

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
**«ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ
И РАДИОЭЛЕКТРОНИКИ» (ТУСУР)**

Г. Н. Глазов

КОНСПЕКТ ЛЕКЦИЙ
Современные технологии и системы
автоматизированного измерения на СВЧ

для студентов, обучающихся по специальности 210302
«Радиотехника»

2012

ВВЕДЕНИЕ

Параметры курса

Название и предмет курса: Радиоизмерительные приборы СВЧ

Распределение времени: лекций – 30 ч,

лаб. работы – 12 ч,

самост. работа – 28 ч.

Организация занятий и контроль

Содержание лекционной части

Цели и задачи курса

Классификация радиоизмерительных приборов по ГОСТ (см. Данилин)

Подгруппы приборов разделяются по функциональному признаку, обозначаются буквой и числом. Например, подгруппа Р – приборы для измерения параметров цепей с распределенными параметрами:

Р1 – измерительные линии;

Р2 – измерители КСВ и ослабления;

Р3 – измерители комплексных коэффициентов отражения и полных сопротивлений нагрузок СВЧ;

Р4 – измерители комплексных параметров матрицы рассеяния – коэффициентов передачи и коэффициентов отражения.

Подгруппа М – измерители мощности в диапазоне СВЧ:

М2 – ваттметры поглощаемой мощности;

М3 – ваттметры проходящей мощности;

М5 – измерительные преобразователи, входящие в состав ваттметров (болومترические или термисторные головки).

Виды приборов, изучаемых в курсе

Анализаторы цепей (скалярные и векторные), спектроанализаторы, измерители мощности, измерители коэффициента шума.

Литература

Основная литература

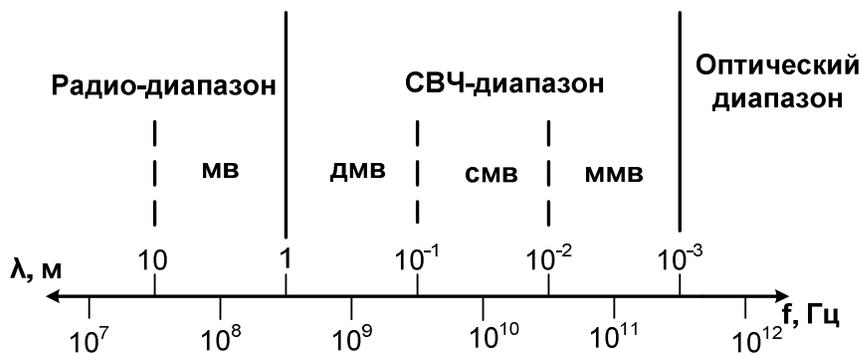
1. Измерения в электронике. Справочник. Под ред. В. А. Кузнецова. – М.: Энергоатомиздат, 1987.
2. М. Хибель. Основы векторного анализа цепей. Пер. с англ. С. М. Смольского. М., Изд. дом МЭИ, 2009.
3. Раушер К. Основы спектрального анализа. Пер. с англ. С. М. Смольского. М., 2006).
4. Данилин А. А. Измерения в технике СВЧ. М., «Радиотехника», 2008.
5. Bryant G. H. Principles of microwave measurements. Peregrinus. Переводы на русский:
Глава 5. Векторные анализаторы цепей (Перевод Глазова Г. Н.)
Глава 6. Скалярные анализаторы цепей (Перевод Глазова Г. Н.)
Глава 8. Измерение мощности на СВЧ (Перевод Глазова Г. Н.)

6. Гусинский А. В., Шаров Г. А., Кострикин А. М. Векторные анализаторы цепей миллиметровых волн. Часть 3, книга 1. Минск, 2008.
7. Гусинский А. В., Шаров Г. А., Кострикин А. М. Векторные анализаторы цепей миллиметровых волн. Часть 3, книга 2. Минск, 2008.
8. Глазов Г. Н. Измерение мощности радиочастотных и микроволновых сигналов. Компиляция. Томск, МИКРАН, 2004.
9. Глазов Г. Н. Характеризация частотной стабильности сигналов. Научно-технический отчет. Томск, 2003.
10. Peterson V. Spectrum analysis basics. Agilent Technologies, Appl. Note 150.
(Перевод на русский: Основы анализа спектра. Под ред. Глазова Г. Н., Панимаскина Е. И., пер. Пинжина П. Ю.).
11. Алмазов К. И. Коэффициент шума и его измерение на СВЧ. – М.: Научный мир, 2000.
12. Кудрявцев А. М. и др. Радиоизмерительная аппаратура СВЧ и КВЧ: узловая и элементная базы. Под ред. А. М. Кудрявцева.– М.: изд. «Радиотехника», 2006.

Диапазон СВЧ

Место диапазона СВЧ в спектре электромагнитных волн

К *диапазону сверхвысоких частот (СВЧ)* обычно относят волны частотой от 300 МГц до 300 ГГц, то есть длиной волны от 1 м до 1 мм. Этот диапазон условно делят на поддиапазоны *дециметровых волн* (частоты от 300 МГц до 3 ГГц, длины волн от 0.1 м до 1 м), *сантиметровых волн* (частоты от 3 ГГц до 30 ГГц, длины волн от 0.01 м до 0.1 м), *миллиметровых волн* (частоты от 30 ГГц до 300 ГГц, длины волн от 0.001 м до 0.01 м). Место диапазона СВЧ в общем спектре электромагнитных волн показано на рис. 1.1.1.



Диапазоны электромагнитных волн

Особенности диапазона СВЧ

В дальнейшем основное внимание будет уделено измерениям в дециметровом и сантиметровом поддиапазонах, т.к. техника миллиметровых волн отличается большой специфичностью и требует отдельного рассмотрения. С другой стороны, по практическим причинам, рабочий диапазон *анализаторов цепей (АЦ)*, которым посвящена эта глава, обычно шире указанного объединения поддиапазонов за счет спектрального участка *метровых волн* (частоты от 30 МГц до 300 МГц, длины волн от 1 м до 10 м), который формально не относится к диапазону СВЧ, а является частью диапазона УКВ. Иногда нижняя частота рабочего диапазона АЦ еще ниже и доходит до 10 МГц и менее.

В диапазоне СВЧ квантовые эффекты еще пренебрежимы, поэтому закономерности протекания электромагнитных явлений в различных средах составляют предмет изучения *классической электродинамики*. Ее технические аспекты составляют область науки и техники, иногда называемую *технической электродинамикой*. В этой области в основном изучается и разрабатывается *техника СВЧ*, которая объединяет разделы радиотехники и радиоэлектроники, касающиеся вопросов анализа, синтеза и практической реализации *устройств СВЧ*.

Устройства СВЧ, соединенные отрезками *линий передачи*, составляют *тракт СВЧ*. Тракт СВЧ любой радиотехнической системы, в том числе анализатора цепей или анализируемого устройства, состоит из большого числа различных устройств СВЧ. К их числу относятся отрезки линий передачи, разъемы, изгибы и скрутки, согласующие устройства, фазовращатели, фильтры, делители мощности, направленные ответвители, переключатели и многие другие. Общим для этих и им подобных устройств является то, что они относятся к *устройствам с распределенными параметрами*. Геометрические размеры этих устройств сравнимы с длиной волны электромагнитных колебаний. Это определяет всю специфику расчета и проектирования устройств СВЧ, т.к. происходящие в них процессы имеют *волновой характер*.

Теория устройств СВЧ тесно связана с электродинамикой и включает в себя два больших раздела: *анализ устройств СВЧ* и *синтез устройств СВЧ*. Задача анализа состоит в изучении внешних характеристик устройств СВЧ, а также в определении этих внешних характеристик из решения соответствующей внутренней задачи методами прикладной электродинамики или из эксперимента. Задача синтеза устройств СВЧ состоит в определении структуры и геометрических размеров элементов устройства СВЧ по заданным его характеристикам.

Изучение внешних характеристик устройств СВЧ может производиться без конкретизации их внутренней структуры, что позволяет рассматривать устройство СВЧ как некий «черный ящик», имеющий определенное число выходящих из него линий передачи СВЧ. Каждая из этих линий передачи также является устройством с распределенными параметрами, для которого неизменным является волновой характер электромагнитных процессов. Это приводит к необходимости фиксировать продольные координаты поперечных сечений линий передачи или, как говорят, фиксировать *опорные плоскости*. Относительно этих опорных плоскостей проводится отсчет фаз, а в некоторых случаях и амплитуд падающих и отраженных волн. Смещение опорных плоскостей вдоль входных линий передачи приводит к изменению внешних характеристик устройств СВЧ. В большинстве случаев во входных линиях передачи таких устройств единственной распространяющейся волной является волна основного типа. Остальные типы волн находятся в закритическом режиме, то есть быстро затухают при отдалении от неоднородности, где они могут возникнуть. Опорные плоскости устройства СВЧ стремятся расположить таким образом, чтобы амплитудами закритических волн в них можно было пренебречь.

Основой теоретического и расчетного анализа и проектирования СВЧ-устройств является *теория цепей СВЧ*. В силу упомянутого выше представления цепи в виде «черного ящика», эта теория основана на *матричном представлении* описания внешних параметров устройств, использует *метод декомпозиции* анализа сложных объектов и компьютерные методы их моделирования. Для расчета параметров базовых блоков *декомпозиционных схем* привлекаются методы *прикладной электродинамики* и *численные модели* систем автоматического проектирования (САПР) СВЧ-устройств. При практической реализации ключевую роль играют *экспериментальные методы* исследования, настройки и контроля характеристик и параметров спроектированных устройств. Измерения используют также в случаях, когда для моделирования отдельных узлов и блоков нет достоверных расчетных данных.

