0$. На рис. 3.22 показана величина $20 \log_{10} |H_a(\Omega)|$, полученная

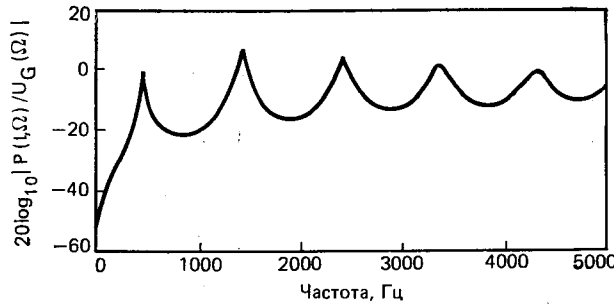


Рис. 3.22. Частотная характеристика, связывающая звуковое давление около губ и скорость потока в голосовой щели, для однородной трубы

с учетом потерь на стенках тракта и потерь на излучение при плоской и бесконечно протяженной отражательной поверхности. Сравнивая рис. 3.21 и 3.22, можно сделать вывод, что характеристики равны нулю при $\Omega = 0$ и имеют подъем на высоких частотах!

3.2.5. Передаточная функция голосового тракта для гласных

Распространение звуковых волн и их излучение при речеобразовании описываются уравнениями, приведенными в 3.2.3 и 3.2.4. Применяя методы численного интегрирования, можно решить эти уравнения в частотной или временной области и найти характеристики голосового тракта. Эти характеристики позволяют изучить подробнее процесс речеобразования и структуру речевого сигнала. В качестве примера в [18] решены уравнения (3.25), (3.26в), (3.28) и (3.29) и найдены частотные характеристики голосового тракта для функций площадей, измеренных в работе Фанта [1]. На рис. 3.23—3.26 показаны функции площади поперечного сечения голосового тракта и частотные характеристики $(U(l, \Omega)/U_G(\Omega))$ для гласных $|a|$, $|e|$, $|i|$ и $|u|$ русского языка. Эти рисунки иллюстрируют влияние потерь, которые изучались в 3.2.3 и 3.2.4. Формантные частоты и ширина формантных областей хорошо совпадают с результатами измерений, проведенных на реальных гласных Петерсоном, Барни [11], Данном [27].

В заключение сделаем следующие выводы.

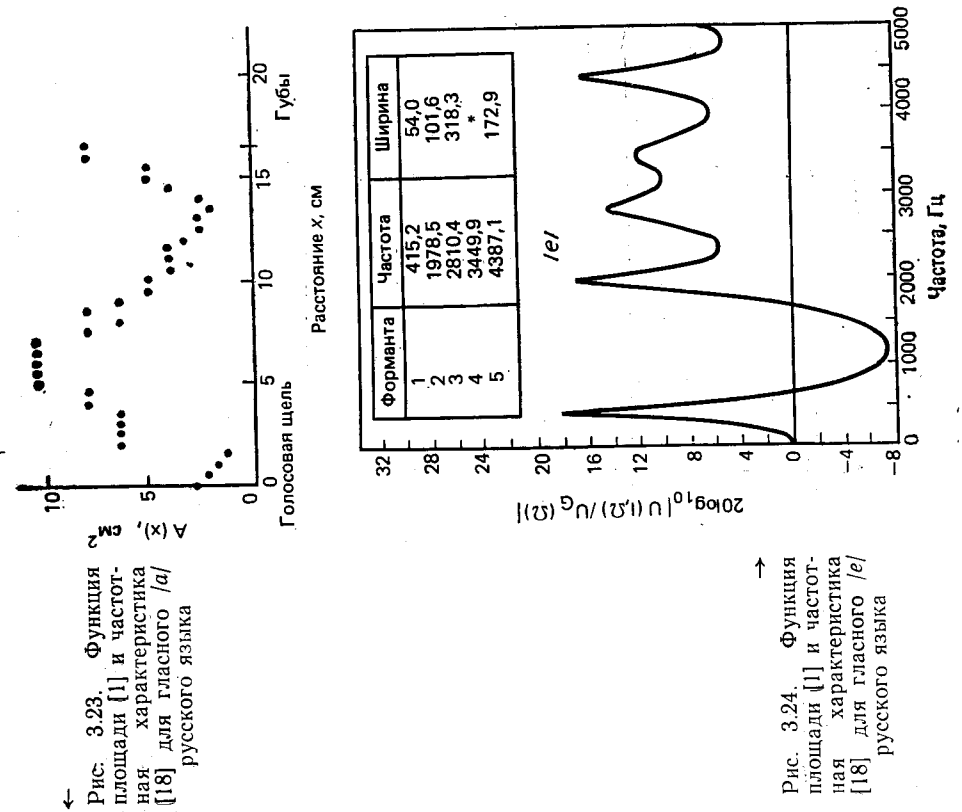
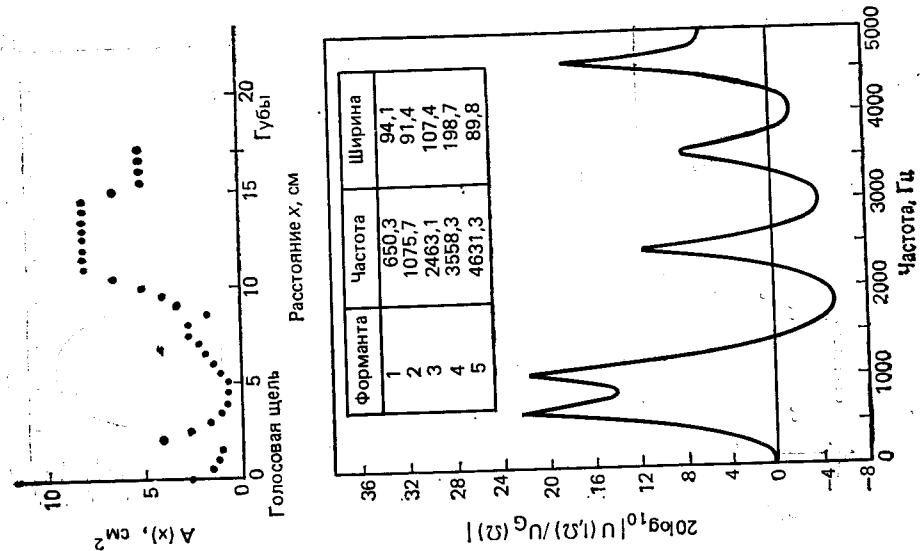


Рис. 3.23. Функция площади [1] и частотная характеристика [18] для гласного $|a|$ русского языка

Рис. 3.24. Функция площади [1] и частотная характеристика [18] для гласного $|e|$ русского языка



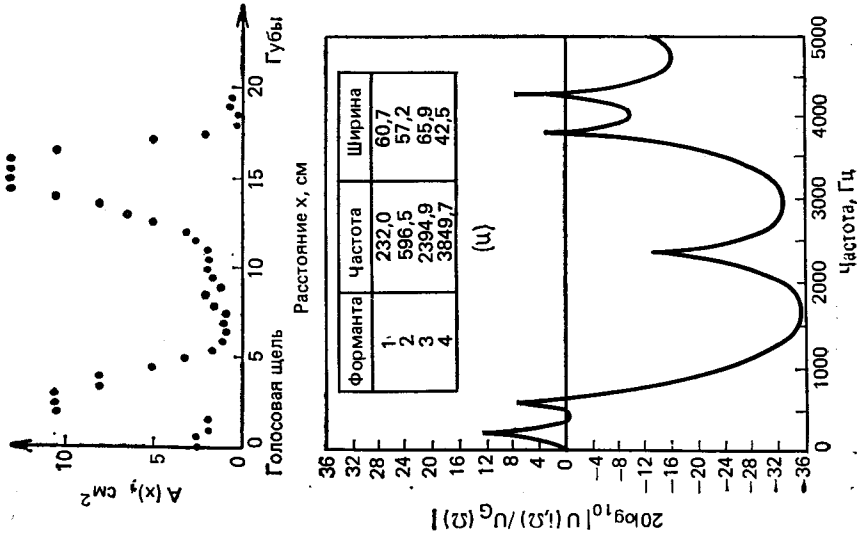


Рис. 3.25. Функция площади [1] и частотная характеристика [18] для гласного /i/ русского языка

Рис. 3.26. Функция площади [1] и частотная характеристика [18] для гласного /u/ русского языка

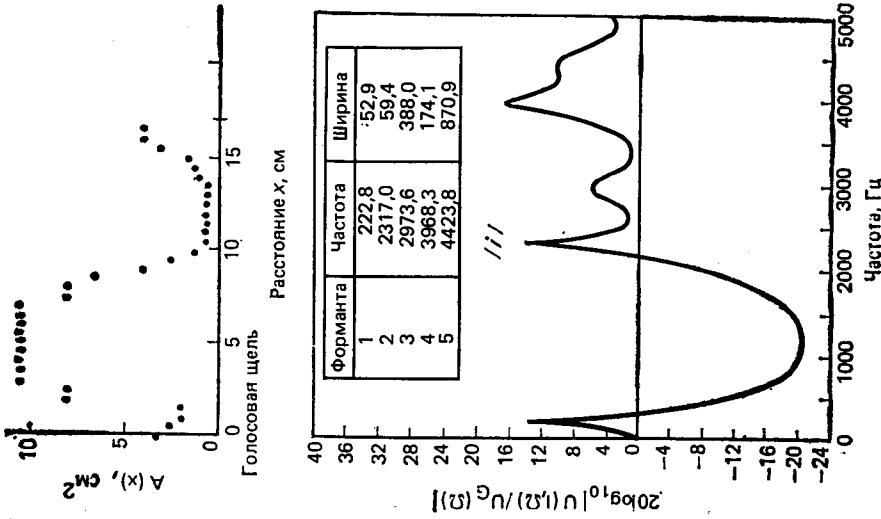


Рис. 3.25. Функция площади [1] и частотная характеристика [18] для гласного /i/ русского языка

1. Система речеобразования описывается набором резонансов (формант), которые определяются в первую очередь функцией площади поперечного сечения голосового тракта. Некоторый сдвиг резонансных частот возникает за счет потерь.

2. Ширина низкочастотных формантных областей (первой и второй) зависит от потерь на стенках голосового тракта¹.

3. Ширина высокочастотных формантных областей зависит в первую очередь от потерь на вязкое трение, теплопроводность и излучение.

3.2.6. Влияние носовой полости

При образовании носовых согласных $|m|$, $|n|$ и $|\eta|$ небная занавеска играет роль «двери» для подключения носового тракта к гортани. Одновременно в ротовой полости формируется полная смычка (перекрытие), например между губами при произнесении звука $|m|$. Такая конфигурация голосового тракта представлена на рис. 3.27а, где показаны два ответвления трубы, одно из кото-



Рис. 3.27. Модель образования носовых звуков (а) и ее электрический аналог (б)

рых полностью перекрыто. В точке соединения звуковое давление такое же, как на входе каждой трубы. Скорость звукового потока в этой точке должна быть непрерывной, т. е. скорость потока на выходе гортани должна быть равна сумме скоростей потоков на входах носовой и ротовой полостей. Соответствующий аналог в виде электрической линии показан на рис. 3.27б. Заметим, что непрерывность скоростей потока в месте соединения трех труб соответствует закону Кирхгофа для токов в точке соединения длинных линий.

Излучение звуковых волн при произнесении носовых звуков происходит в основном через ноздри. Таким образом, труба носового тракта имеет сопротивление излучения, зависящее от размера отверстия в ноздрях. Ротовой тракт, который полностью перекрыт, соответствует разомкнутой электрической линии, т. е. ток по линии не протекает. Носовые гласные формируются в подобной же системе, но ротовой тракт при этом имеет ту же форму, что и для обычных гласных. Речевой сигнал определяется взаимодей-

¹ В 3.2.7 будет показано, что потери, связанные с источником возбуждения, также влияют на низкочастотные форманты.

нием звуковых волн на выходах носовой и ротовой полостей. Математическое описание голосового тракта такой конфигурации сводится к трем системам уравнений в частных производных с граничными условиями, зависящими от формы возбуждения, конфигурации ротовой и носовой полостей и соотношения непрерывности в точке их соединения. В итоге получается весьма сложная система уравнений, которую в принципе можно решить, если располагать измерениями функции площади поперечного сечения для всех трех труб. Передаточная функция этой сложной системы будет обладать свойствами, весьма сходными со свойствами характеристик уже рассмотренных примеров. Так, система будет обладать набором формант, которые зависят от формы и длины трех труб. Существенным отличием этой системы является то, что перекрытая ротовая полость способна задерживать акустический поток при определенных его частотах, не допуская прохождения потока в носовую полость. В эквивалентной электрической линии эти частоты соответствуют частотам, на которых входное сопротивление разомкнутой линии равно нулю. На этих частотах место соединения эквивалентных линий накоротко замкнуто линией, отображающей ротовую полость. В результате в передаточной функции речеобразующей системы для носовых звуков наряду с резонансами появляются антирезонансы (нули). В результате измерений было выяснено, что для носовых звуков формантные области являются более широкими, чем для неносовых звуков. Это объясняется большими потерями за счет вязкого трения и теплопроводности, возникающими вследствие большой площади поверхности носовой полости.

3.2.7. Возбуждение звуков в голосовом тракте

В предыдущих разделах показано, как законы физики могут быть применены для описания распространения и излучения звуковых волн при речеобразовании. Для завершения изучения акустических аспектов речеобразования необходимо рассмотреть механизм возбуждения звуковых волн в речеобразующей системе. Напомним, что в общем обзоре в 3.1.1 выделено три способа возбуждения:

1. Воздушный поток, нагнетаемый из легких, модулируется за счет вибраций голосовых связок. В результате возникает квазипериодический импульсный поток.
2. Воздушный поток из легких становится турбулентным при прохождении сужения голосового тракта. В результате возникает шумоподобное возбуждение.
3. Воздушный поток сжимается легкими перед смычкой голосового тракта. Далее этот воздух внезапно высвобождается при устранении смычки, вызывая шумоподобное возбуждение.

Подробная схема возбуждения звуковых волн включает подглоточную систему (легкие, бронхи, трахею), голосовую щель и голосовой тракт. Безусловно полная модель описывает не только

речеобразование, но и процесс дыхания [2]!. Первая попытка создания физической модели возбуждения звуков в речеобразующей системе сделана в работе Фланагана [2, 28]. В последующих исследованиях разработана более совершенная модель, которая подробно описывает процесс образования вокализованной и невокализованной речи [28—31]. Эта модель основана на классической механике и механике жидкостей, и ее анализ выходит за рамки настоящей книги. Однако даже краткое качественное описание основных принципов возбуждения звуков оказывается весьма полезным для объяснения упрощенных моделей, которые широко используются при обработке речевых сигналов.

Вибрацию голосовых связок при образовании вокализованной речи можно упрощенно представить в виде рис. 3.28. Голосовые

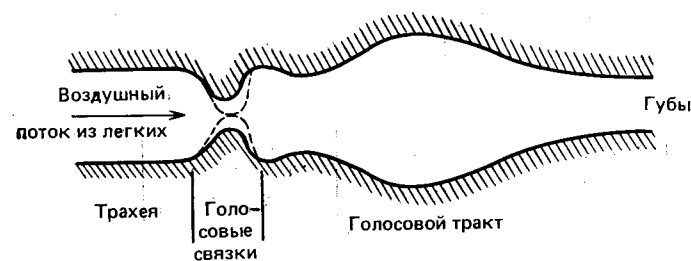


Рис. 3.28. Схематическое изображение речеобразующей системы

связки преграждают путь воздушному потоку из легких в голосовой тракт. Когда звуковое давление в легких возрастает, воздушный поток, нагнетаемый из легких, проходит через отверстие между голосовыми связками. По мере нарастания воздушного потока давление в голосовой щели падает в соответствии с законом Бернулли. Вследствие натяжения голосовых связок и уменьшения давления в голосовой щели связки соединяются, образуя полное перекрытие. На рис. 3.28 это показано пунктирными линиями. В результате давление звукового потока перед связками начинает возрастать. Когда давление повышается до уровня, достаточного чтобы раздвинуть связки, голосовая щель раскрывается и воздушный поток вновь проходит в голосовой тракт. Давление в голосовой щели снова падает, и цикл повторяется. Таким образом возникают условия, при которых голосовые связки начинают вибрировать. Частота вибрации связок зависит от давления потока, нагнетаемого из легких, массы и упругости голосовых связок, а также площади голосовой щели в свободном состоянии. Эти параметры могут быть приняты за основу создания модели голосовых связок. Такие модели должны учитывать и влияние голосового тракта, так как изменение звукового давления в голосовом тракте влияет на давление в голосовой щели. С точки зрения электрических аналогий голосовой тракт играет роль нагрузки генератора звукового возбуждения.

Схематическая диаграмма модели голосовых связок [30] показана на рис. 3.29а. Модель описывается системой сложных нелинейных дифференциальных уравнений. Объединение этих уравнений с дифференциальными уравнениями в частных производных,

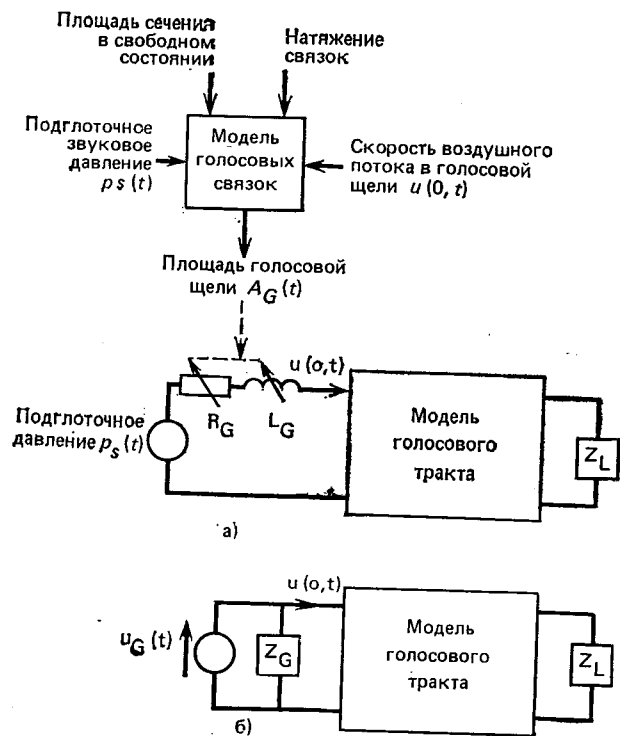


Рис. 3.29. Схематическое изображение модели голосовых связок

которыми описывается голосовой тракт, можно выполнить путем введения переменного во времени акустического сопротивления и индуктивности [30]. Это сопротивление является функцией $1/A_G(t)$. Например, когда $A_G(t) = 0$ (голосовая щель закрыта), сопротивление равно бесконечности, а скорость воздушного потока равна нулю. Таким образом, воздушный поток автоматически приобретает импульсный характер. Пример сигналов, формируемых в таких моделях, показан на рис. 3.30 [30]. В верхней части рисунка показана скорость воздушного потока, а в нижней — давление около губ для конфигурации голосового тракта, соответствующей гласной [a]. Импульсная структура потока в голосовой щели согласуется с ранее изложенным материалом и с результатами высокоскоростной киносъемки [2]. Естественно также, что затухающие колебания на выходе согласуются с изложенной трактовкой природы распространения звука в голосовом тракте.

Так как площадь голосовой щели является функцией потока в голосовом тракте, система, изображенная на рис. 3.29а, в общем случае нелинейна, хотя голосовой тракт и тракт излучения являются линейными. Взаимодействие между голосовым трактом и голосовой щелью невелико, и, как правило, им пренебрегают¹. При

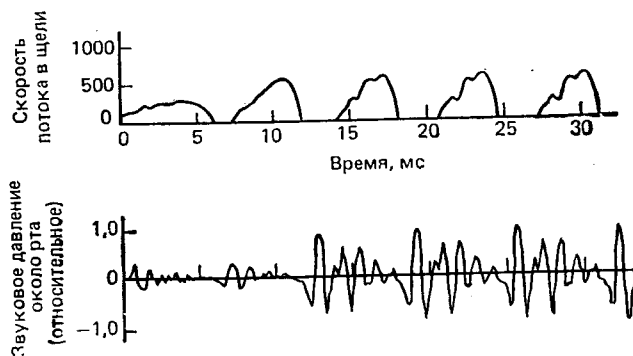


Рис. 3.30. Скорость звукового потока в голосовой щели и звуковое давление около рта для гласной [a] [30]

этом возникает возможность разделения систем возбуждения и преобразования звуковых волн и их линеаризации, как это показано на рис. 3.29б. В этом случае $u_G(t)$ является источником скорости звукового потока (источником тока), сигнал которого показан в верхней части рис. 3.30. Акустическое сопротивление голосовой щели Z_G можно получить путем линеаризации соотношений, связывающих давление и скорость воздушного потока в голосовой щели [2]. Это сопротивление равно

$$Z_G(\Omega) = R_G + i\Omega L_G, \quad (3.33)$$

где R_G и L_G постоянны. В этом случае идеальные граничные условия в частотной области $U(0, \Omega) = U_G(\Omega)$ следует заменить соотношением

$$U(0, \Omega) = U_G(\Omega) - P(0, \Omega)/Z_G(\Omega). \quad (3.34)$$

Сопротивление источника оказывает значительное влияние на ширину резонансных областей речеобразующей системы. Наиболее сильно это влияние сказывается на ширине низкочастотной резонансной области. Это происходит потому, что $Z_G(\Omega)$ растет с увеличением частоты так, что на высоких частотах Z_G соответствует разомкнутой цепи и весь поток от источника возбуждения проходит в голосовой тракт. Таким образом, вибрация стенок голосового тракта и потери в голосовой щели влияют на ширину

¹ Строго говоря, такие рассуждения неверны, так как в нелинейной системе малому по величине взаимодействию могут соответствовать существенные изменения выходного сигнала. По-видимому, именно это и проявляется при речеобразовании, что объясняет несовершенство современных систем обработки и передачи речи, в основу которых положено это «мелкое» упрощение. (Прим. ред.)

низкочастотных формантных областей, в то время как потери на излучение, трение и теплопроводность влияют на ширину высокочастотных формантных областей.

Механизм образования невокализованных звуков основан на формировании турбулентного воздушного потока. Он формируется в месте сужения голосового тракта, когда скорость потока возрастает до определенного критического уровня [2, 29]. Такое возбуждение можно имитировать путем введения источника случайного нестационарного шума в область сужения. Мощность возбуждения должна зависеть от скорости потока в трубе. Это позволяет учесть потери на трение [2, 29, 31]. При образовании фриктивных звуков параметры голосовых связей принимают такие значения, при которых связи не вибрируют. При образовании вокализованных фриктивных звуков голосовые связи вибрируют и одновременно, когда скорость потока достигает критического значения, в месте сужения голосового тракта возникает турбулентный поток. Обычно это сказывается в моменты пиков скорости импульсного воздушного потока. При произнесении взрывных звуков голосовой тракт перекрывается на период времени, когда перед смычкой воздух, нагнетаемый из легких, сжимается. Голосовые связи в это время неподвижны. Далее воздух за смычкой внезапно высвобождается, поток приобретает большую скорость и, таким образом, возникает турбулентность.

3.2.8. Модели сигнала, основанные на акустической теории

В § 3.2 изложены основные положения акустической теории речеобразования. Модели возбуждения, распространения и излучения звуковых волн описываются сложными уравнениями. Для определения речевого колебания на выходе эти уравнения можно разрешить при соответствующих значениях параметров возбуждения и голосового тракта. Естественно, что такой способ синтеза речи является наиболее эффективным [31]. Однако во многих случаях такой сложный синтез оказывается неприемлемым. В этих ситуациях на основе акустической теории можно получить упрощенные модели синтеза. На рис. 3.31 показана общая структурная



Рис. 3.31. Модель речеобразования

схема, по которой разработано множество моделей, применяемых при обработке речевых сигналов. Основной особенностью этих моделей является то, что источник возбуждения и голосовой тракт рассматриваются как отдельные системы¹. Голосовой тракт с уче-

¹ См. прим. ред. на с. 81.

том излучения представлен линейной системой с переменными параметрами. Эта система отображает резонансные явления в головном тракте. Генератор возбуждения формирует сигнал либо в виде последовательности импульсов, либо в форме шумоподобного процесса. Параметры источника возбуждения и линейной системы выбираются так, что формируемый на выходе сигнал оказывается речеподобным. Если удастся достигнуть этого, то полученная модель может быть использована при обработке речевого сигнала. В последующей части главы изучаются модели такого типа.

3.3. Модели с трубами без потерь

Одна из наиболее распространенных моделей речеобразования основана на предположении, что голосовой тракт можно представить соединением акустических труб без потерь (рис. 3.32). Площади поперечного сечения труб выбираются так, чтобы результирующая оказалась равной функции площади поперечного сечения голосового тракта $A(x)$. Если количество труб велико, а их длина достаточно мала, то можно ожидать,

что резонансные частоты такого соединения будут близки к резонансным частотам трубы с непрерывной функцией площади сечения. Поскольку, однако, при таком приближении пренебрегают потерями на трение, теплопроводность и вибрацию стенок, можно ожидать, что ширина резонансных областей будет отличаться от ширины таких областей в сложной модели, учитывающей и потери. Потери в голосовой щели и около губ могут быть учтены и, как будет показано в гл. 8, это позволяет достаточно точно описать резонансные свойства речевого сигнала.

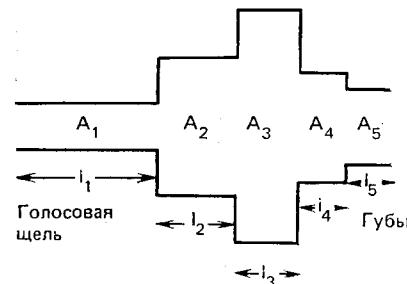


Рис. 3.32. Соединение пяти труб без потерь

Более важным вопросом является то обстоятельство, что модели с трубами без потерь позволяют перейти от непрерывного времени к дискретному. Иначе говоря, предметом последующего изучения является модель рис. 3.32.