Свойства электромагнитного поля существенно зависят от характера его изменения во времени; для модели монохроматических волн – от частоты волны. Особенности техники измерений на СВЧ можно условно разделить на особенности излучения, распространения, отражения и поглощения СВЧ-волн и особенности СВЧ-техники применительно к измерениям. Освое-

ние СВЧ-диапазона началось позже, чем соседних диапазонах – только в 30-х годах двадцатого века, что объясняется особенностями СВЧ-волн, к важнейшим из которых относятся:

- Соизмеримость длины волны λ и характерного размера приборов, неоднородностей тракта и других, что делает малоэффективным применение обычных линий передачи и колебательных контуров. Традиционные элементы аппаратуры – соединительные проводники, резонансные контуры, конденсаторы, катушки индуктивности и т.д. становятся распределенными системами. Вместо соединительных проводников используют линии передачи конечной длины, колебательные системы представляют собой объемные резонаторы; сосредоточенные компоненты заменяются различными неоднородностями в линии передачи, расчет которых проводят электродинамическими методами.

- Соизмеримость периода колебаний T и времени пролета носителей заряда в активной области прибора τ . Проявляющаяся в этих условиях инерция носителей заряда нарушает нормальную работу активных приборов, эффективно функционирующих в радиодиапазоне.

- Малая энергия кванта электромагнитного поля $E_k = hf$ по сравнению с энергией теплового движения частиц, заданной в единицах $E_T = kT$, при обычной температуре, исключает проявление и возможное использование квантовых эффектов, успешно используемых в оптическом диапазоне. Например, на частоте $f = 20$ ГГц $E_k = 1.33 \cdot 10^{-23}$ Дж, а при температуре $T = 290$ К $E_T = 4 \cdot 10^{-21}$ Дж, то есть их отношение $E_T / E_k = 300$. Интересно отметить, что этот вывод меняется при очень низких температурах, например, на той же частоте, но температуре жидкого гелия $T = 4.2$ К отношение $E_T / E_k = 4.34$, а при частоте 50 ГГц даже $E_T / E_k = 1.74$.

Для реализации генерации, усиления и преобразования электромагнитных колебаний СВЧ-диапазона необходимо было разработать специальные методы, что потребовало значительных усилий многочисленных коллективов инженеров и ученых. В частности, при переходе на измерения в диапазоне СВЧ потребовалось, в ряде случаев, разработать специальные измерительные методы и средства. Измерения на СВЧ, по сравнению с традиционными измерениями на более низких частотах, имеют свои особенности:

- Ряд проблем вызывает подключение измерительной аппаратуры к исследуемому устройству и элементам калибровки. Пришлось разработать элементы подключения измерительных датчиков и преобразователей к соединительным линиям, вообще, в большой мере заново создавать измерительный тракт.

- Излучение электромагнитных волн из отверстий и щелей в измерительной аппаратуре приводит к потерям и искажениям сигналов. Требуется тщательное экранирование измерительных элементов и датчиков, что, однако осложняет введение измерительных элементов в исследуемые объекты.

- Отражения на соединениях, неоднородностях, сгибах, концах линий и т.д. (*несогласованность*) вызывают искажение сигналов, ухудшают точность измерений, требуют изолированной калибровки измерительного средства. Вообще, *проблема согласования* «нагрузок» в диапазоне СВЧ настолько всеобъемлюща, что иногда высказывается мнение, что «проблема оптимального конструирования тракта СВЧ сводится к проблеме согласования» [1.1].

- Неоднозначность отсчета напряжений и токов на СВЧ заставляет переходить к мощности как единице уровня сигналов.

- В технике СВЧ необходимо измерять новые параметры – КСВ, коэффициент отражения и т.д. Вообще, в рамках матричной характеристики цепей приходится переходить от матриц, связанных с импедансами, токами и напряжениями, к матрицам рассеяния; появляются также новые виды графов – *потокосые графы*, и т.д.

- Измерениям на СВЧ свойственны ограничения технического порядка, связанные с несовершенством аппаратуры на этих частотах. В частности, трудно или невозможно добиться достаточно высоких добротностей резонаторов, трудно создавать измерительные генераторы достаточной мощности со сверхширокополосной перестройкой, трудно обеспечить высокую частотную стабильность генераторов, трудно создавать измерительные усилители со стабильными метрологическими характеристиками и т.д.

- В диапазоне СВЧ ограничен выбор невзаимных устройств, позволяющих разделять волны разных направлений и улучшать согласование трактов. На малых частотах построение невзаимных устройств облегчается возможностями различных соединений в цепях, в диапазоне СВЧ, с его полевой структурой волн, приходится применять специальные волновые среды (например, ферриты) или комбинации линий передачи с полевой связью.

В силу перечисленных особенностей измерений на СВЧ, для решения каждой конкретной измерительной задачи в этом диапазоне часто приходится индивидуально подбирать или заново разрабатывать свою собственную методику, а в ряде случаев и уникальную измерительную аппаратуру. Всегда требуется предварительный теоретический и расчетный анализ измерительной схемы, особенно в отношении калибровки и оценки погрешностей измерения.

Вышеперечисленные особенности СВЧ-волн могут создать впечатление, что сам диапазон «невыгоден» для радиоэлектронных приложений и в измерениях на СВЧ, в частности, в анализе СВЧ-цепей нет большой необходимости. Это впечатление в корне неверно: практически диапазон СВЧ очень важен, а в некоторых случаях – незаменим, для целого ряда приложений, вследствие важных преимуществ СВЧ-волн. Эти преимущества таковы:

- Как известно, информационная емкость канала связи пропорциональна произведению полосы пропускания канала Δf на отношение сигнал-шум по мощности P_S / P_N . Полосу пропускания можно оценить как 1-10% от несущей частоты f_0 , следовательно, чем выше несущая частота, тем выше информационная емкость канала связи. Поэтому в СВЧ-диапазоне информационная емкость канала, в частности, скорость передачи информации, при том же отношении сигнал-шум, в сотни и тысячи раз больше, чем в радиодиапазоне. Но этот выигрыш еще больше вследствие минимума шумов атмосферного, космического и искусственного происхождения в диапазоне 1-10 ГГц, что повышает отношение сигнал-шум.

- С увеличением частоты возрастают возможности направленной передачи электромагнитной энергии, так как они зависят от отношений диаметра антенны к длине волны. В диапазоне СВЧ эти возможности, важные для многих приложений, значительно выше, чем в радиодиапазоне.

- В целом земная атмосфера весьма прозрачна в диапазоне СВЧ; это является ценным энергетическим фактором для многочисленных радиоэлектронных приложений.

- Ряд веществ проявляет селективно-частотное (избирательное) поглощение СВЧ-излучения, особенно при низких температурах, вследствие полосового, линейчатого характера спектра поглощения на этих частотах, а также, в ряде случаев, высоких градиентов поглощения по частоте. Это создает ряд возможностей для создания оригинальных устройств техники СВЧ.

Радиотехнические системы, работающие в диапазоне СВЧ

Отмеченные преимущества диапазона обусловили широкое применение волн СВЧ в различных областях науки и техники. Перечислим некоторые из них:

- Радиолокация – первая исторически и до сих пор наиболее важная область применения колебаний СВЧ. Эта область в последние десятилетия испытала большой прогресс, как в военном, так и научном и народно-хозяйственном аспектах. Многообразие объектов локации (сосредоточенные: наземные, надводные, воздушные, космические; распределенные: элементы поверхности Земли, облака, полярные льды, северные сияния, метеорные ионизированные следы,

планеты Солнечной системы и т.д.), многообразие способов формирования и обработки радиолокационных сигналов (когерентная радиолокация, радиолокация бокового обзора, сжатие сигналов и т.д.), многообразие носителей радиолокационной техники (наземные, надводные, воздушные, космические) и степени их подвижности, многообразие антенных систем управления лучами (одиночные антенны, синтезированные антенны, пассивные и активные антенные решетки) и видов их сканирования – все это и многое другое предопределило огромное поле для создания разнообразных устройств и систем техники СВЧ, нуждающихся в их тестировании и измерении.

- Радиорелейная связь, обеспечивающая передачу пакетов телефонных, телевизионных каналов и профессиональной информации различного назначения на различные, в том числе – большие расстояния с высокой скоростью передачи.
- Кабельная связь, в том числе распределенные системы вещательного телевидения, промышленного и охранного телевидения.
- Космическая связь, то есть передача больших объемов информации через спутники связи (непосредственное телевизионное вещание, связь с космическими кораблями и орбитальными станциями, с автоматическими космическими аппаратами).
- Радионавигация.
- Радиоастрономия (прием и анализ электромагнитного излучения космических объектов).
- Радиоразведка и радиоконтроль.
- Экспериментальная физика (нагрев и диагностика плазмы, молекулярная спектроскопия газов, жидкостей и твердых тел).
- Мазеры и их применение.
- Метрология (атомные стандарты времени и частоты).
- Технология (использование СВЧ-излучения для нагрева различных материалов с целью их сушки и полимеризации, приготовления пищи и т.п.).
- Медицина и биология (воздействие СВЧ-колебаний на биологические объекты с целью изучения и коррекции процессов в них, анализ собственного излучения объектов).

Особенности измерений на СВЧ

Измерения на СВЧ, по сравнению с традиционными измерениями на более низких частотах, имеют свои особенности. В первую очередь, они являются следствиями свойств СВЧ волн, их распространения и преобразования.

В диапазоне СВЧ квантовые эффекты еще пренебрежимы, поэтому закономерности протекания электромагнитных явлений в различных средах составляют предмет изучения *классической электродинамики*. Ее технические аспекты составляют область науки и техники, иногда называемую *технической электродинамикой*. В этой области в основном изучается и разрабатывается *техника СВЧ*, которая объединяет разделы радиотехники и радиоэлектроники, касающиеся вопросов анализа, синтеза и практической реализации *устройств СВЧ*.