3.3.1. Распространение звуковых волн в соединении труб без потерь

Так как потери в трубах (см. рис. 3.32) отсутствуют, распространение звуковых волн в каждой трубе описывается уравнениями (3.2) с соответствующими значениями площади поперечного сечения. Если рассмотреть k -ю трубу с площадью сечения A_k , то звуковое давление и скорость потока в трубе

$$p_k(x, t) = \frac{\rho c}{A_k} [u_k^+(t - x/c) + u_k^-(t + x/c)]; \quad (3.35a)$$

$$u_k(x, t) = u_k^+(t - x/c) - u_k^-(t + x/c), \quad (3.35b)$$

где x — расстояние от левого конца соединения до k -й трубы ($0 \leq x \leq l_k$) и $u_k^+(\cdot)$, $u_k^-(\cdot)$ — прямая и отраженная волны. Взаимосвязь между волнами в соседних трубах можно установить на ос-

нове следующего соображения. Звуковое давление и скорость потока должны быть непрерывными в каждый момент времени и в любой точке системы. Это позволяет задать граничные условия для обоих концов трубы.

Рассмотрим, в частности, соединение k -й и $(k+1)$ -й труб (рис. 3.33). Из условия непрерывности получаем

$$p_k(l_k, t) = p_{k+1}(0, t); u_k(l_k, t) = u_{k+1}(0, t). \quad (3.36a); (3.36b)$$

Подставляя (3.35) в (3.36), имеем

$$\frac{A_{k+1}}{A_k} [u_k^+(t - \tau_k) + u_k^-(t + \tau_k)] = u_{k+1}^+(t) + u_{k+1}^-(t); \quad (3.37a)$$

$$u_k^+(t - \tau_k) - u_k^-(t + \tau_k) = u_{k+1}^+ - u_{k+1}^-(t), \quad (3.37b)$$

где $\tau_k = l_k/c$ — время прохождения волны через k -ю трубу. Из рис. 3.33 видно, что часть прямой волны распространяется далее (направо), а часть — отражается. Аналогично часть отраженной вол-

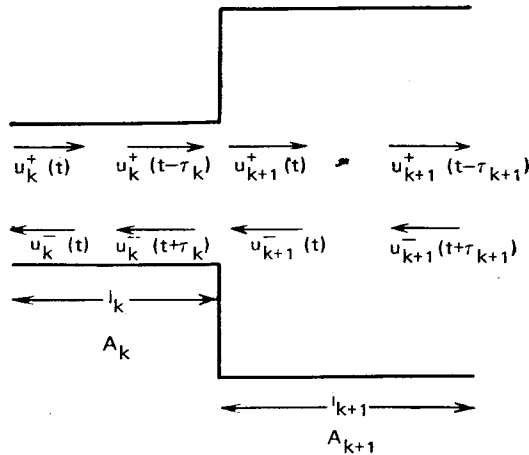


Рис. 3.33. Соединение двух труб без потерь

ны распространяется далее (налево), а другая часть — вновь отражается. Таким образом, если выразить $u_{k+1}^+(t)$ и $u_{k+1}^-(t + \tau_{k+1})$ через $u_k^+(t - \tau_k)$ и $u_k^-(t + \tau_k)$, то можно подробно изучить, как распространяются прямые и отраженные волны через всю систему. Решая (3.37b) для $u_k^-(t + \tau_k)$ и подставляя результат в (3.37a), получаем

$$u_{k+1}^+(t) = \left[\frac{2A_{k+1}}{A_{k+1} + A_k} \right] u_k^+(t - \tau_k) + \left[\frac{A_{k+1} - A_k}{A_{k+1} + A_k} \right] u_{k+1}^-(t). \quad (3.38a)$$

Вычитая (3.37b) из (3.37a), имеем

$$u_k^-(t + \tau_k) = - \left[\frac{A_{k+1} - A_k}{A_{k+1} + A_k} \right] u_k^+(t - \tau_k) + \left[\frac{2A_k}{A_{k+1} + A_k} \right] u_{k+1}^-(t). \quad (3.38b)$$

Из (3.38a) видно, что величина

$$r_k = (A_{k+1} - A_k) / (A_{k+1} + A_k). \quad (3.39)$$

определяет величину отраженной волны $u_{k+1}^-(t)$. Поэтому величину r_k называют коэффициентом отражения для k -го соединения труб. Можно показать (см. задачу 3.4), что так как площади поперечного сечения положительны, то

$$-1 \leq r_k \leq 1. \quad (3.40)$$

На основе определения r_k уравнения (3.38) можно записать так:

$$u_{k+1}^+(t) = (1 + r_k) u_k^+(t - \tau_k) + r_k u_{k+1}^-(t); \quad (3.41a)$$

$$u_k^-(t + \tau_k) = -r_k u_k^+(t - \tau_k) + (1 - r_k) u_{k+1}^-(t). \quad (3.41b)$$

Эти уравнения впервые были использованы для синтеза речи Келли и Лохбаумом [32]. Удобно изображать эти уравнения в виде графа рис. 3.34. На рисунке для отображения операций сложения

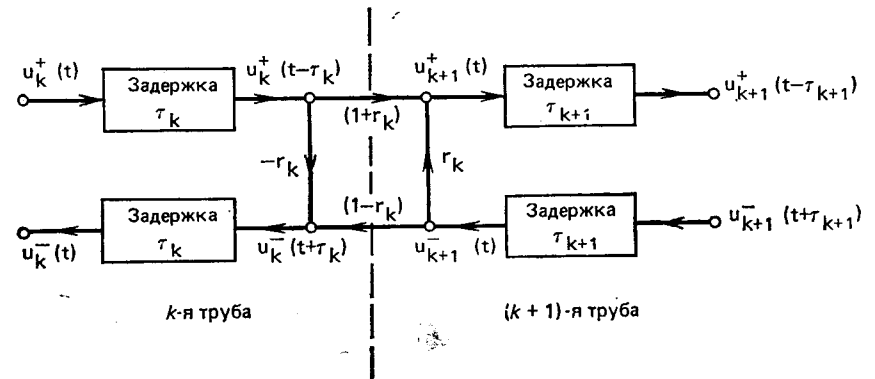


Рис. 3.34. Граф прохождения сигнала через соединение двух труб без потерь

и умножения в (3.41) использовано графическое изображение преобразований сигнала¹. Любое соединение труб на рис. 3.32 можно представить аналогично рис. 3.34, если анализу подвергаются только значения звукового давления и скорости потока на входе и выходе труб. Такой анализ не вносит ограничений так как нас интересует только соотношение между выходом последней трубы и входом первой трубы. Модель из пяти труб рис. 3.32 будет состоять из пяти элементов задержки в прямом и обратном направлениях и четырех соединений, каждое из которых описывается коэффициентом отражения. Для полного описания распространения волны в такой системе остается ввести граничные условия со стороны губ и голосовой щели.

¹ В качестве введения к использованию сигнальных графов при обработке сигналов рекомендуется [32].

3.3.2. Граничные условия

Пронумеруем N секций модели числами от 1 до N , начиная со стороны голосовой щели. Тогда граничные условия со стороны губ будут определять звуковое давление $p_N(l_N, t)$ и скорость потока $u_N(l_N, t)$ на выходе N -й трубы с учетом излучения. В частотной области справедливо соотношение

$$P_N(l_N, \Omega) = Z_L U_N(l_N, \Omega). \quad (3.42)$$

Если предположить, что Z_L действительная величина, то во временной области справедливо соотношение

$$\rho c (u_N^+(t - \tau_N) + u_N^-(t + \tau_N)) / A_N = Z_L (u_N^+(t - \tau_N) - u_N^-(t + \tau_N)). \quad (3.43)$$

(Если Z_L — комплексная величина, то вместо (3.43) получается дифференциальное уравнение, связывающее $p_N(l_N, t)$ и $u_N(l_N, t)$.) Решая уравнение для $u_N^-(t + \tau_N)$, получаем

$$u_N^-(t + \tau_N) = -r_L u_N^+(t - \tau_N), \quad (3.44)$$

где коэффициент отражения около губ равен

$$r_L = \left[\frac{\rho c / A_N - Z_L}{\rho c / A_N + Z_L} \right]. \quad (3.45)$$

Скорость выходного потока около губ

$$u_N(l_N, t) = u_N^+(t - \tau_N) - u_N^-(t + \tau_N) = (1 + r_L) u_N^+(t - \tau_N). \quad (3.46)$$

На выходе трубы справедливы уравнения (3.44), (3.46). Модель окончания трубы изображена на рис. 3.35. Заметим, что если Z_L комплексная величина, то (3.45) остается справедливым, но величина r_L здесь также будет комплексной и для получения соотношения, аналогичного (3.44), надо перейти в частотную область. Во временной области $u_N^-(t + \tau_N)$ и $u_N^+(t - \tau_N)$ будут связаны дифференциальным уравнением (см. задачу 3.5).

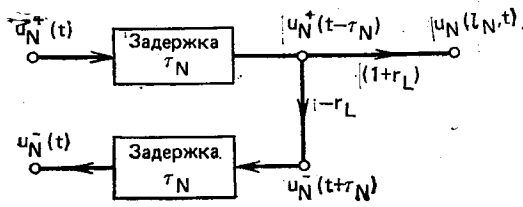


Рис. 3.35. Нагрузка со стороны губ

Соотношения в частотной области в предположении, что источник возбуждения линейно независим от голосового тракта, приведены в § 3.2.7. Применяя это предположение, здесь также можно получить соотношение для давления и скорости входного потока первой трубы

$$U_1(0, \Omega) = U_G(\Omega) - P_1(0, \Omega) / Z_G. \quad (3.47)$$

Снова предположив, что Z_G — действительная величина, имеем

$$u_1^+(t) - u_1^-(t) = u_G(t) - \frac{\rho c}{A_1} \left[\frac{u_1^+(t) + u_1^-(t)}{Z_G} \right]. \quad (3.48)$$

Решая для $u_1^+(t)$, получаем (см. задачу 3.6)

$$u_1^+(t) = \frac{(1+r_G)}{2} u_G(t) + r_G u_1^-(t), \quad (3.49)$$

где коэффициент отражения у голосовой щели равен

$$r_G = \left[\frac{Z_G - \rho c / A_1}{Z_G + \rho c / A_1} \right]. \quad (3.50)$$

Выражение (3.49) изображено в виде схемы на рис. 3.36. Так же, как и в случае излучения, если Z_G комплексная величина, то (3.50) остается справедливым. Однако r_G будет комплексным и (3.49) следует заменить аналогичным соотношением в частотной области или $u^+(t)$ и $u_G(t)$, $u_1^-(t)$ могут быть связаны дифференциальным уравнением. Обычно для упрощения полагают, что Z_G и Z_L — действительные величины.

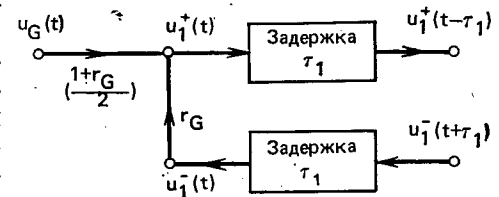


Рис. 3.36. Нагрузка со стороны голосовой щели

В качестве примера на рис. 3.37 приведена полная схема распространения волны в системе из двух труб. Скорость потока около губ обозначена через $u_L(t) = u_2(l_2, t)$. В частотной области

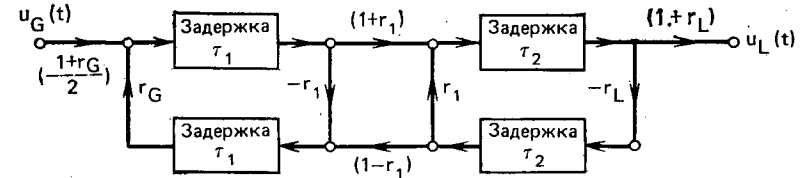


Рис. 3.37. Полный граф двухтрубной модели

можно получить (см. задачу 3.7) выражение для частотной характеристики

$$V_a(\Omega) = \frac{U_L(\Omega)}{U_G(\Omega)} = \frac{0,5(1+r_G)(1+r_L)(1+r_1)e^{-i\Omega(\tau_1+\tau_2)}}{1+r_1r_Ge^{-i\Omega 2\tau_1} + r_1r_Le^{-i\Omega 2\tau_2} + r_Lr_Ge^{-i\Omega 2(\tau_1+\tau_2)}}. \quad (3.51)$$

Отметим несколько особенностей (3.51). Во-первых, в числителе имеется коэффициент $e^{-i\Omega(\tau_1+\tau_2)}$. Он отображает задержку прохождения потока от голосовой щели до губ. Системную функцию можно найти путем замены $i\Omega$ на s :

$$V_a(s) = \frac{0,5(1+r_G)(1+r_L)(1+r_1)e^{-s(\tau_1+\tau_2)}}{1+r_1r_Ge^{-s 2\tau_1} + r_1r_Le^{-s 2\tau_2} + r_Lr_Ge^{-s 2(\tau_1+\tau_2)}}. \quad (3.52)$$

Полюса $V_a(s)$ являются комплексными резонансными частотами системы. Видно, что количество полюсов бесконечно, так как зависимость от s экспоненциальная. В работах Фанта [1] и Фланагана [2] показано, что можно получить хорошее соответствие резонансных частот модели реальным формантным частотам гласных, если подобрать длины труб и площади их поперечного сечения (см. также задачу 3.8).