Устройства СВЧ, соединенные отрезками *линий передачи*, составляют *тракт СВЧ*. Тракт СВЧ любой радиотехнической системы, в том числе анализатора сигналов, анализатора цепей или анализируемого устройства, состоит из большого числа различных устройств СВЧ. К их числу относятся отрезки линий передачи, разъемы, изгибы и скрутки, согласующие устройства, фазовращатели, фильтры, делители мощности, направленные ответвители, переключатели и многие другие. Общим для этих и им подобных устройств является то, что они относятся к *устройствам с распределенными параметрами*. Геометрические размеры этих устройств срав-

нимы с длиной волны электромагнитных колебаний. Это определяет всю специфику расчета и проектирования устройств СВЧ, т.к. происходящие в них процессы имеют *волновой характер*.

Теория устройств СВЧ тесно связана с электродинамикой и включает в себя два больших раздела: *анализ устройств СВЧ* и *синтез устройств СВЧ*. Задача анализа состоит в изучении внешних характеристик устройств СВЧ, а также в определении этих внешних характеристик из решения соответствующей внутренней задачи методами прикладной электродинамики или из эксперимента. Задача синтеза устройств СВЧ состоит в определении структуры и геометрических размеров элементов устройства СВЧ по заданным его характеристикам.

Изучение внешних характеристик устройств СВЧ может производиться без конкретизации их внутренней структуры, что позволяет рассматривать устройство СВЧ как некий «черный ящик», имеющий определенное число выходящих из него линий передачи СВЧ. Каждая из этих линий передачи также является устройством с распределенными параметрами, для которого неизменным является волновой характер электромагнитных процессов. Это приводит к необходимости фиксировать продольные координаты поперечных сечений линий передачи или, как говорят, фиксировать *опорные плоскости*. Относительно этих опорных плоскостей проводится отсчет фаз, а в некоторых случаях и амплитуд падающих и отраженных волн. Смещение опорных плоскостей вдоль входных линий передачи приводит к изменению внешних характеристик устройств СВЧ. В большинстве случаев во входных линиях передачи таких устройств единственной распространяющейся волной является волна основного типа. Остальные типы волн находятся в закритическом режиме, то есть быстро затухают при отдалении от неоднородности, где они могут возникнуть. Опорные плоскости устройства СВЧ стремятся расположить таким образом, чтобы амплитудами закритических волн в них можно было пренебречь.

Основой теоретического и расчетного анализа и проектирования СВЧ-устройств является *теория цепей СВЧ*. В силу упомянутого выше представления цепи в виде «черного ящика», эта теория основана на *матричном представлении* описания внешних параметров устройств, использует *метод декомпозиции* анализа сложных объектов и компьютерные методы их моделирования. Для расчета параметров базовых блоков *декомпозиционных схем* привлекаются методы *прикладной электродинамики* и *численные модели* систем автоматического проектирования (САПР) СВЧ-устройств. При практической реализации ключевую роль играют *экспериментальные методы* исследования, настройки и контроля характеристик и параметров спроектированных устройств. Измерения используют также в случаях, когда для моделирования отдельных узлов и блоков нет достоверных расчетных данных.

Свойства электромагнитного поля существенно зависят от характера его изменения во времени; для модели монохроматических волн – от частоты волны. Особенности техники измерений на СВЧ можно условно разделить на особенности излучения, распространения, отражения и поглощения СВЧ-волн и особенности СВЧ-техники применительно к измерениям.

Фундаментальные особенности СВЧ волн это, во-первых, соизмеримость длины волны и характерного размера приборов и неоднородностей тракта, что делает малоэффективным применение обычных линий передачи и колебательных контуров; во-вторых, соизмеримость периода колебаний и времени пролета носителей заряда в активной области прибора; проявляющаяся в этих условиях инерция носителей заряда нарушает нормальную работу активных приборов, эффективно функционирующих в радиодиапазоне.

Измерения на СВЧ сталкиваются с рядом проблем. «Вечная» проблема не только измерений, но и техники СВЧ в целом – качественное, надежное, устойчивое подключение измерительной аппаратуры к исследуемому устройству и элементам калибровки. В историческом развитии пришлось разработать элементы подключения измерительных датчиков и преобразователей к соединительным линиям, вообще, в большой мере заново создавать измерительный тракт.

Излучение электромагнитных волн из отверстий и щелей в измерительной аппаратуре приводит к потерям и искажениям сигналов. Требуется тщательное экранирование измеритель-

ных элементов и датчиков, что, однако осложняет введение измерительных элементов в исследуемые объекты.

Отражения на соединениях, неоднородностях, сгибах, концах линий и т.д. (*несогласованность*) вызывают искажение сигналов, ухудшают точность измерений, требуют изощренной калибровки измерительного средства. Вообще, *проблема согласования* «нагрузок» в диапазоне СВЧ настолько всеобъемлюща, что иногда высказывается мнение, что «проблема оптимального конструирования тракта СВЧ сводится к проблеме согласования» [0.1].

Неоднозначность отсчета напряжений и токов на СВЧ заставляет переходить к мощности как характеристике уровня сигналов.

В технике СВЧ необходимо измерять новые параметры – КСВ, коэффициент отражения и т.д. Вообще, в рамках матричной характеристики цепей приходится переходить от матриц, связанных с импедансами, токами и напряжениями, к матрицам рассеяния; появляются также новые виды графов – *потокосые графы*, и т.д.

Измерениям на СВЧ свойственны ограничения технического порядка, связанные с несовершенством аппаратуры на этих частотах. В частности, трудно или невозможно добиться достаточно высоких добротностей резонаторов, трудно создавать измерительные генераторы достаточной мощности со сверхширокополосной перестройкой, трудно обеспечить высокую частотную стабильность генераторов, трудно создавать измерительные усилители со стабильными метрологическими характеристиками и т.д.

В диапазоне СВЧ ограничен выбор невзаимных устройств, позволяющих разделять волны разных направлений и улучшать согласование трактов. На малых частотах построение невзаимных устройств облегчается возможностями различных соединений в цепях, в диапазоне СВЧ, с его полевой структурой волн, приходится применять специальные волновые среды (например, ферриты) или комбинации линий передачи с полевой связью.

В силу перечисленных особенностей измерений на СВЧ, для решения каждой конкретной измерительной задачи в этом диапазоне часто приходится индивидуально подбирать или заново разрабатывать свою собственную методику, а в ряде случаев и уникальную измерительную аппаратуру. Всегда требуется предварительный теоретический и расчетный анализ измерительной схемы, особенно в отношении калибровки и оценки погрешностей измерения.

В диапазоне СВЧ углубляется различие между двумя сторонами измерений на СВЧ – анализом сигналов и анализом устройств.

Портовое представление цепей

Радиосистемы СВЧ-диапазона (а также более высокочастотные) обычно можно представить в виде некоторых устройств, соединенных линиями передачи. Часть такой системы, расположенную между начальным и оконечным устройствами (например, между антенной и радиопередающим или радиоприемным устройством, называют *трактом СВЧ* (иногда *цепью СВЧ*). Подобный тракт осуществляет передачу электромагнитной энергии от передатчика к антенне, или от антенны к приемнику, или от генератора сигналов специальной формы ко входу некоторого устройства и т.п. Тракт СВЧ может, например, обеспечивать требуемый режим работы выходных и входных цепей передатчика или приемника, выполнять частотное и поляризационное разделение и объединение передаваемых сигналов и ряд других функций. Отдельные части тракта также могут называться цепями СВЧ. Наиболее распространенными элементами СВЧ-цепей являются отрезки линий передачи, переходные и стыковочные узлы между линиями различных типов, согласующие и настроечные элементы, сумматоры, делители и ответвители мощности, поляризационные устройства, фильтры, фазовращатели, коммутаторы и переключатели, невзаимные устройства с намагниченными ферритами и др.

Процессы передачи электромагнитных сигналов в цепях СВЧ и в образующих их элементах весьма сложны. Можно было бы попытаться проанализировать эти процессы на основе со-

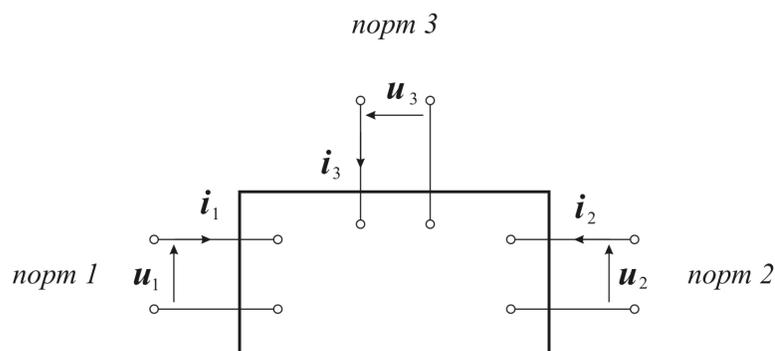
ответствующих *краевых задач электродинамики*. Однако строгая постановка и решение таких задач возможны только для очень упрощенных форм и конструкций, почти всегда выходящих за рамки реальных. А для применяемых на практике цепей СВЧ из-за их конфигурационной сложности решение краевых задач в строгой постановке в настоящее время практически невозможно.

Некоторую помощь в решении проблемы передачи электромагнитных сигналов в сколь-нибудь сложных цепях СВЧ оказывает всегда применяемый на практике *метод декомпозиции* (разбиения). Он состоит в том, что тракт СВЧ или цепь СВЧ разбивается на ряд *элементов*, которые анализируются независимо, но теперь также нужно адекватно учесть взаимодействия между элементами через распространяющиеся между ними волны. При этом каждый такой элемент рассматривается как независимая электродинамическая система.