3.3.3. Связь с цифровыми фильтрами

Анализ $V_a(s)$ показывает, что модель из двух труб без потерь имеет свойства, близкие к цифровым фильтрам. Чтобы убедиться в этом, рассмотрим систему, состоящую из N труб без потерь длиной $\Delta x = l/N$, где l — длина голосового тракта. Такая система изображена на рис. 3.38 для $N=7$. Распространение волн в такой системе можно представить схемой, изображенной на рис. 3.34, с задержками на интервал $\tau = \Delta x/c$, равный времени прохождения волны через одну трубу. Изучение отклика системы на возбуждение в виде единичного импульса $u_G(t) = \delta(t)$. Импульс возбуждения проходит через все трубы, частично отражаясь от мест их соединений. Подробный анализ распространения импульса показывает, что импульсная характеристика системы (скорость потока около губ при импульсном возбуждении со стороны голосовой щели)

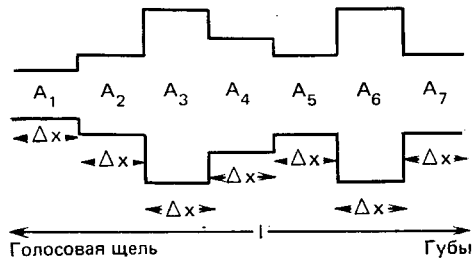


Рис. 3.38. Соединение ($N=7$) труб без потерь одинаковой длины

импульса показывает, что импульсная характеристика системы (скорость потока около губ при импульсном возбуждении со стороны голосовой щели)

$$v_a(t) = \alpha_0 \delta(t - N\tau) + \sum_{k=1}^{\infty} \alpha_k \delta(t - N\tau - 2k\tau). \quad (3.53)$$

Очевидно, что импульс может появиться на выходе только через $N\tau$ с. Импульсы, отраженные от мест соединений, достигнут выхода на 2τ с позже. Величина 2τ — время распространения импульса вдоль одной трубы в прямом и обратном направлениях. Системная функция имеет вид

$$V_a(s) = \sum_{k=0}^{\infty} \alpha_k e^{-s(N+2k)\tau} = e^{-sN\tau} \sum_{k=0}^{\infty} \alpha_k e^{-s2\tau k}. \quad (3.54)$$

Коэффициент $e^{-sN\tau}$ отображает задержку, т. е. время распространения импульса вдоль всех N секций. Выражение

$$\hat{V}_a(s) = \sum_{k=0}^{\infty} \alpha_k e^{-sk2\tau} \quad (3.55)$$

является передаточной функцией линейной системы, импульсная характеристика которой равна $\hat{v}_a(t) = v_a(t + N\tau)$. Эта функция определяет резонансные свойства системы. На рис. 3.39а приведена структурная схема модели с трубами без потерь, в которой задержка отделена от системы $\hat{v}_a(t)$. Частотная характеристика $\hat{V}_a(\Omega)$ равна

$$\hat{V}_a(\Omega) = \sum_{k=0}^{\infty} \alpha_k e^{-i\Omega k2\tau}. \quad (3.56)$$

Легко показать, что

$$\hat{V}_a(\Omega + 2\pi/2\tau) = \hat{V}_a(\Omega). \quad (3.57)$$

Очевидно, что эта характеристика напоминает частотную характеристику системы в дискретном времени. Действительно, если спектр входного сигнала сосредоточен в диапазоне ниже $\pi/(2\tau)$, то можно дискретизировать входной сигнал с периодом $T=2\tau$ и затем пропустить сигнал через цифровой фильтр с импульсной характеристикой

$$v(n) = \begin{cases} \alpha_n, & n \geq 0, \\ 0, & n < 0. \end{cases} \quad (3.58)$$

Для периода $T=2\tau$ время задержки на $N\tau$ с соответствует сдвигу на $N/2$. Эквивалентная система в дискретном времени для входного сигнала с ограниченным по частоте спектром показана на рис. 3.39б. Отметим, что если N — четное, то $N/2$ — целое число и задержка может быть реализована как простой сдвиг выходной последовательности первой системы. Если N нечетное, то для получения выходного сигнала в схеме рис. 3.39а необходимо проводить интерполяцию. Однако задержкой можно пренебречь, так как в большинстве приложений она не имеет существенного значения.

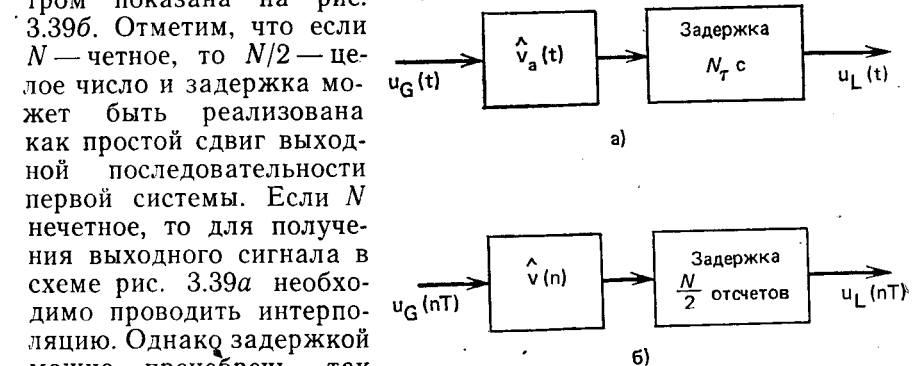


Рис. 3.39. Модель акустической трубы без потерь (а) и эквивалентная дискретная система (б)

Выполнив замену e^{sT} на z в $\hat{V}_a(s)$, получим z -преобразование $\hat{V}(z)$:

$$\hat{V}(z) = \sum_{k=0}^{\infty} \alpha_k z^{-k}. \quad (3.59)$$

Граф прохождения сигнала в системе с дискретным временем можно построить по графу аналоговой системы. Каждую переменную аналоговой системы надо заменить на соответствующую последовательность отсчетов. Кроме того, каждую задержку на τ с надо заменить элементом задержки на половину периода дискретизации, так как $\tau = T/2$. Пример такого графа показан на рис. 3.40. Отметим, что задержка отображена на рис. 3.40б ветвью с коэффициентом передачи $z^{-1/2}$.

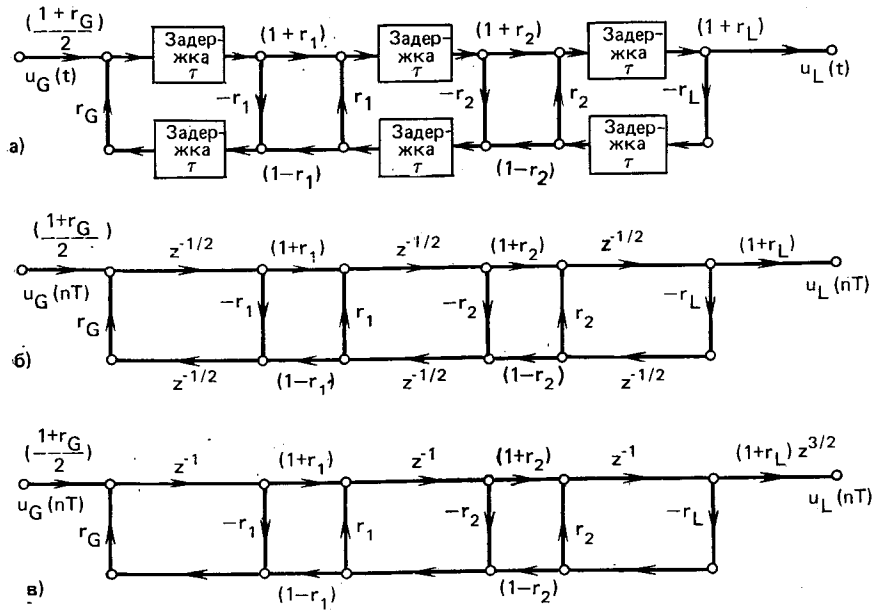


Рис. 3.40. Граф модели голосового тракта (а), эквивалентная дискретная система (б) и эквивалентная дискретная система с объединенными элементами задержки (в)

Задержка на половину периода дискретизации (рис. 3.40б) предполагает интерполяцию сигнала между двумя отсчетами. Такая интерполяция не может быть реализована непосредственно. Можно получить более удобную структуру, если использовать то обстоятельство, что схема рис. 3.40б имеет лестничную форму с элементами задержки в верхних и нижних ветвях. Сигнал распространяется в прямом направлении (вправо) в верхних ветвях и в обратном направлении (влево) — в нижних ветвях. Из рисунка видно, что задержка внутри замкнутого пути сохранится, если задержку в нижней ветви объединить с задержкой в верхней ветви. Общая задержка прохождения сигнала от входа к выходу при этом изменится, однако это не имеет большого значения на практике, а теоретически это изменение может быть устранено введе-

нием корректирующего упреждения $z^{N/2}$. На рис. 3.40в показано, как эти рассуждения можно реализовать на примере модели из трех труб. Достоинством такой структуры является то, что для этой системы могут быть записаны разностные уравнения, которые позволяют путем итераций получить выходной сигнал при заданном входном.

Цифровые системы [33], подобные изображенной на рис. 3.40в, могут быть использованы для формирования отсчетов синтезированного речевого сигнала по отсчетам сигнала возбуждения [32]. Система такой структуры является довольно сложной. Каждая ветвь, коэффициент передачи которой не равен единице, предполагает выполнение операции умножения. Таким образом, в каждой секции требуется выполнить четыре операции умножения и две — сложения. В общем для модели с N трубами (рис. 3.40в) надо выполнять $4N$ операций умножения и $2N$ — сложения. Так как выполнение операции умножения требует много времени, представляет интерес разработка других структур (или другой организации вычислений), которые требуют выполнения меньшего числа операций умножения. Такие структуры могут быть получены в результате подробного анализа схемы рис. 3.41а. Разностные уравнения для этой схемы имеют вид

$$u^+(n) = (1+r)w^+(n) + ru^-(n); \quad (3.60a)$$

$$w^-(n) = -rw^+(n) + (1-r)u^-(n). \quad (3.60б)$$

Эти уравнения можно переписать:

$$u^+(n) = w^+(n) + rw^+(n) + ru^-(n); \quad (3.61a)$$

$$w^-(n) = -rw^+(n) - ru^-(n) + u^-(n). \quad (3.61б)$$

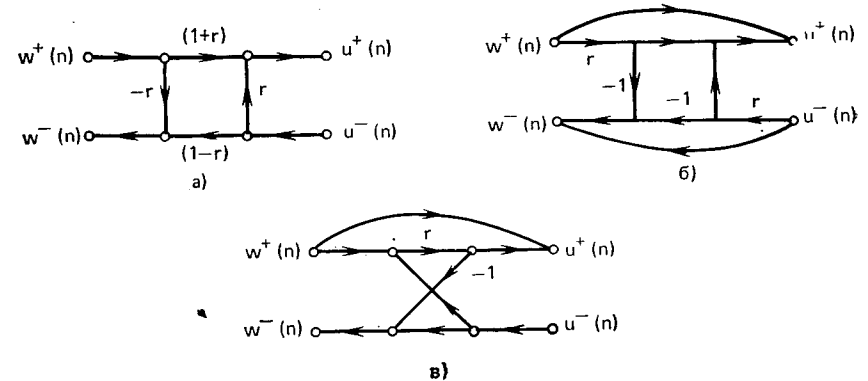


Рис. 3.41. Модель с четырьмя операциями умножения (а), двумя операциями умножения (б) и одной операцией умножения (в)

¹ Все задержки можно объединить в нижних ветвях. Тогда потребуется ввести корректирующую задержку на $N/2$.

Так как члены $r\omega^+(n)$ и $r\omega^-(n)$ входят в оба уравнения, два выхода четырех ветвей с операциями умножения можно устранить так, как это сделано на рис. 3.41б. В полученной структуре имеется две операции умножения и четыре — сложения. Другую структуру можно получить, группируя слагаемые с коэффициентом r следующим образом:

$$u^+(n) = \omega^+(n) + r[\omega^+(n) + u^-(n)]; \quad (3.62a)$$

$$\omega^-(n) = u^-(n) - r[\omega^+(n) + u^-(n)]. \quad (3.62b)$$

Так как слагаемое $r[\omega^+(n) + u^-(n)]$ входит в оба уравнения, здесь требуется выполнить одну операцию умножения и три — сложения (рис. 3.41в). Такая форма представления модели с трубами без потерь впервые была получена Итакурой и Саито [34]. При использовании модели с трубами без потерь для синтеза речевых сигналов выбор вычислительной схемы зависит от скорости, с которой могут быть выполнены операции умножения и сложения, а также от простоты выполнения контроля над вычислениями.