Постановка и решение краевых электродинамических задач, соответствующих отдельным элементам, существенно проще, чем для всего устройства в целом. Используя или решение электродинамической задачи или результаты экспериментального исследования, если подобное решение получить не удастся, для каждого выделенного элемента строят такое описание, которое позволяет находить влияние этого элемента на передаваемые электромагнитные сигналы. Обычно описание элементов цепи представляют либо формально-математически с помощью матриц специального вида (матрицы рассеяния, или матрицы передачи и т.п.), либо в виде *эквивалентной схемы*, состоящей из отрезков эквивалентной линии передачи, в которую тем или иным способом включены сосредоточенные элементы L, C, R и трансформаторы. Имея подобные универсальные описания всех элементов тракта или цепи СВЧ, можно определить все требуемые характеристики тракта или цепи.

Обычно при построении *математической модели* тракта или цепи СВЧ для упрощения анализа отрезки линий передачи, соединяющие входящие в этот тракт устройства заменяют отрезками *эквивалентной линии*, а устройства рассматриваются как некоторые *многополюсники*.

Сделаем одно терминологическое замечание. Понятия полюсов, клемм и т.п., так хорошо подходившие к концепции «черного ящика» на достаточно низких частотах, где цепи строятся на сосредоточенных элементах, а соединения производятся с помощью проводов, становятся неудобными в диапазоне СВЧ, где по крайней мере линии передачи суть распределенные элементы, а соединения отрезков этих линий производятся с помощью коаксиальных разъемов, фланцевых соединений волноводов и т.п. В этом диапазоне больше подходит понятие *порта*, которым и пользуются во всей научно-технической радиоэлектронной литературе, кроме русскоязычной. С точки зрения формально-математического описания порт полностью эквивалентен паре полюсов или клемм, но физическая интерпретация этого понятия более прозрачна. В этой работе мы будем пользоваться именно этим понятием.



Токи и напряжения в трехпортовом устройстве

ЭКВИВАЛЕНТНЫЕ ЛИНИИ ПЕРЕДАЧИ

Предисловие: электродинамика ЛП

Тракт СВЧ, согласование, классификация ЛП

Измерительные приборы СВЧ, в частности, векторный и скалярный анализаторы цепей, спектроанализатор и др., представляют собой сложную *радиотехническую систему*. Со структурной точки зрения такая система включает большое число различных *устройств СВЧ*. К ним относятся отрезки линий передачи, разъемы, изгибы и скрутки, согласующие устройства, фильтры СВЧ, делители мощности, направленные ответвители, коммутирующие устройства и т.п. Среди устройств СВЧ особое место занимают *линии передачи* волн СВЧ. Они не только соединяют устройства СВЧ между собой в *тракт СВЧ* и канализируют электромагнитную энергию, несущую информацию, но, в определенном смысле, являются координаторами конструкции и «законодателями мод»: в схемотехническом смысле «основные средства для конструирования ... находятся из понимания передающих линий ... и методов согласования импедансов» [13.6]. Этот тезис как нельзя лучше относится к векторному анализатору цепей, который должен быть сконструирован так, чтобы информативные, подлежащие измерению, потоки в линиях передачи минимально искажались, а неинформативные, обязанные факторам несогласованности, максимально подавлялись.

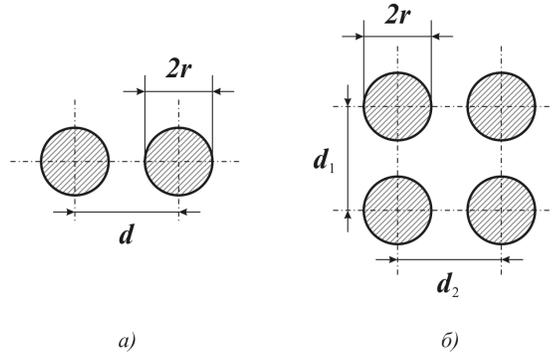
Строго говоря, для расчета и конструирования линии передачи необходимо знать структуру электромагнитного поля в рассматриваемой части пространства, т. е. решить прямую внутреннюю краевую задачу электродинамики. Однако, возникающие на практике электродинамические задачи обычно весьма сложны, поэтому прибегают к рассмотрению некоторой модельной задачи, в той или иной мере отражающей реальную ситуацию. Применительно к линиям передачи такие модели доставляет концепция *эквивалентных длинных линий*, применение которой каждый раз должно иметь некоторые основания.

С позиций электродинамики линия передачи (ЛП) это *направляющая система*, т. е. устройство канализации электромагнитной энергии в определенном направлении, называемом ниже *продольным*, без излучения в окружающее пространство. Направление распространения определяется взаимным расположением *источника* электромагнитных волн и *нагрузки* в ЛП. Источником может служить генератор, подключенный к ЛП, приемная антенна, *элемент связи* с другой ЛП или с неким устройством СВЧ, и т. п. Нагрузкой ЛП может служить сопротивление (в общем случае комплексное), передающая антенна, входная цепь приемника и т. п.

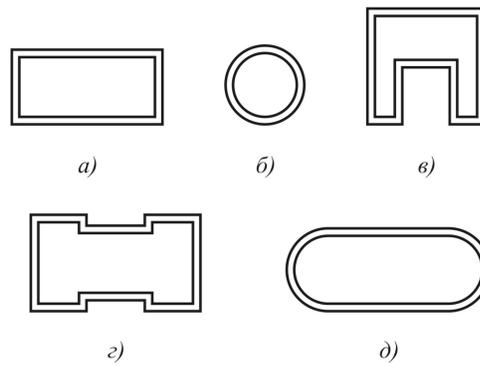
Существует некоторая классификация ЛП. Если электромагнитное поле ЛП не ограничено в поперечном направлении, ее называют *открытой*, в противном случае – *закрытой*; обычно ограничителем является некая металлическая оболочка той или иной формы. В открытых ЛП, вообще говоря, электромагнитное поле распределено во всем пространстве, окружающем линию, поэтому она конструируется так, чтобы подавляющее большинство передаваемой энергии было сосредоточено в непосредственной близости к линии. Все же открытые линии подвержены влиянию внешних воздействий: электромагнитных полей, созданных другими источниками, метеоусловий и т. п. Закрытые ЛП иногда называют волноводами [13.4], в такой терминологии *коаксиальную* ЛП следует относить к волноводам специального типа. Мы будем считать коаксиальные ЛП отдельным видом. ЛП называется *регулярной*, если ее геометрические и электродинамические свойства не меняются в продольном направлении (*продольно-регулярная ЛП*) или меняются по периодическому закону (*периодическая ЛП*), в противном случае ЛП называется *нерегулярной*. Ниже рассматриваются только продольно-регулярные ЛП.

В зависимости от материалов тел, образующих ЛП, они могут быть *металлическими*, *диэлектрическими* или *металлодиэлектрическими*. Электрические свойства ЛП зависят от *порядка связности* их поперечного сечения; он является геометрической характеристикой поперечного сечения и определяется числом проводящих поверхностей ЛП, в зависимости от которого

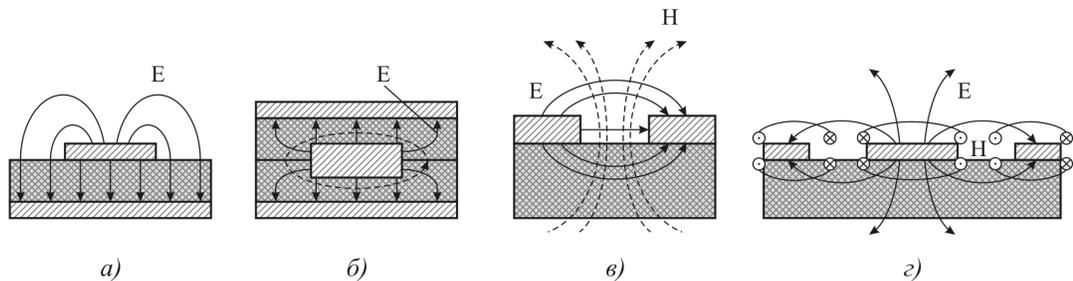
различают ЛП *нулевой связности* (при отсутствии проводящих поверхностей, например, диэлектрические ЛП), *односвязные* (например, металлические волноводы), *двухсвязные* (например, коаксиальные ЛП), и т. д. С точки зрения рассмотренных классификационных признаков, волновод – закрытая односвязная жесткая металлическая ЛП, коаксиальная линия – закрытая двухсвязная гибкая металлодиэлектрическая ЛП, *микророскопическая линия* – открытая двухсвязная жесткая металлодиэлектрическая ЛП.



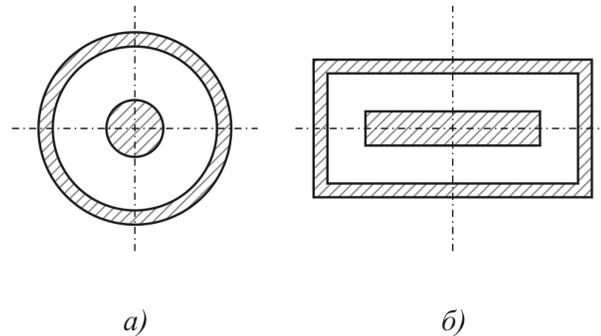
Проволочная линия передачи; а) – из двух проводов; б) – из четырех проводов



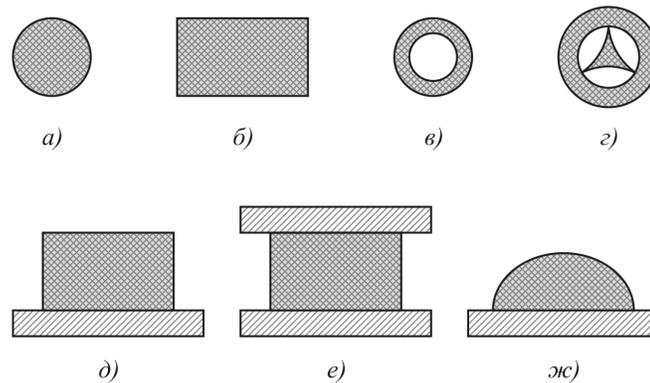
Виды сечений волноводных ЛП



Виды полосковых линий и характер их полей



Коаксиальные ЛП; а)–круглые проводники; б)–прямоугольные проводники



Виды сечений диэлектрических ЛП

Продолжим классификацию ЛП по конструктивному признаку, перечислив наиболее часто применяемые *виды ЛП*. *Проволочная ЛП* – комплекс из двух или четырех проводников, иногда покрытых диэлектриком (рис. 13.1.1). *Волноводная ЛП* имеет одну замкнутую проводящую поверхность, имеющую вид цилиндра (трубы) прямоугольного, круглого или эллиптического сечения. (рис. 13.1.2). *Полосковые ЛП* – *симметричная, несимметричная, щелевая, копланарная, микрополосковая* (рис. 13.1.3). Последняя отличается тем, что ее *диэлектрическая подложка* имеет большую относительную диэлектрическую проницаемость ϵ_r (более 10) и малые потери, вследствие чего геометрические размеры устройств, выполненных на основе таких линий, уменьшаются в $\sqrt{\epsilon_r}$ раз. Материалом диэлектрической подложки в таких линиях служат поликор, ситалл, кремний, сапфир и др. Если для уменьшения потерь в качестве диэлектрика в полосковой линии используется воздух, то такая линия называется *воздушной*. *Коаксиальная ЛП* –

двухсвязная линия, жесткая или гибкая, имеющая два проводника прямоугольного или круглого сечения – внутренний и внешний, разделенных диэлектриком или воздухом (в последнем случае линия называется воздушной) (рис. 13.1.4). *Диэлектрические ЛП* подразделяются по форме поперечного сечения (рис. 13.1.5). *Зеркальные диэлектрические ЛП* имеют металлический экран.

Волны, распространяющиеся посредством ЛП, называются *направляемыми*, в отличие от *свободных волн*. По своей структуре направляемые волны делятся на *поперечные*, или *T*-волны (от английского transverse – поперечный); *электрические*, или *E*-волны; *магнитные*, или *H*-волны; *гибридные*. В *T*-волнах векторы **E** и **H** не имеют продольных составляющих, т. е. $E_z = 0$, $H_z = 0$. В *E*-волнах вектор **H** не имеет продольной составляющей, т. е. $H_z = 0$, а вектор **E** имеет как поперечные, так и продольную составляющие. В *H*-волнах вектор **E** не имеет продольной составляющей, т. е. $E_z = 0$, а вектор **H** имеет как поперечные, так и продольную составляющие. В гибридных, или смешанных волнах и вектор **E**, и вектор **H** наряду с поперечными составляющими имеют и продольные составляющие.

Направляющая система

Рассмотрим произвольную бесконечно протяженную однородную направляющую систему, ориентированную вдоль оси *Z*. Будем считать, что она состоит из однородных изотропных материалов, не вносит потерь, в ней возбуждены только монохроматические волны одной частоты и сторонние источники отсутствуют. В этих условиях уравнения Максвелла для комплексных амплитуд векторов $\dot{\mathbf{E}}$, $\dot{\mathbf{H}}$ имеют вид:

$$\text{rot } \dot{\mathbf{H}} = i\omega\epsilon\dot{\mathbf{E}}, \quad \text{rot } \dot{\mathbf{E}} = -i\omega\mu\dot{\mathbf{H}}, \quad (13.2.1)$$

они сводятся к векторным однородным уравнениям Гельмгольца

$$\nabla^2 \dot{\mathbf{E}} + k^2 \dot{\mathbf{E}} = 0, \quad \nabla^2 \dot{\mathbf{H}} + k^2 \dot{\mathbf{H}} = 0 \quad (13.2.2)$$

Для произвольной *m*-ой моды волн, бегущих вдоль ЛП, зависимость комплексных амплитуд $\dot{\mathbf{E}}_m$, $\dot{\mathbf{H}}_m$ от *z* имеет вид:

$$\dot{\mathbf{E}}_m = \dot{\mathbf{E}}^0(\xi, \eta) \exp(\mp i\beta z), \quad \dot{\mathbf{H}}_m = \dot{\mathbf{H}}^0(\xi, \eta) \exp(\mp i\beta z), \quad (13.2.3)$$

где $\beta = \text{const}$ – коэффициент фазы, ξ, η – координаты в поперечном сечении ЛП, знак «минус» в показателе экспоненты соответствует, бегущей в положительном направлении оси *Z*, знак «плюс» – бегущей в обратном направлении; пока для определенности рассматриваем волну, бегущую в положительном направлении. С учетом этих выражений, уравнения Гельмгольца примут вид

$$\nabla_{\perp}^2 \dot{\mathbf{E}}_m + \gamma_{\perp}^2 \dot{\mathbf{E}}_m = 0, \quad \nabla_{\perp}^2 \dot{\mathbf{H}}_m + \gamma_{\perp}^2 \dot{\mathbf{H}}_m = 0, \quad (13.2.4)$$

где

$$\gamma_{\perp}^2 = k^2 - \beta^2 = \omega^2 \epsilon \mu - \beta^2 \quad (13.2.5)$$

– квадрат *поперечного волнового числа*,

$$\nabla_{\perp}^2 = \partial^2 / \partial x^2 + \partial^2 / \partial y^2 \quad (13.2.6)$$

– дифференциальный оператор Лапласа в поперечном сечении.

Если вектор $\dot{\mathbf{E}}_m$ и (или) вектор $\dot{\mathbf{H}}_m$ имеют продольные составляющие (т. е. если волна в ЛП электрическая, магнитная или гибридная), то нет необходимости решать уравнения Гельмгольца для всех шести составляющих этих векторов $\dot{E}_{mx}, \dot{E}_{my}, \dot{E}_{mz}, \dot{H}_{mx}, \dot{H}_{my}, \dot{H}_{mz}$, т. к. в этом случае существуют соотношения между поперечными и продольными составляющими этих векторов:

$$\begin{aligned}\gamma_{\perp}^2 \dot{E}_{mx} &= -i(\beta \frac{\partial \dot{E}_{mz}}{\partial x} + \omega \mu \frac{\partial \dot{H}_{mz}}{\partial y}), \\ \gamma_{\perp}^2 \dot{E}_{my} &= -i(\beta \frac{\partial \dot{E}_{mz}}{\partial y} - \omega \mu \frac{\partial \dot{H}_{mz}}{\partial x}), \\ \gamma_{\perp}^2 \dot{H}_{mx} &= i(\omega \epsilon \frac{\partial \dot{E}_{mz}}{\partial y} - \beta \frac{\partial \dot{H}_{mz}}{\partial x}), \\ \gamma_{\perp}^2 \dot{H}_{my} &= -i(\omega \epsilon \frac{\partial \dot{E}_{mz}}{\partial x} + \beta \frac{\partial \dot{H}_{mz}}{\partial y}),\end{aligned}\tag{13.2.7}$$

и остается только решить скалярные уравнения для продольных составляющих:

$$\nabla_{\perp}^2 \dot{E}_{mz} + \gamma_{\perp}^2 \dot{E}_{mz} = 0, \quad \nabla_{\perp}^2 \dot{H}_{mz} + \gamma_{\perp}^2 \dot{H}_{mz} = 0,\tag{13.2.8}$$

с учетом краевых условий, соответствующих рассматриваемой направляющей системе, а затем найти поперечные составляющие из приведенных равенств.

Электрические и магнитные волны

При решении краевой задачи (13.2.8) одновременно определяется поперечное волновое число γ_{\perp} , отличное от нуля, если волна электрическая, магнитная или гибридная. Постоянная γ_{\perp} зависит от формы и размеров поперечного сечения ЛП и от типа распространяющейся волны, но не зависит от частоты. Из (13.2.5) следует, что

$$\beta = \sqrt{k^2 - \gamma_{\perp}^2}\tag{13.3.1}$$

и

$$k = 2\pi f \sqrt{\epsilon \mu}.\tag{13.3.2}$$

В зависимости от частоты f подкоренное выражение в (13.3.1) может быть положительным (при $k > \gamma_{\perp}$), равным нулю (при $k = \gamma_{\perp}$) или отрицательным (при $k < \gamma_{\perp}$). В первом случае параметр β – действительное число, фазы составляющих векторов поля в фиксированный момент времени линейно зависят от координаты z , волна распространяется вдоль оси Z с постоянной фазовой скоростью $v_{\phi} = \omega/\beta$ и переносит энергию.

В третьем случае $k < \gamma_{\perp}$ параметр β чисто мнимый: $\beta = -i|\beta|$, множитель в (13.2.3):

$$\exp(-i\beta z) = \exp(-|\beta|z),\tag{13.3.3}$$

амплитуды составляющих векторов поля экспоненциально убывают вдоль оси Z , а фазы этих векторов не зависят от координат, т. е. поле имеет характер стоячей волны, экспоненциально

убывает вдоль ЛП (что в данном случае не означает потерь энергии), переноса энергии не происходит.

Во втором случае параметр $\beta=0$, частота, соответствующая условию $k = \gamma_{\perp}$, равная

$$f = f_{кр} = \gamma_{\perp} / (2\pi\sqrt{\epsilon\mu}), \quad (13.3.4)$$

и отвечающая ей длина волны

$$\lambda = \lambda_{кр} = 2\pi / \gamma_{\perp} = 1 / f_{кр} \sqrt{\epsilon\mu}, \quad (13.3.5)$$

называются *критическими*. Выясним смысл этого термина: выражая γ_{\perp} из (13.3.5) и подставляя в (13.3.1), получаем:

$$\beta = \sqrt{\omega^2 \epsilon\mu - (2\pi / \lambda_{кр})^2} = k \sqrt{1 - (\lambda / \lambda_{кр})^2}, \quad (13.3.6)$$

т. е. коэффициент фазы β действителен, а поле представляет собой распространяющуюся волну, переносящую энергию, только при условии

$$\lambda < \lambda_{кр}, \quad f > f_{кр}, \quad (13.3.7)$$

называемом *условием распространения волны в ЛП*; в противном случае энергия не переносится, а поле в форме стоячей волны быстро затухает вдоль Z . Отметим, что значение $f_{кр}$ (также как соответствующее ему значение $\lambda_{кр}$) зависит от формы и размеров поперечного сечения ЛП и типа волны (моды).