3.3.4. Передаточная функция модели с трубами без потерь

Для полного завершения изучения модели с трубами без потерь в дискретном времени полезно получить общее выражение для передаточной функции через коэффициенты отражения. Рассматриваемые далее уравнения впервые были получены Аталом и Ханауером [35], Маркелом, Греем [36] и Вакитой [37] при исследовании анализа речи с помощью линейного предсказания. В гл. 8 будет рассмотрена взаимосвязь модели с трубами без потерь с моделью линейного предсказания. Основными вопросами данного раздела являются вывод общего выражения для передаточной функции модели с трубами без потерь и выявление вытекающих из этой модели разнообразных представлений сигнала.

Передаточная функция определена выражением

$$V(z) = U_L(z)/U_G(z). \quad (3.63)$$

Таким образом, для получения $V(z)$ достаточно выразить $U_G(z)$ через $U_L(z)$ и записать их отношение. С этой целью рассмотрим рис. 3.42, на котором показана одна секция модели трубы без потерь. Уравнения в z -плоскости для этой секции имеют вид

$$U_{k+1}^+(z) = (1+r_k)z^{-1/2}U_k^+(z) + r_k U_{k+1}^-(z); \quad (3.64a)$$

$$U_k^-(z) = -r_k z^{-1}U_k^+(z) + (1-r_k)z^{-1/2}U_{k+1}^-(z). \quad (3.64b)$$

Решая эти уравнения относительно $U_k^+(z)$, $U_k^-(z)$, получим

$$U_k^+(z) = \frac{z^{1/2}}{1+r_k}U_{k+1}^+(z) - \frac{r_k z^{1/2}}{1+r_k}U_{k+1}^-(z); \quad (3.65a)$$

$$U_k^-(z) = \frac{-r_k z^{-1/2}}{1+r_k}U_{k+1}^+(z) + \frac{z^{-1/2}}{1+r_k}U_{k+1}^-(z). \quad (3.65b)$$

Уравнения (3.65) позволяют выразить $U_G(z)$ через $U_L(z)$ на основе связи между выходом и входом модели с трубами без потерь.

Для получения результата в компактной форме удобно записать граничные условия около губ в том же виде, что и для стыков между трубами. Чтобы сделать это, введем $U_{N+1}(z)$ как z -

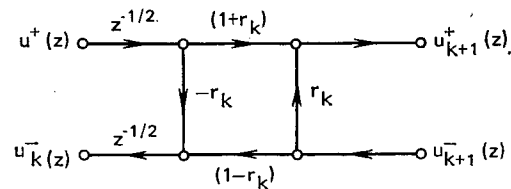


Рис. 3.42. Граф для z -преобразований

преобразования входного сигнала фиктивной $(N+1)$ -й трубы, которая имеет бесконечную длину, так что отраженная волна в ней отсутствует. Можно рассуждать и иначе, полагая, что нагрузкой $(N+1)$ -й трубы является характеристическое сопротивление. В обоих случаях $U_{N+1}^+(z) = U_L(z)$ и $U_{N+1}^-(z) = 0$. Тогда из (3.39) и (3.45) видно, что если $A_{N+1} = \rho c/Z_L$, то можно положить $r_N = r_L$.

Уравнение (3.65) можно записать в матричной форме:

$$U_k = Q_k U_{k+1}, \quad (3.66)$$

где

$$U_k = \begin{bmatrix} U_k^+(z) \\ U_k^-(z) \end{bmatrix} \quad (3.67)$$

и

$$Q_k = \begin{bmatrix} \frac{z^{1/2}}{1+r_k} & \frac{-r_k z^{1/2}}{1+r_k} \\ \frac{-r_k z^{-1/2}}{1+r_k} & \frac{z^{-1/2}}{1+r_k} \end{bmatrix}. \quad (3.68)$$

Итерируя (3.66), можно показать, что входная переменная первой трубы может быть выражена через переменную на выходе с помощью матричного произведения:

$$U_1 = Q_1 Q_2 \dots Q_N U_{N+1} = \prod_{k=1}^N Q_k U_{N+1}. \quad (3.69)$$

Из рис. 3.36 видно, что граничные условия со стороны голосовой щели можно записать в виде

$$U_G(z) = \frac{2}{(1+r_G)} U_1^+(z) - \frac{2r_G}{1+r_G} U_1^-(z). \quad (3.70)$$

Это выражение можно переписать:

$$U_G(z) = \left\| \frac{2}{1+r_G}, -\frac{2r_G}{1+r_G} \right\| U_1. \quad (3.71)$$

Так как

$$U_{N+1} = \begin{vmatrix} U_L(z) \\ 0 \end{vmatrix} = \begin{vmatrix} 1 \\ 0 \end{vmatrix} U_L(z), \quad (3.72)$$

и окончательно можно записать

$$\frac{U_G(z)}{U_L(z)} = \begin{vmatrix} 2 \\ 1+r_G \end{vmatrix}, -\frac{2r_G}{1+r_G} \begin{vmatrix} \prod_{k=1}^N Q_k \\ 0 \end{vmatrix}. \quad (3.73)$$

Это решение равно $1/V(z)$.

Для изучения свойств $V(z)$ полезно записать

$$Q_k = z^{1/2} \begin{vmatrix} 1 & -r_k \\ 1+r_k & 1+r_k \\ -r_k z^{-1} & z^{-1} \\ 1+r_k & 1+r_k \end{vmatrix} = z^{1/2} \hat{Q}_k. \quad (3.74)$$

Выражение (3.73) можно записать в форме

$$\frac{1}{V(z)} = z^{N/2} \begin{vmatrix} 2 \\ 1+r_G \end{vmatrix}, -\frac{2r_G}{1+r_G} \begin{vmatrix} \prod_{k=1}^N \hat{Q}_k \\ 0 \end{vmatrix}. \quad (3.75)$$

Так как элементы \hat{Q}_k либо постоянные величины, либо пропорциональны z^{-1} , то полное матричное произведение можно представить в виде полинома от z^{-1} степени N . Например, можно показать (см. задачу 3.9), что для $N=2$ справедливо выражение

$$\frac{1}{V(z)} = \frac{2(1+r_1 r_2 z^{-1} + r_1 r_G z^{-1} + r_2 r_G z^{-2})z}{(1+r_G)(1+r_1)(1+r_2)} \quad (3.76)$$

или

$$V(z) = \frac{0,5(1+r_G)(1+r_1)(1+r_2)z^{-1}}{1+(r_1 r_2 + r_1 r_G)z^{-1} + r_2 r_G z^{-2}}. \quad (3.77)$$

Из (3.74) и (3.75) видно, что в общем случае передаточная функция модели с трубами без потерь может быть записана как

$$V(z) = \frac{0,5(1+r_G) \prod_{k=1}^N (1+r_k) z^{-N/2}}{D(z)}, \quad (3.78a)$$

где $D(z)$ — полином переменной z^{-1} — задается матрицей

$$D(z) = \begin{vmatrix} 1 & -r_1 \\ -r_1 z^{-1} & z^{-1} \end{vmatrix} \cdots \begin{vmatrix} 1 & -r_N \\ -r_N z^{-1} & z^{-1} \end{vmatrix} \begin{vmatrix} 1 \\ 0 \end{vmatrix}. \quad (3.78b)$$

Из (3.78b) для $D(z)$ можно получить

$$D(z) = 1 - \sum_{k=1}^N \alpha_k z^{-k}. \quad (3.79)$$

Другими словами, передаточная функция модели с трубами без потерь имеет столько элементов задержки, сколько секций в модели. Передаточная функция имеет полюса и не имеет нулей. Эти полюса определяют резонансы или форманты модели.

В специальном случае $r_G=1$ ($Z_G=\infty$) полином $D(z)$ можно найти по рекурсивной формуле, вытекающей из (3.78b). Если вычислять матричное произведение слева, то вначале надо вычислить произведение матрицы-строки размером 1×2 и матрицы размером 2×2 , а в конце вычислений умножить результат на матрицу-столбец размером 2×1 в правой части (3.78b). Рекурсивная формула может быть получена путем вычисления первых нескольких матричных произведений. Введем

$$P_1 = \begin{vmatrix} 1 & -r_1 \\ -r_1 z^{-1} & z^{-1} \end{vmatrix} = \begin{vmatrix} 1+r_1 z^{-1} & -(r_1+z^{-1}) \end{vmatrix}. \quad (3.80)$$

Если далее ввести

$$D_1(z) = 1 + r_1 z^{-1}, \quad (3.81)$$

то легко показать, что

$$P_1 = \begin{vmatrix} D_1(z) & -z^{-1} D_1(z^{-1}) \end{vmatrix}. \quad (3.82)$$

Аналогично введем матрицу-строку

$$P_2 = P_1 \begin{vmatrix} 1 & -r_2 \\ -r_2 z^{-1} & z^{-1} \end{vmatrix}. \quad (3.83)$$

Если выполнить умножение, то можно показать, что

$$P_2 = \begin{vmatrix} D_2(z) & -z^{-2} D_2(z^{-1}) \end{vmatrix}, \quad (3.84)$$

где

$$D_2(z) = D_1(z) + r_2 z^{-2} D_1(z^{-1}). \quad (3.85)$$

По индукции получаем

$$P_k = P_{k-1} \begin{vmatrix} 1 & -r_k \\ -r_k z^{-1} & z^{-1} \end{vmatrix} = \begin{vmatrix} D_k(z) & -z^{-k} D_k(z^{-1}) \end{vmatrix}. \quad (3.86)$$

где

$$D_k(z) = D_{k-1}(z) + r_k z^{-k} D_{k-1}(z^{-1}). \quad (3.87)$$

Окончательно

$$D(z) = P_N \begin{vmatrix} 1 \\ 0 \end{vmatrix} = D_N(z). \quad (3.88)$$

Таким образом, нет необходимости вычислять все матричные произведения. Достаточно выполнить вычисления по рекурсивной формуле:

$$D_0(z) = 1; \quad (3.89a)$$

$$D_k(z) = D_{k-1}(z) + r_k z^{-k} D_{k-1}(z^{-1}), \quad k = 1, 2, \dots, N; \quad (3.89b)$$

$$D(z) = D_N(z). \quad (3.89b)$$

Эффективность модели с трубами без потерь можно продемонстрировать путем вычисления передаточной функции для функций площадей поперечного сечения, использованных для получения рис. 3.23—3.26. При выполнении этого расчета нужно учесть нагрузку со стороны губ и количество секций в модели. В наших определениях нагрузка излучения была представлена трубой с площадью сечения A_{N+1} , в которой отсутствует отраженная волна. Значение A_{N+1} выбирается из соображений получения требуемого коэффициента отражения на выходе. Влияние излучения является единственным источником потерь в системе (если $r_G=1$), и, таким образом, можно ожидать, что выбором величины A_{N+1} можно изменять ширину резонансных областей $V(z)$. Например, при $A_{N+1}=\infty$ имеем $r_N=r_L=1$, т. е. коэффициент отражения соответствует акустическому короткому замыканию. Этот случай соответствует полному отсутствию потерь. Обычно A_{N+1} выбирают так, чтобы получить коэффициент отражения со стороны губ, при котором получается требуемая ширина резонансных областей. Пример подобного расчета излагается ниже.

Выбор числа секций зависит от частоты дискретизации речевого сигнала. Пусть частотная характеристика модели с трубами без потерь периодическая. В этом случае модель правильно отображает поведение голосового тракта только в диапазоне частот $|F| < 1/(2T)$, где T — период дискретизации. Ранее было установлено, что $T=2\tau$, где τ — время распространения звуковой волны через одну секцию в одном направлении. Если имеется N секций общей длиной l , то $\tau=l/(cN)$. Так как знаменатель передаточной функции является полиномом степени N , существует $N/2$ комплексных полюсов, соответствующих резонансам в полосе частот $|F| < 1/(2T)$. Если принять $l=17,5$, $c=35\,000$ см/с, то получаем

$$1/2T = 1/4\tau = NC/4l = 1000 N/2 \text{ Гц.} \quad (3.90)$$

Это означает, что в полосе частот до 1000 Гц в голосовом тракте длиной 17,5 см будет $N/2$ резонансов (формант). Если $1/T = 10\,000$ Гц, то общая полоса частот составляет 5000 Гц. Это означает, что N должно быть равно 10. Просматривая рис. 3.21—3.26, можно сделать вывод о том, что резонансы голосового тракта расположены с плотностью примерно одна форманта на 1000 Гц. Голосовому тракту меньшей длины будет соответствовать меньшая плотность расположения резонансов и наоборот.