Поскольку длина волны и фазовая скорость в направляющей системе отличаются от соответствующих величин в вакууме, необходимо получить выражения для них. *Длиной направляемой волны* Λ называют расстояние между двумя поперечными сечениями ЛП, в которых в один момент времени фазы составляющих вектора $\dot{\mathbf{E}}_m$ отличаются на 2π , или, то же самое, расстояние, на которое поверхность равной фазы перемещается за период. С учетом экспоненциальной зависимости (13.2.3) всех составляющих векторов $\dot{\mathbf{E}}_m, \dot{\mathbf{H}}_m$ поля, это определение приводит к соотношениям:

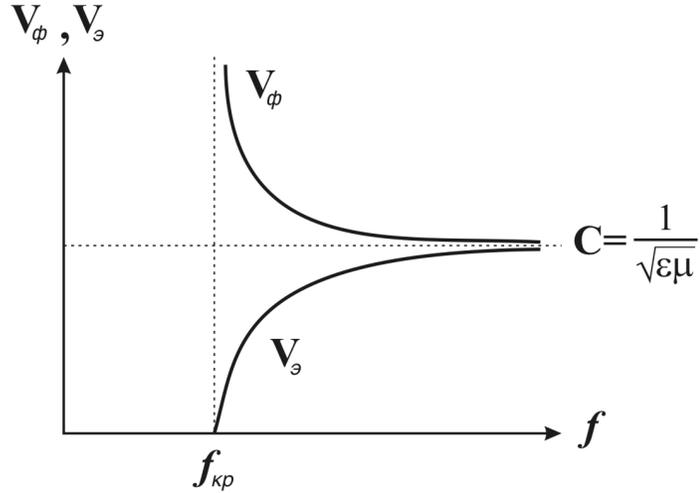
$$\Lambda = 2\pi / \beta = \lambda / \sqrt{1 - (\lambda / \lambda_{кр})^2} \quad (13.3.8)$$

$$v_{\phi} = \omega / \beta = c / \sqrt{1 - (\lambda / \lambda_{кр})^2} \quad (13.3.9)$$

Поскольку длина волны и фазовая скорость свободно распространяющейся волны в безграничной однородной среде без потерь с параметрами ϵ, μ равны

$$\lambda = c / f, \quad v_{\phi} = c = 1 / \sqrt{\epsilon\mu}, \quad (13.3.10)$$

то сравнение с (13.3.8), (13.3.9) показывает, что эти величины в ЛП больше в $1 / \sqrt{1 - (\lambda / \lambda_{кр})^2}$ раз (см. рис. 13.3.1), причем этот множитель неограниченно возрастает при приближении λ к $\lambda_{кр}$ слева. Само по себе явление зависимости фазовой скорости волны от частоты называется *дисперсией* и служит причиной различия фазовой и групповой скоростей, а также искажения формы некогерентного сигнала при распространении в среде.



Зависимость фазовой и групповой скоростей от частоты

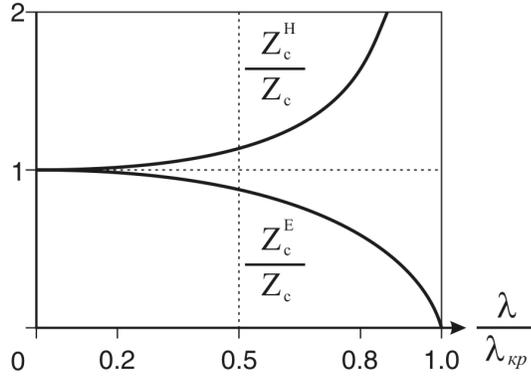
Характеристическим сопротивлением направляемой волны называют отношение поперечных (к направлению Z) составляющих вектора $\dot{\mathbf{E}}_m$ и вектора $\dot{\mathbf{H}}_m$. Можно показать [13.1], что для E -волны характеристическое сопротивление равно

$$Z_c^E = \frac{\beta}{\omega \epsilon} = Z_c \sqrt{1 - (\lambda / \lambda_{кр})^2}, \quad (13.3.11)$$

где $Z_c = \sqrt{\mu / \epsilon}$ – характеристическое сопротивление свободной волны, а для H -волны оно равно

$$Z_c^H = \omega \mu / \beta = Z_c / \sqrt{1 - (\lambda / \lambda_{кр})^2}. \quad (13.3.12)$$

Как видим, характеристические сопротивления E -волны и H -волны зависят от частоты (длины волны). В режиме переноса энергии ($\lambda < \lambda_{кр}$) Z_c^E всегда меньше Z_c и при приближении λ к $\lambda_{кр}$ слева Z_c^E стремится к нулю (рис. 13.3.2); в закритическом режиме ($\lambda > \lambda_{кр}$) Z_c^E чисто мнимое. Наоборот, Z_c^H в режиме переноса энергии всегда больше Z_c и при приближении λ к $\lambda_{кр}$ слева Z_c^H стремится к бесконечности; в закритическом режиме Z_c^H чисто мнимое.



Характеристические сопротивления для E -волны и H -волны

При рассмотрении формул (13.3.4), (13.3.5), (13.3.6), (13.3.8), (13.3.9) может показаться, что величины $\lambda_{кр}$, $f_{кр}$, β , ν_ϕ , Λ одинаковы для E -волн, H -волн и гибридных волн. Однако, это неверно: величины $\lambda_{кр}$, $f_{кр}$ зависят от поперечного волнового числа η_\perp , которое, в свою очередь зависит от формы и размеров поперечного сечения ЛП и от вида направляемой волны; остальные величины зависят от $\lambda_{кр}$. Таким образом, все перечисленные величины зависят от формы и размеров поперечного сечения ЛП и вида волны.

Поперечные волны

В сравнении с рассмотренными выше волнами, имеющими продольные составляющие, поперечные T -волны устроены намного проще. Полагая в (13.2.7) $\dot{E}_{mz} = 0$, $\dot{H}_{mz} = 0$, получаем для T -волны:

$$\gamma_\perp^2 \dot{E}_{mx} = 0, \quad \gamma_\perp^2 \dot{E}_{my} = 0, \quad \gamma_\perp^2 \dot{H}_{mx} = 0, \quad \gamma_\perp^2 \dot{H}_{my} = 0, \quad (13.4.1)$$

или в векторном виде, справедливом для произвольной системы координат в поперечном сечении ЛП:

$$\gamma_\perp^2 \dot{\mathbf{E}}_m = 0, \quad \gamma_\perp^2 \dot{\mathbf{H}}_m = 0 \quad (13.4.2)$$

Эта однородная система уравнений имеет нетривиальные решения $\dot{\mathbf{E}}_m \neq \mathbf{0}$, $\dot{\mathbf{H}}_m \neq \mathbf{0}$ только при $\gamma_\perp^2 = 0$. С учетом этого из (13.3.4), (13.3.5) следует, что для T -волны $f_{кр} = 0$, $\lambda_{кр} = \infty$. В свою очередь, это означает, что если в направляющей системе может распространяться T -волна, то эта возможность имеется на всех частотах (длинах волн), включая $f = 0$, т. е. на постоянном токе. Отсюда следует, что T -волна может распространяться только в направляющих системах не менее чем второго порядка связности; например, она не может распространяться в волноводах любого сечения (связность первого порядка) или в диэлектрических ЛП (связность нулевого порядка).

В уравнения (13.2.8) не входит частота, значит функции поперечных координат $\mathbf{E}^0(r, \phi)$, $\mathbf{H}^0(r, \phi)$ не зависят от частоты и могут быть найдены на основе решения рассматриваемой задачи при $f = 0$. Отсюда следует важный *квазистатический принцип*, позволяющий

упростить решение задачи: для определения векторов $\mathbf{E}^0(r, \varphi)$, $\mathbf{H}^0(r, \varphi)$ достаточно решить двумерные электростатическую и магнитостатическую задачи для той же линии.

Предостережем читателя от слишком буквального понимания квазистатического принципа: во-первых, статическая аналогия относится только к распределению поля в поперечном сечении линии, распределение вдоль оси Z в корне отличается от постоянного и носит волновой характер; во-вторых, аналогия позволяет найти комплексные амплитуды электрического и магнитного полей, полные поля включают еще фактор $\exp(-i\omega t)$, задающий колебания во времени, отсутствующие в статических полях.

Подставляя в (13.3.6), (13.3.8), (13.3.9) условие $\gamma_{\perp}^2 = 0$ ($\lambda_{\text{кр}} = \infty$), и в (13.2.7) – условия $\dot{E}_{mz} = 0$, $\dot{H}_{mz} = 0$, обнаруживаем, что для T -волны коэффициент фазы β_T , фазовая скорость $v_{\phi T}$, длина волны в линии Λ_T и характеристическое сопротивление Z_c^T равны:

$$\beta_T = k = \omega\sqrt{\epsilon\mu}, \quad (13.4.3)$$

$$v_{\phi T} = \omega/\beta_T = 1/\sqrt{\epsilon\mu} = c, \quad (13.4.4)$$

$$\Lambda_T = 2\pi/\beta_T = c/f = \lambda, \quad (13.4.5)$$

$$Z_c^T = \beta_T/(\omega\epsilon) = \sqrt{\mu/\epsilon} = Z_c, \quad (13.4.6)$$

т. е. в направляемой T -волне эти параметры те же самые, что в свободной волне, распространяющейся в безграничной однородной изотропной среде без потерь с теми же параметрами ϵ, μ , что в диэлектрике линии.