На рис. 3.43 приведен пример для $N=10$ и $1/T=10$ кГц. На рис. 3.43а показаны значения функций площадей поперечного сечения, взятые из рис. 3.23 и дискретизированные для получения десятисекционной модели гласного /а/. На рис. 3.43б показаны десять коэффициентов отражения для $A_{11}=30$ см². При этом коэффициент отражения около губ был равен $r_N=0,714$. Отметим, что наибольшая величина коэффициента отражения соответствует месту наибольшего изменения функции площади. На рис. 3.43в показана частотная характеристика для $r_N=1$ и 0,714 (пунктирная линия). Сравнивая пунктирную линию рис. 3.43в с рис. 3.23, можно

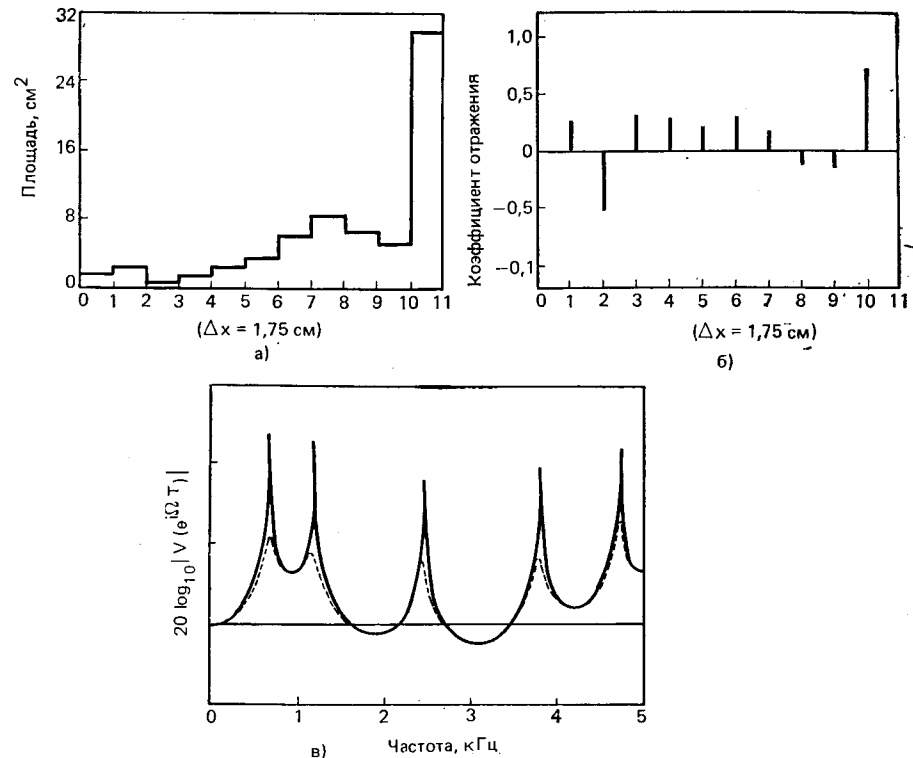


Рис. 3.43. Функция площади десятисекционной трубы без потерь, нагруженной трубой с площадью 30 см² (а); коэффициенты отражения десятисекционной трубы (б); частотная характеристика десятисекционной трубы (в). Пунктирная линия соответствует условиям (б); сплошная линия соответствует короткому замыканию. Данные (а) взяты из [1] для гласного /а/ русского языка

сделать вывод, что при наличии потерь на излучение частотная характеристика модели с трубами без потерь очень близка к характеристике сложной модели с потерями.

3.4. Цифровые модели речевых сигналов

В § 3.2 было показано, что можно получить довольно подробное математическое описание акустического процесса речеобразования. Наша цель при изучении этой теории состоит в выяснении основных особенностей речевого сигнала и в установлении того, как эти особенности согласуются с физикой речеобразования. Было рассмотрено три способа возбуждения звуков и выяснено, что каждому способу соответствует свой тип выходного сигнала. Установлено также, что сигнал возбуждения проходит через голосовой тракт и преобразуется в нем в соответствии с резонансами тракта, образуя звуки речи. Это явление составляет предмет дальнейшего изучения.

Из проведенного обстоятельного рассмотрения моделей речеобразования можно сделать важный вывод. Уже ясно, что для общего описания речевых сигналов можно разработать «эквивалентную модель», такую, как показано на рис. 3.31. Здесь линейная система, выходной сигнал которой обладает свойствами, близкими к свойствам речи, управляется множеством параметров примерно так, как это происходит при речеобразовании. Таким образом, выходной сигнал этой модели эквивалентен выходному сигналу физической модели, но ее внутренняя структура не связана с физикой речеобразования. Представляет интерес вопрос о построении подобных эквивалентных моделей в дискретном времени для описания дискретизированных речевых сигналов. Для формирования речеподобного сигнала тип возбуждения и резонансные свойства линейной системы должны изменяться во времени. Характер этих изменений рассматривался в § 3.1. В частности, колебания (см. рис. 3.3а) показывают, что свойства сигнала изменяются довольно медленно. Для многих звуков речи можно считать, что тип возбуждения и свойства голосового тракта остаются неизменными в течение 10—20 мс. Таким образом, эквивалентная аналоговая модель должна состоять из линейной системы с медленно изменяющимися во времени параметрами, возбуждаемой сигналом от источника возбуждения. Сигнал возбуждения представляет собой квазипериодическую последовательность импульсов для вокализованной речи и шумоподобный процесс для невокализованной речи.

Примером такой подходящей модели является модель с трубами без потерь в дискретном времени, которая рассматривалась выше. Основные особенности этой модели отображены на рис. 3.4а. Пусть система, отображающая голосовой тракт, описывается

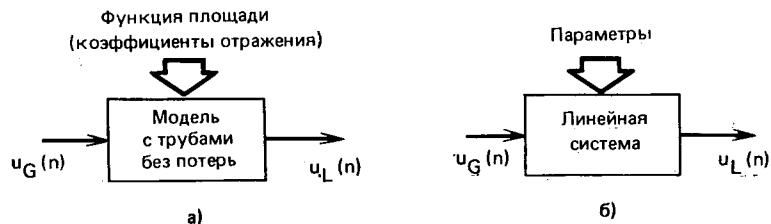


Рис. 3.44. Диаграмма модели с трубами без потерь (а) и эквивалентная модель (б)

набором площадей или коэффициентами отражения. Системы вида рис. 3.40в могут быть применены для формирования речевого сигнала по соответствующему возбуждению. Соотношение между входным и выходным сигналами может быть записано через передаточную функцию

$$V(z) = G \left/ \left(1 - \sum_{k=1}^N \alpha_k z^{-k} \right) \right. \quad (3.91)$$

где G и $\{\alpha_k\}$ зависят от функции площади (постоянная задержка

из (3.78а) здесь опущена). Нас интересует только выходной сигнал, поэтому любая система, которая обладает данной передаточной функцией, способна формировать такой выходной сигнал по заданному воздействию. (Это не вполне справедливо для систем с переменными параметрами, но возможные различия можно свести к минимуму.) Таким образом, эквивалентные модели в дискретном времени в общем виде имеют структуру рис. 3.44б. Это указывает на возможность описания голосового тракта с помощью фильтра.

Для получения полной модели речевого сигнала необходимо также учесть изменение звукового возбуждения и эффект излучения через губы. В заключительной части этой главы будут рассмотрены отдельно составляющие такой полной модели.

3.4.1. Голосовой тракт

Резонансы (форманты) речевого сигнала соответствуют полюсам передаточной функции $V(z)$. Полюсная модель дает хорошее описание голосового тракта для большинства звуков речи. Однако акустическая теория показывает, что для носовых и фрикативных звуков надо учитывать и резонансы и антирезонансы (полюса и нули). В этом случае в передаточную функцию следует ввести нули или, следуя работе Атала [35], полагать, что наличие нулей можно учесть, увеличивая количество полюсов (см. задачу 3.10). Последнее соображение применяется наиболее часто.

Так как коэффициенты знаменателя $V(z)$ в (3.91) действительные, корни полинома в знаменателе могут быть действительными или комплексно-сопряженными. Типичная комплексная резонансная частота голосового тракта имеет вид

$$s_k, s_k^* = -\sigma_k \pm i 2\pi F_k T \quad (3.92)$$

Соответствующая пара комплексно-сопряженных полюсов в модели с дискретным временем может быть записана так:

$$z_k, z_k^* = e^{-\sigma_k T} e^{\pm i 2\pi F_k T} = e^{-\sigma_k T} \cos(2\pi F_k T) \pm i e^{-\sigma_k T} \sin(2\pi F_k T) \quad (3.93)$$

Ширина резонансной области голосового тракта равна примерно $2\sigma_k$, а центральная частота — $2\pi F_k$ [26]. В z -плоскости длина линии от начала координат до координаты полюса определяет ширину резонансной области:

$$|z_k| = e^{-\sigma_k T}, \quad (3.94a)$$

а угол наклона линии определяется соотношением

$$\theta_k = 2\pi F_k T. \quad (3.94b)$$

Таким образом, если знаменатель $V(z)$ факторизован, то значения формантных частот и ширины формантных областей могут быть найдены по (3.94). Как показано на рис. 3.45, комплексные частоты голосового тракта расположены в левой полуплоскости s ,

так как голосовой тракт является устойчивой системой. Следовательно, $\sigma_k > 0$ и $|z_k| < 1$. Это означает, что все полюса модели в дискретном времени располагаются внутри единичного круга. На рис. 3.45 показаны типичные комплексные резонансные частоты в s - и z -плоскостях.

В § 3.3 было показано, как для модели с трубами без потерь получить передаточную функцию (3.91). Можно показать [35, 36], что если площади поперечного сечения труб положительные, все полюса $V(z)$ будут находиться внутри единичного круга. Обратно, по $V(z)$ из (3.91) можно определить модель с трубами без потерь [35, 36]. Один из путей реализации требуемой передаточной функции состоит в использовании лестничной структуры рис. 3.40в, одну секцию которой можно представить в виде рис. 3.41.

Другим способом реализации передаточной функции является применение

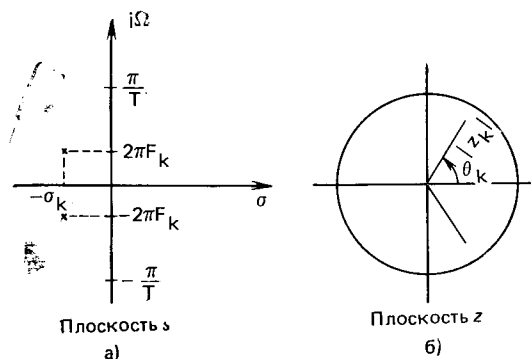


Рис. 3.45. Резонансы голосового тракта: а) в s -плоскости; б) в z -плоскости

описанных в гл. 2 структур, используемых в обычных цифровых фильтрах. Например, можно применить прямую форму реализа-

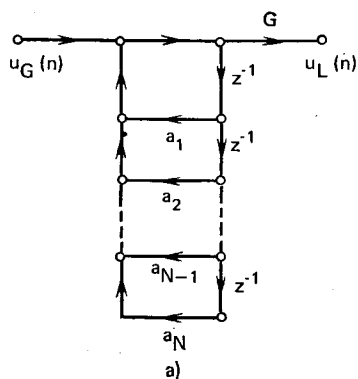
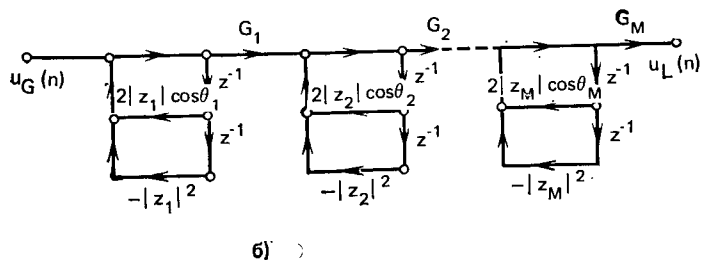


Рис. 3.46. Прямая (а) и каскадная (б) формы реализации $(G_k = 1 - 2|z_k| \cos \theta_k + |z_k|^2)$



б) >

ции (рис. 3.46а); можно использовать и каскадное соединение систем второго порядка (резонаторов), для которых

$$V(z) = \prod_{k=1}^M V_k(z), \quad (3.95)$$

где M равно $((N+1)/2)$, округленному до большего целого, и

$$V_k(z) = \frac{(1 - 2|z_k| \cos(2\pi F_k T) + |z_k|^2)}{(1 - 2|z_k| \cos(2\pi F_k T) z^{-1} + |z_k|^2 z^{-2})}. \quad (3.96)$$

Числитель $V_k(z)$ следует выбирать так, чтобы коэффициент усиления фильтра совпадал с коэффициентом усиления модели с трубами без потерь. Заметим, что на нулевой частоте ($z=1$) $V_k(1)=1$. Каскадная модель показана на рис. 3.46б. В задаче 3.11 указан путь уменьшения числа операций умножения в такой модели. Другой способ реализации $V(z)$ основан на ее разложении на простые дроби и построении параллельной структуры. Этот способ рассматривается в задаче 3.12.