Выше предполагалось, что проводимость металлических частей ЛП идеальна. Если это условие не выполняется, электромагнитное поле проникает в металл, появляется продольная составляющая вектора $\dot{\mathbf{E}}_m$ и существование T -волны невозможно. Однако при высокой проводимости металла структура поля волны мало отличается от структуры T -волны и этим отличием пренебрегают. Это же замечание относится и к волнам с продольной составляющей вектора $\dot{\mathbf{E}}_m$ и (или) вектора $\dot{\mathbf{H}}_m$: отличие структуры реальных волн при реальной высокой проводимости металла от рассмотренной выше структуры настолько мало, что им обычно пренебрегают.

Электрическая прочность ЛП

Средняя по периоду мощность, переносимая волной по ЛП, равна интегралу по поперечному сечению S_{\perp} от действительной части продольной составляющей вектора Умова-Пойнтинга [13.1]:

$$P_{\text{cp}} = \frac{E_0^2}{2Z_c^T} \int_{S_{\perp}} |\dot{\mathbf{e}}_{\perp}^0|^2 dS, \quad (13.5.1)$$

где E_0 – максимальное значение напряженности электрического поля в ЛП;

$\dot{\mathbf{e}}_{\perp}^0 = \dot{\mathbf{E}}_{\perp}^0 / E_0$ – безразмерная векторная комплекснозначная функция, зависящая только от поперечных координат и определяющая структуру электрического поля в поперечном сечении ЛП;

$\dot{\mathbf{E}}_{\perp}^0$ – поперечная составляющая комплексного вектора $\dot{\mathbf{E}}^0$, введенного в (13.2.3);

Z_c^n – характеристическое сопротивление распространяющейся волны:
 $Z_c^n = Z_c, Z_c^E, Z_c^H$ для $T-, E-, H$ – волн соответственно.

Как следует из (13.5.1), увеличение средней по периоду передаваемой мощности P_{cp} связано с увеличением максимальной напряженности электрического поля E_0 . Это, в свою очередь, приближает E_0 к предельному значению $E_{пред}$, при котором происходит *электрический пробой* воздуха или диэлектрического заполнения ЛПП, характеризующийся большими токами и резким уменьшением активного сопротивления в разрядном промежутке, большим выделением тепла и резким повышением температуры в месте пробоя. Поступление мощности в нагрузку практически прекращается, а большая часть энергии падающей волны отражается от места пробоя и может вывести из строя генератор и вызвать другие нежелательные явления. Таким образом, электрический пробой в ЛПП недопустим.

С другой стороны, увеличение P_{cp} приводит к возрастанию мощности потерь в металлических элементах линии и заполняющем диэлектрике, т. е. их нагреву, что приближает материалы к разрушению и так называемому *тепловому пробую*. *Предельной мощностью* $P_{пред}$ называется меньшая из мощностей, вызывающих электрический или тепловой пробой. Для электрического пробоя это мощность, определяемая предельной напряженностью поля E_0 , равной при нормальных давлении воздуха и ионизации примерно 30 кВ/см; для теплового пробоя $P_{пред}$ определяют по температуре, при которой возникает тепловое разрушение материалов, образующих линию. Появление отраженных волн в реальной линии может привести к электрическому или тепловому пробую при мощности, существенно меньшей, чем $P_{пред}$. Поэтому *допустимую мощность* принимают в несколько раз меньшей, чем предельная, например,
 $P_{доп} = (0.2...0.3)P_{пред}$.

Затухание в ЛПП

Выше анализ свойств направляемых волн для простоты проводился в предположении, что ЛПП не вносит потерь. Зависимость векторов поля от координаты z задавалась множителем $\exp(-i\beta z)$, где в режиме переноса энергии β – действительное положительное и имеет смысл коэффициента фазы, а в закритическом режиме β – мнимое положительное, обуславливая быстрое затухание волны, не отвечающей условию распространения $\lambda < \lambda_{кр}$. В реальной ЛПП и в режиме переноса энергии волна ослабевает по мере движения вдоль линии, или *затухает*. Затухание волны связано, во-первых, с поглощением части передаваемой мощности в металлических частях линии, во-вторых, с поглощением в заполняющем диэлектрике, в-третьих, с излучением части мощности в окружающее пространство (в ЛПП открытого типа). Поэтому в реальной ЛПП в режиме переноса энергии зависимость векторов поля волны от z описывается множителем

$$\exp(-\gamma z), \quad (13.6.1)$$

где

$$\gamma = \alpha + i\beta \quad (13.6.2)$$

– *постоянная распространения*, комплексная величина, ее действительная часть α – коэффициент ослабления, имеющий размерность 1/м. Теперь зависимость амплитуды вектора поля от z описывается произведением двух экспонент:

$$\exp(-\gamma z) = \exp(-\alpha z) \exp(-i\beta z), \quad (13.6.3)$$

первая из которых задает затухание амплитуды с логарифмическим декрементом α , а вторая – отставание по фазе. Поскольку мощность пропорциональна квадрату амплитуды, затухание по мощности задается множителем $\exp(-2\alpha z)$.

Можно показать [13.1], что коэффициент ослабления, обусловленный потерями в металлических элементах ЛП, равен:

$$\alpha_m = Z_c R_s \left[\oint_{\Gamma} |\dot{\mathbf{H}}_{m\tau}^0|^2 dl + \int_{S_{\perp}} |\dot{\mathbf{E}}_{m\perp}^0|^2 dS \right], \quad (13.6.4)$$

где $R_s = \sqrt{\pi f \mu_0 / \sigma}$ – активная часть поверхностного сопротивления проводника;

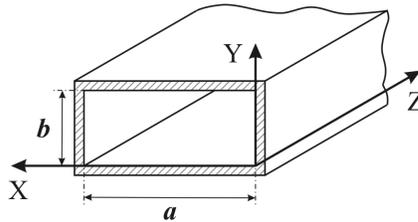
$\mu_0 = 4\pi \cdot 10^{-7}$ Гн/м, σ – удельная проводимость материала стенки волновода;

$\dot{\mathbf{H}}_{m\tau}^0$ – касательная составляющая вектора $\dot{\mathbf{H}}_m^0$;

Γ – контур поперечного сечения металлических элементов ЛП (односвязный в случае полого металлического волновода, двусвязный из двух концентрических окружностей – в случае коаксиальной линии и т. д.).

Прямоугольный волновод (E-волны, H-волны, основная волна, одноволновая передача, затухание)

Будем считать, что стенки волновода имеют бесконечную проводимость, а заполняющая его среда – идеальный (без потерь) диэлектрик с параметрами ϵ, μ . Поскольку эта направляющая система имеет порядок связности 1, в ней могут существовать E -волны и H -волны, но не T -волны (см. 13.4). На рис. 13.7.1 показаны применяемые далее система координат и размеры a, b широкой и узкой стенок ($a > b$). Источники, создающие поле в волноводе, расположены со стороны отрицательных значений координаты z , а созданные ими волны распространяются в положительном направлении оси Z .



Система координат для прямоугольного волновода

Т. к. поперечные составляющие векторов поля выражаются через продольные (см. 13.2), достаточно решить уравнения Гельмгольца относительно последних:

$$\nabla_{\perp}^2 \dot{E}_{mz} + \gamma_{\perp}^2 \dot{E}_{mz} = 0, \quad \nabla_{\perp}^2 \dot{H}_{mz} + \gamma_{\perp}^2 \dot{H}_{mz} = 0, \quad (13.2.8)$$

при соответствующих граничных условиях. Как видим, это две скалярные краевые задачи на собственные значения и на собственные функции (собственные волны) поперечного оператора Лапласа. Их решения стандартным *методом разделения переменных* имеют вид [13.1]:

а) для E -волн:

$$\begin{aligned}\dot{E}_{mv}(x, y, z) &= E_v^0(x, y) \exp(-i\beta z), \quad v = x, y, z, \\ \dot{H}_{mv}(x, y, z) &= H_v^0(x, y) \exp(-i\beta z), \quad v = x, y,\end{aligned}\quad (13.7.1)$$

где для m, n -ой моды ($m, n=1, 2, \dots$):

$$\begin{aligned}\dot{E}_z^0(x, y) &= E_{0z} \sin(m\pi x/a) \sin(n\pi y/b), \\ \dot{E}_x^0(x, y) &= -i(\beta/\gamma_\perp^2) E_{0z} (m\pi/a) \cos(m\pi x/a) \sin(n\pi y/b), \\ \dot{E}_y^0(x, y) &= -i(\beta/\gamma_\perp^2) E_{0z} (n\pi/b) \sin(m\pi x/a) \cos(n\pi y/b), \\ \dot{H}_x^0(x, y) &= i(\omega\varepsilon/\gamma_\perp^2) E_{0z} (n\pi/b) \sin(m\pi x/a) \cos(n\pi y/b), \\ \dot{H}_y^0(x, y) &= -i(\omega\varepsilon/\gamma_\perp^2) E_{0z} (m\pi/a) \cos(m\pi x/a) \sin(n\pi y/b), \\ \dot{H}_z^0(x, y) &= 0.\end{aligned}\quad (13.7.2)$$

б) для H -волн:

$$\begin{aligned}\dot{H}_{mv}(x, y, z) &= H_v^0(x, y) \exp(-i\beta z), \quad v = x, y, z, \\ \dot{E}_{mv}(x, y, z) &= E_v^0(x, y) \exp(-i\beta z), \quad v = x, y,\end{aligned}\quad (13.7.3)$$