Интересно отметить, что каскадные и параллельные структуры вначале использовались как аналоговые модели. Они обладают существенным ограничением: аналоговые системы второго порядка (резонаторы) имеют частотные характеристики, быстро убывающие с ростом частоты. Это потребовало введения звеньев «коррекции высокочастотных полюсов», соединенных последовательно с формантными резонаторами для обеспечения необходимого наклона частотной характеристики на высоких частотах. В результате цифрового моделирования Голд и Рабинер [38] показали, что цифровые резонаторы, благодаря присущим им свойствам периодичности частотных характеристик, обеспечивают требуемый наклон характеристики на высоких частотах. В этом можно убедиться на примере модели с трубами без потерь. Таким образом, в цифровых моделях нет необходимости вводить цепи коррекции.

3.4.2. Излучение

Ранее рассматривалась передаточная функция $V(z)$, описывающая взаимосвязь скорости потока источника и скорости выходного потока около губ. Если нужно получить модель, описывающую характер звукового давления около губ (как это обычно и требуется), необходимо учитывать эффект излучения. В 3.2.4 показано, что в аналоговых моделях звуковое давление и скорость потока связаны уравнением (3.29). Аналогичное соотношение в z -плоскости имеет вид

$$P_L(z) = R(z) U_L(z). \quad (3.97)$$

Из 3.2.4 и рис. 3.20 можно сделать вывод, что звуковое давление связано со скоростью воздушного потока операцией высокочастотной фильтрации. В частности, можно считать, что на низких частотах звуковое давление примерно равно производной скорости

потока. Для представления этой взаимосвязи в дискретном времени необходимо использовать методы цифровой техники, позволяющие избежать явления «наложения частот». Например, метод билинейного преобразования (обычный метод проектирования цифровых фильтров) обеспечивает достаточно хорошее описание эффекта излучения (см. задачу 3.13) при

$$R(z) = R_0(1 - z^{-1}), \quad (3.98)$$

т. е. при вычислении разности первого порядка (более точное приближение рассматривается в задаче 3.13). «Грубое» дифференцирование путем вычисления разности первого порядка соответствует обычно применяемому «грубому» описанию связи давления со скоростью потока посредством операции дифференцирования на низких частотах.

Эта модель излучения может быть включена последовательно с моделью голосового тракта (рис. 3.47). Передаточную функцию

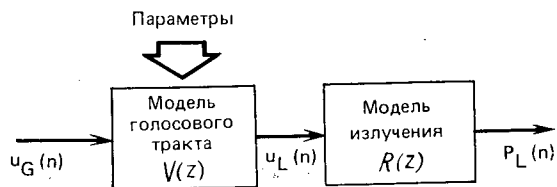


Рис. 3.47. Эквивалентная модель речеобразования с учетом излучения

$V(z)$ можно реализовать любым подходящим способом с параметрами, соответствующими требуемой конфигурации голосового тракта: функции площади сечения модели с трубами без потерь или формантным областям каскадной модели.

3.4.3. Возбуждение

Для завершения построения эквивалентной модели необходимо изучить способы описания возбуждения голосового тракта. Будем считать, что большинство звуков речи можно отнести либо к вокализованным, либо к невокализованным. В первом случае источник возбуждения должен формировать квазипериодическую последовательность импульсов, а во втором — шумоподобное случайное колебание.

Для синтеза вокализованной речи сигнал возбуждения должен иметь вид, изображенный на рис. 3.30. Один из способов получения такого сигнала показан на рис. 3.48. Генератор последовательности импульсов формирует единичные импульсы, повторяющиеся через период основного тона. Этот сигнал поступает на линейную систему, импульсная характеристика которой $g(n)$ соответствует форме колебания в голосовой щели. Коэффициент усиления A_v определяет интенсивность голосового возбуждения.

Форма $g(n)$ несущественна и требуется только, чтобы преобразование Фурье от $g(n)$ обладало «правильными» свойствами. Розенберг [39], исследуя влияние формы импульса возбуждения на

качество восприятия речевого сигнала, выяснил, что импульс голосового возбуждения может иметь следующий вид:

$$g(n) = \begin{cases} \frac{1}{2} [1 - \cos(\pi n/N_1)], & 0 \leq n \leq N_1; \\ \cos(\pi(n - N_1)/2N_2), & N_1 \leq n \leq N_1 + N_2; \\ 0, & \text{в противном случае.} \end{cases} \quad (3.99)$$

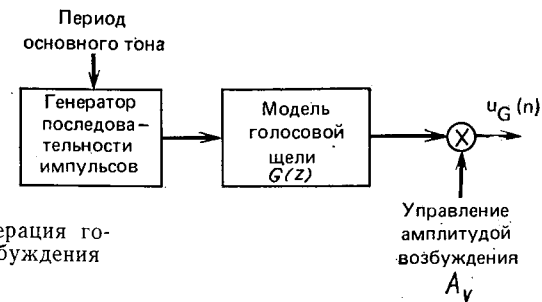


Рис. 3.48. Генерация голосового возбуждения

Этот импульс очень напоминает импульсы, изображенные на рис. 3.30. На рис. 3.49 показаны импульсное колебание и его амплитудный спектр для типичных значений N_1 и N_2 . Из рисунка видно, что, как и следовало ожидать, спектр импульсного возбуждения сосредоточен в низкочастотном диапазоне.

Так как $g(n)$ в (3.99) имеет конечную длительность, z -преобразование $G(z)$ имеет только нули. Однако в большинстве случаев требуется располагать полюсной моделью, поэтому обычно применяют двухполюсную функцию $G(z)$ [36].

Для невокализованных звуков модель возбуждения гораздо проще. Здесь достаточно располагать источником шума и изменять коэффициент усиления для получения требуемой мощности возбуждения. Для моделей в дискретном времени в качестве такого источника мо-

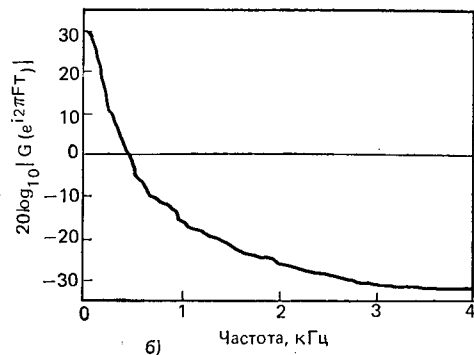
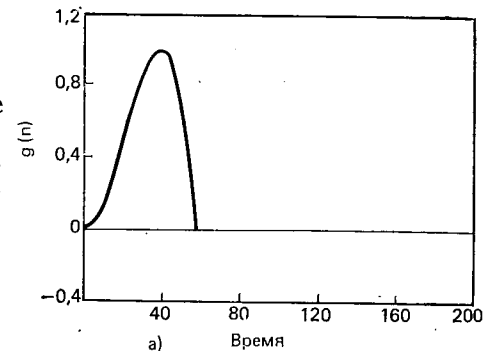


Рис. 3.49. Аппроксимация импульса голосового возбуждения Розенберга (а) и спектр (б)

жет быть использован генератор случайных чисел, формирующих последовательность с равномерным спектром, функция распределения шумового возбуждения при этом несущественна.

3.4.4. Полная модель

Объединим все компоненты модели (рис. 3.50). Здесь переключением источников возбуждения можно изменять характер сигнала возбуждения. Голосовой тракт можно представить различным

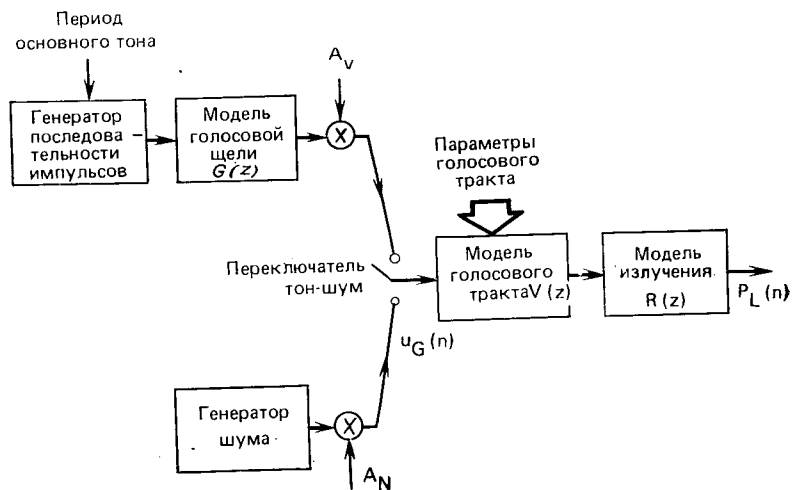


Рис. 3.50. Общая дискретная модель речеобразования

образом. В некоторых случаях удобно объединить модели голосового тракта, возбуждения и излучения в одну систему. В частности, далее будет показано, что в случае анализа на основе линейного предсказания модели голосового возбуждения, излучения и голосового тракта удобно объединить вместе, записав общую передаточную функцию в виде

$$H(z) = G(z)V(z)R(z). \quad (3.100)$$

Эта функция имеет только полюса. Другими словами, схема, изображенная на рис. 3.50, дает лишь общее представление о речеобразовании. Существует много разновидностей этой модели.

Важным вопросом является выяснение ограничений этой модели. Очевидно, что модель весьма далека от тех дифференциальных уравнений в частных производных, с которых было начато изучение. Можно выделить несколько ограничений. Первое состоит в характере изменения параметров. Для протяжных звуков, таких как гласные, параметры изменяются довольно медленно и в этом случае модель оказывается достаточно точной. При произнесении кратковременных, например, взрывных звуков модель уже не яв-

ляется адекватной. Следует подчеркнуть, что использование понятий «передаточная функция» и «частотная характеристика» предполагает «кратковременный» анализ сигнала. Таким образом, предполагается, что параметры модели постоянны на интервалах 10—20 мс. Передаточная функция $V(z)$ хорошо отображает структуру звуков, для которых параметры медленно изменяются во времени. В последующих главах это обстоятельство будет часто упоминаться. Второе ограничение состоит в отсутствии нулей передаточной функции, необходимых для точного описания носовых и фрикативных звуков. Это ограничение имеет большее значение для носовых звуков и несколько меньшее для фрикативных. При необходимости нули можно ввести в передаточную функцию модели. Третье ограничение состоит в упрощенном дихотомическом разделении типов возбуждения: вокализованное или невокализованное; такое разделение не соответствует вокализованным фрикативным звукам. Устранить это ограничение путем простого сложения сигналов возбуждения двух типов не удастся, так как для фрикативных звуков импульсы основного тона коррелированы с шумовым возбуждением. Смешанная модель возбуждения для вокализованных фрикативных звуков разработана в [40] и может быть использована там, где это необходимо. Наконец, еще одним недостатком модели, изображенной на рис. 3.50, является то, что импульсы голосового возбуждения повторяются с периодом, кратным интервалу дискретизации T . В [41] рассмотрены пути устранения этого ограничения для ситуаций, в который требуется точное управление основным тоном.

3.5. Заключение

В данной главе изложены три основных вопроса: звуки речи, физика речеобразования и дискретные модели речеобразования. Обзор акустической фонетики и теории речеобразования был довольно обстоятельным, но далеко не полным. Основная задача заключалась в изучении таких характерных особенностей речи, располагая которыми можно осознанно вводить модели, полезные для обработки речевых сигналов.

Модели, рассмотренные в § 3.3 и 3.4, являются основой всего последующего содержания книги. Эти модели будут использоваться в двух направлениях. Первое направление называется анализом речи, второе — синтезом. В первом случае предмет изучения охватывает методы оценивания параметров модели в предположении, что выходным сигналом модели является речь. При синтезе речи модели используются для формирования речевого сигнала путем управления соответствующими параметрами модели. Такие задачи встречаются во многих приложениях. Основой изучения являются модели данной главы. Располагая обзором теории цифровой обработки сигналов (гл. 2) и акустической теории речеобразования (гл. 3), можно приступить к изучению способов применения методов цифровой обработки к речевым сигналам.

Задачи

- 3.1. На рис. 3.3.1 приведена временная диаграмма речевого сигнала длительностью 500 мс (100 мс на каждом отрезке).
а) Указать области вокализованной, невокализованной речи и пауз (шума).

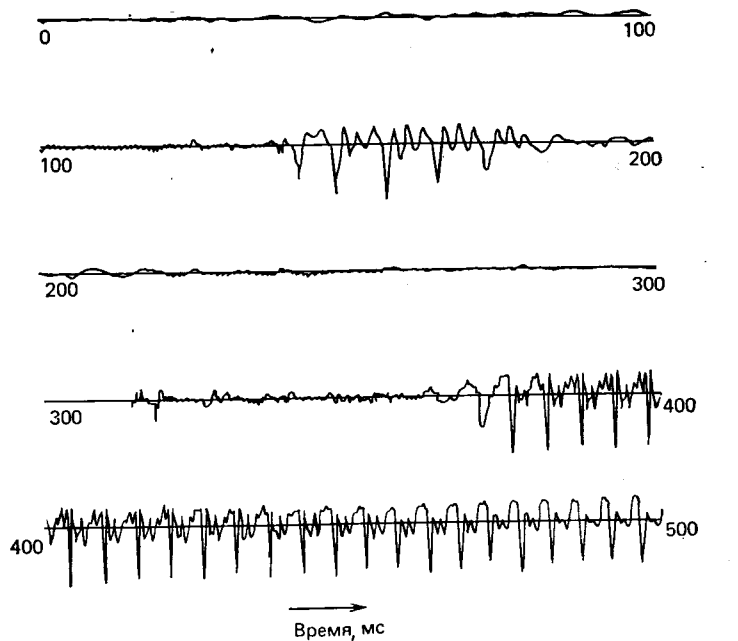


Рис. 3.3.1

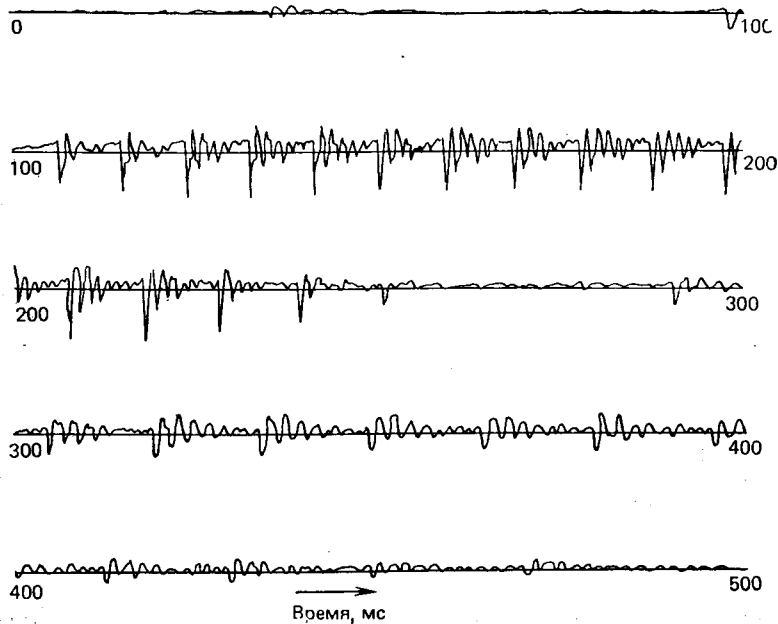


Рис. 3.3.2

б) Для участков вокализированной речи оценить период основного тона и изобразить зависимость периода от времени (для невокализированных участков и пауз положить, что период равен нулю).

3.2. Колебание рис. 3.3.2 соответствует слову *cattle*. Каждый отрезок соответствует 100 мс.

а) Отметить границы между фонемами, т. е. указать временные границы между звуками $[c|a|t|t|le]$.

б) Отметить моменты, где частота основного тона (i) — максимальна, (ii) — минимальна. Определить частоту основного тона в эти моменты времени.

в) Является ли диктор мужчиной, женщиной или ребенком, и почему?

3.3. Показать что (3.3) есть решение дифференциальных уравнений в частных производных (3.2). Подставить для этого (3.3) в (3.2).

3.4. Коэффициенты отражения в месте соединения двух труб без потерь с площадями сечения A_k и A_{k+1} равны

$$r_k = \frac{A_{k+1}/A_k - 1}{A_{k+1}/A_k + 1} \quad \text{или} \quad r_k = \frac{1 - A_k/A_{k+1}}{1 + A_k/A_{k+1}}$$

Показать, что, если A_k и A_{k+1} положительны, то $-1 \leq r_k \leq 1$.

3.5. При анализе влияния нагрузочного сопротивления излучения в модели с трубами без потерь предполагалось, что Z_L — действительная постоянная величина. Более точным является соотношение (3.296).

а) Задавая граничные условия $P_N(l_N, \Omega) = Z_L U_N(l_N, \Omega)$, найти соотношение между преобразованиями Фурье $u^-_N(t + \tau_N)$ и $u^+_N(t - \tau_N)$.

б) Используя соотношение в частотной области, полученное в а) и (3.296), показать, что $u^-_N(t + \tau_N)$ и $u^+_N(t - \tau_N)$ удовлетворяют обычному дифференциальному уравнению:

$$L_r \left(R_r + \frac{\rho c}{A_N} \right) \frac{du^-_N(t + \tau_N)}{dt} + \frac{\rho c}{A_N} R_r u^-_N(t + \tau_N) = \\ = L_r \left(R_r - \frac{\rho c}{A_N} \right) \frac{du^+_N(t - \tau_N)}{dt} - \frac{\rho c}{A_N} R_r u^+_N(t - \tau_N).$$

3.6. Подставляя (3.50) в (3.49), показать, что (3.48) и (3.49) эквивалентны.

3.7. Пусть имеется модель из двух труб (см. рис. 3.37). Написать для этой модели уравнения в частотной области и показать, что передаточная функция, связывающая скорости входного и выходного потоков, определяется выражением (3.51).

3.8. Рассмотрим идеальную двухсекционную модель без потерь для образования гласных (рис. 3.3.3). Предположим, что и со стороны голосовой щели, и со стороны губ потери отсутствуют. Для этих условий системная функция модели может быть получена из

(3.52) подстановкой $r_G = r_L = 1$ и $r_1 = (A_2 - A_1)/(A_2 + A_1)$.

а) Показать, что полюса системы на оси $i\Omega$ располагаются на частотах Ω , удовлетворяющих уравнению $\cos \Omega(\tau_1 + \tau_2) + r_1 \cos \Omega(\tau_2 - \tau_1) = 0$ или $(A_1/A_2) \tan(\Omega\tau_2) = \cot(\Omega\tau_1)$, где $\tau_1 = l_1/c$; $\tau_2 = l_2/c$ и c — скорость звука.

б) Значения частоты Ω , найденные решением уравнений а), являются формантными частотами модели с трубами без потерь. Путем соответствующего выбора параметров l_1 , l_2 , A_1 и A_2 можно аппроксимировать конфигурацию голосового тракта для гласных и, решая приведенные выше уравнения, получить формантные частоты модели. В табл. 3.3.1 даны параметры для нескольких гласных из [2]. Найти формантные частоты для каждого случая (не-

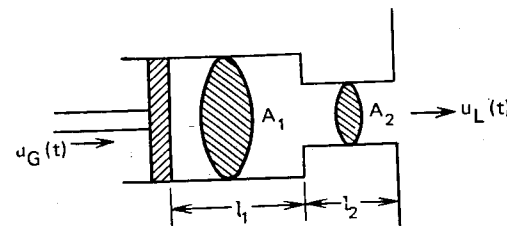


Рис. 3.3.3

Таблица 3.3.1

Гласный звук	l_1 , см	A_1 , см ²	l_2 , см	A_2 , см ²
$ i $	9	8	6	1
$ ae $	4	1	13	8
$ a $	9	1	8	7
$ \Lambda $	17	6	0	6

линейные уравнения можно решить графически или путем итераций). Принять $c=35000$ см/с.

3.9. Подставляя \hat{Q}_1 и \hat{Q}_2 в (3.75), показать, что передаточная функция дискретной модели голосового тракта с двумя трубами равна выражению (3.77).

3.10. Показать, что если $|a| < 1$, то

$$1 - az^{-1} = 1 \left/ \sum_{n=0}^{\infty} a^n z^{-n} \right.,$$

и, таким образом, наличие нуля в передаточной функции можно заменить бесконечным количеством полюсов.

3.11. Передаточная функция цифрового формантного резонатора

$$V_k(z) = \frac{1 - 2|z_k| \cos \theta_k + |z_k|^2}{1 - 2|z_k| \cos \theta_k z^{-1} + |z_k|^2 z^{-2}},$$

где $|z_k| = e^{-\sigma_k T}$ и $\theta_k = 2\pi F_k T$.

а) Изобразить полюса $V_k(z)$ на z -плоскости. Отметить соответствующие полюса аналоговой системы в s -плоскости.

б) Написать разностные уравнения, связывающие выходной сигнал $y_k(n)$, $V_k(z)$ и входной сигнал $x_k(n)$.

в) Изобразить схему цифровой системы формантного резонатора с тремя операциями умножения.

г) Используя разностные уравнения п. б), изобразить схему цифрового формантного резонатора с двумя операциями умножения.

3.12. Пусть имеется системная функция дискретной модели голосового тракта

$$V(z) = G \left/ \prod_{k=1}^N (1 - z_k z^{-1}) \right.$$

а) Показать, что $V(z)$ можно разложить на простые дроби:

$$V(z) = \sum_{k=1}^M \left[\frac{G_k}{1 - z_k z^{-1}} + \frac{G_k^*}{1 - z_k^* z^{-1}} \right],$$

где M равно $(N+1)/2$ с округлением до большего целого. Предполагается, что все полюса $V(z)$ — комплексные. Получить выражение для G_k .

б) Объединяя слагаемые в а), показать, что

$$V(z) = \sum_{k=1}^M \frac{B_k - C_k z^{-1}}{1 - 2|z_k| \cos \theta_k z^{-1} + |z_k|^2 z^{-2}}.$$

Это выражение задает параллельную форму реализации $V(z)$, где $z_k = |z_k| e^{i\theta_k}$. Получить формулы для B_k и C_k через G_k и z_k .

в) Изобразить схему параллельной реализации $V(z)$ при $M=3$.
г) Пусть задана полюсная система с передаточной функцией $V(z)$. Какая форма реализации потребует выполнения большего числа операций умножения — параллельная или каскадная (см. задачу 3.11)?

3.13. Взаимосвязь между звуковым давлением и скоростью потока около губ задается соотношением $P(l, s) = Z_L(s)U(l, s)$, где $P(l, s)$ и $U(l, s)$ — преобразования Лапласа $p(l, t)$ и $u(l, t)$ соответственно, $Z_L(s) = sR_r L_r / (R_r + sL_r)$, где $R_r = 128/9\pi^2$ и $L_r = 8a/3\pi c$. Здесь c — скорость звука; a — радиус отверстия между губами. В дискретной модели аналогичное соотношение имеет вид (3.97): $P_L(z) = R(z)U_L(z)$, где $P_L(z)$ и $U_L(z)$ есть z -преобразования $p_L(n)$ и $u_L(n)$ — дискретные аналоги звукового давления и скорости потока, спектр которых ограничен по частоте. Один из способов получения состоит в использовании билинейного преобразования [33]

$$R(z) = Z_L(s) \Big|_s = \frac{2}{T} \left[\frac{1-z^{-1}}{1+z^{-1}} \right].$$

а) Для заданного $Z_L(s)$ определить $R(z)$.

б) Записать разностное уравнение для $p_L(n)$ и $u_L(n)$.

в) Найти полюса и нули $R(z)$.

г) Если $c=35000$ см/с, $T=10^{-4}$ с, $0,5$ см $< a < 1,3$ см, то где располагаются полюса?

д) Простым приближением $R(z)$ является соотношение $R(z) = R_0(1 - z^{-1})$. Для $a=1$ см и $T=10^{-4}$ найти R_0 , при котором $R(-1) = Z_L(\infty) = R(-1)$.

ж) Нарисовать частотные зависимости $Z_L(\Omega)$, $R(e^{i\Omega T})$ и $R(e^{i\Omega T})$ как функции Ω для $a=1$ см и $T=10^{-4}$, $0 \leq \Omega \leq \pi/T$.

3.14. Простое приближение импульса голосовой щели изображено на рис. 3.3.4а.

а) Найти z -преобразование $G_1(z)$ этой последовательности. (Указание: $g_1(n)$ можно записать как свертку двух последовательностей $p(n)$:

$$p(n) = \begin{cases} 1, & 0 \leq n \leq N-1; \\ 0, & \text{в противном случае.} \end{cases}$$

б) Показать полюса и нули $G_1(z)$ на z -плоскости при $N=10$.

в) Изобразить амплитудный спектр $g_1(n)$ как функцию ω . Рассмотрим приближение импульса голосовой щели рис. 3.3.4б.

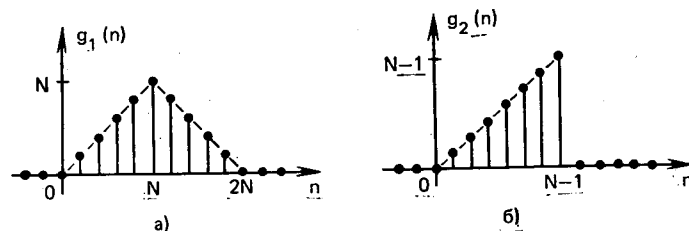


Рис. 3.3.4

г) Показать, что z -преобразование $G_2(z)$ равно

$$G_2(z) = z^{-1} \sum_{n=0}^{N-2} (n+1) z^{-n} = z^{-1} \left[\frac{1 - N z^{-(N-1)} + (N-1) z^{-N}}{(1 - z^{-1})^2} \right].$$

(Указание: z -преобразование $nx(n)$ равно $-z \frac{dX(z)}{dz}$)

д) Показать, что в общем случае $G_2(z)$ должно иметь по крайней мере один нуль вне единичного круга. Найти нули $G_2(z)$ для $N=4$.