где для m, n -ой моды ($m, n=0, 1, 2, \dots$, кроме одновременного $m = n = 0$):

$$\begin{aligned}H_z^0(x, y) &= H_{0z} \cos(m\pi x/a) \cos(n\pi y/b), \\ H_x^0(x, y) &= i(\beta/\gamma_\perp)(m\pi/a) H_{0z} \sin(m\pi x/a) \cos(n\pi y/b), \\ H_y^0(x, y) &= i(\beta/\gamma_\perp)(n\pi/b) H_{0z} \cos(m\pi x/a) \sin(n\pi y/b), \\ E_x^0(x, y) &= i(\omega\mu/\gamma_\perp)(n\pi/b) H_{0z} \cos(m\pi x/a) \sin(n\pi y/b), \\ E_y^0(x, y) &= -i(\omega\mu/\gamma_\perp)(m\pi/a) H_{0z} \sin(m\pi x/a) \cos(n\pi y/b), \\ E_z^0(x, y) &= 0.\end{aligned}\quad (13.7.4)$$

Как для E -волны, так и для H -волны постоянная γ_\perp для m, n -ой моды находится из соотношения:

$$\gamma_\perp = \sqrt{(m\pi/a)^2 + (n\pi/b)^2}, \quad (13.7.5)$$

а зная ее, из (13.3.5) находится критическая длина волны:

$$\lambda_{\text{кр}} = 2\pi/\gamma_\perp = \frac{2ab}{\sqrt{(mb)^2 + (na)^2}}, \quad (13.7.6)$$

длина волны в волноводе:

$$\Lambda = 2\pi/\beta = \lambda/\sqrt{1 - (\lambda/\lambda_{\text{кр}})^2}, \quad (13.3.8)$$

и фазовая скорость этой волны:

$$v_{\phi} = \omega / \beta = c / \sqrt{1 - (\lambda / \lambda)^2}. \quad (13.3.9)$$

Характеристическое сопротивление для E -волн равно

$$Z_c^E = \frac{\beta}{\omega \epsilon} = Z_c \sqrt{1 - (\lambda / \lambda_{\text{кр}})^2} \quad (13.3.11)$$

а для H -волн равно

$$Z_c^H = \omega \mu / \beta = Z_c / \sqrt{1 - (\lambda / \lambda_{\text{кр}})^2} \quad (13.3.12)$$

Заметим, что для E -волны определены все параметры, кроме E_{0z} , а для H -волны – все параметры, кроме H_{0z} . Для определения этих параметров нужны дополнительные данные, например, мощность источника.

Соотношения (13.7.1), (13.7.2), (13.7.3), (13.7.4) показывают, что в волноводе могут существовать различные моды E -волн и H -волн, структура поля и параметры которых зависят от двумерного номера m, n . Эти моды принято обозначать E_{mn} и H_{mn} , причем у E -волн m, n пробегает значения 1, 2, ..., а у H -волн один из индексов (но не оба) может равняться нулю.

Чтобы понять смысл номеров m и n , заметим, что структура поля в поперечном сечении волновода (т. е. при фиксированном значении координаты z) аналогична структуре стоячей волны, которую можно характеризовать «длинами волн» $\lambda_x = 2a / m$ и $\lambda_y = 2b / n$ в направлении осей X и Y соответственно. Таким образом, номер m есть число «полуволен» ($\lambda_x / 2$), укладываемых на поперечном размере a стенки, параллельной оси X , а номер n есть число «полуволен» ($\lambda_y / 2$), укладываемых на поперечном размере b стенки, параллельной оси Y . Равенство нулю одного из номеров означает, что поле рассматриваемой волны не зависит от соответствующей координаты (при $m = 0$ – от координаты x , а при $n = 0$ – от координаты y). Изменение всех составляющих комплексных амплитуд векторов $\dot{\mathbf{E}}$ и $\dot{\mathbf{H}}$ вдоль оси Z описывается множителем $\exp(-i\beta z)$. В волноводе без потерь распространение волны происходит только при $\lambda < \lambda_{\text{кр}}$, а критическая длина волны $\lambda_{\text{кр}}$ зависит, в силу (13.7.6), от размеров a и b , от номеров m и n моды. При фиксированных размерах a и b волновода с увеличением номеров m и (или) n значение $\lambda_{\text{кр}}$ уменьшается.

Наибольшую $\lambda_{\text{кр}}$ среди всех возможных волн при $a > b$ имеет волна H_{10} с $\lambda_{\text{кр}} = 2a$. При $a = b$ наибольшую $\lambda_{\text{кр}}$ имеют две волны H_{10} и H_{01} . Волну, имеющую наибольшую $\lambda_{\text{кр}}$, называют *основной волной* ЛП (или *волной низшего типа*). Таким образом, при $a > b$ основной волной прямоугольного волновода является волна H_{10} .

Волны, у которых $\lambda > \lambda_{\text{кр}}$, не распространяются: образуется стоячая волна, в которой $\beta = -i|\beta|$ и амплитуды составляющих векторов $\dot{\mathbf{E}}$ и $\dot{\mathbf{H}}$ экспоненциально убывают вдоль оси Z пропорционально фактору

$$\exp(-i\beta z) = \exp(-|\beta|z), \quad (13.7.7)$$

но это не связано с потерями на поглощение, они по-прежнему предполагаются отсутствующими.

Основная волна прямоугольного волновода

Изучим более подробно основную, при $a > b$, волну прямоугольного волновода H_{10} , имеющую наибольшую критическую длину волны $\lambda_{кр} = 2a$. Использование именно этой волны позволяет провести оптимизацию размеров волновода и структуры поля в нем. Полагая в вышеприведенных формулах $m = 1$ и $n = 0$, получаем для этой волны:

а) выражения для составляющих комплексных амплитуд векторов $\dot{\mathbf{E}}$ и $\dot{\mathbf{H}}$:

$$\begin{aligned}\dot{E}_{my} &= -i(\omega\mu a / \pi)H_{0z} \sin(\pi x / a) \exp(-i\beta_{10}z), \\ \dot{H}_{mx} &= i(\beta_{10}a / \pi)H_{0z} \sin(\pi x / a) \exp(-i\beta_{10}z), \\ \dot{H}_{mz} &= H_{0z} \cos(\pi x / a) \exp(-i\beta_{10}z), \\ \dot{E}_{mx} &= \dot{E}_{mz} = \dot{H}_{my} = 0,\end{aligned}\tag{13.7.8}$$

где $\beta_{10} = k\sqrt{1 - [\lambda/(2a)]^2}$ – коэффициент фазы,

$k = 2\pi/\lambda$ – волновое число,

$\lambda = c/f$ – длина свободной волны в среде с параметрами ϵ, μ ,

$c = 1/\sqrt{\epsilon\mu}$ – скорость света в той же среде;

б) фазовая скорость: $v_{\phi}^{H_{10}} = \frac{c}{\sqrt{1 - [\lambda/(2a)]^2}}$;

в) длина волны в линии: $\Lambda^{H_{10}} = \frac{\lambda}{\sqrt{1 - [\lambda/(2a)]^2}}$;

г) характеристическое сопротивление: $Z_c^{H_{10}} = \frac{Z_c}{\sqrt{1 - [\lambda/(2a)]^2}}$,

где $Z_c = \sqrt{\mu/\epsilon}$ – характеристическое сопротивление свободной волны в среде с параметрами ϵ, μ .

Одноволновая передача в волноводе

Практические соображения передачи сигналов в заданном диапазоне частот диктуют следующую двуединую цель:

а) выбрать наиболее «выгодную» волну;

б) обеспечить условия существования только этой волны (*одноволновой передачи*).

Что касается первой части задачи, то очевидно, что такой волной является основная волна, т. е. для прямоугольного волновода при $a > b$ – волна H_{10} . Действительно, эта волна имеет наибольшую критическую длину волны, равную $2a$, и на заданной частоте можно выбрать наименьшие поперечные размеры волновода, обеспечивающие перенос энергии этой волной, по сравнению с другими волнами. При этом волновод будет иметь наименьшие массу, габариты и стоимость.

Переходим ко второй части задачи. Необходимость условий одноволновой передачи обусловлена, главным образом, тем, что различные типы волн имеют различные фазовые, а следо-

вательно, и групповые скорости (*модовая дисперсия*), поэтому при передаче сигнала двумя или более волнами он приходит в место приема с искажениями, величина и характер которых зависят от вида модуляции, скорости передачи информации и т. п.

Если в качестве единственной волны выбрана основная волна, то одноволновую передачу просто обеспечить: для этого во всем рабочем диапазоне нужно обеспечить неравенства:

$$\lambda_{\text{кр}2} < \lambda < \lambda_{\text{кр}1}, \quad (13.7.9)$$

где $\lambda_{\text{кр}1} = 2a$ – критическая длина волны основной волны H_{10} прямоугольного волновода, λ – длина волны в рабочем диапазоне ($\lambda_{\text{мин}}, \lambda_{\text{макс}}$) волн, $\lambda_{\text{кр}2}$ – критическая длина волны *первого высшего типа*, т. е. волны с критической длиной, большей, чем у всех других волн, не считая основной. Поскольку критическая длина волны H_{20} выражается через размер a широкой стенки:

$$\lambda_{\text{кр}}^{H_{20}} = a, \quad (13.7.10)$$

а критическая длина волны H_{01} – через размер b узкой стенки:

$$\lambda_{\text{кр}}^{H_{01}} = 2b, \quad (13.7.11)$$

то, в зависимости от соотношения a и b , первым высшим типом может оказаться или та, или другая, и если не конкретизировать это соотношение, для обеспечения одноволновой передачи нужно выполнение двух условий:

$$\lambda/2 < a < \lambda; \quad b < \lambda/2. \quad (13.7.12)$$

Отсюда следует, что одноволновая передача обеспечена во всем рабочем диапазоне

$$\lambda_{\text{мин}} < \lambda < \lambda_{\text{макс}}, \quad (13.7.13)$$

если размеры волновода выбираются исходя из условий:

$$\lambda_{\text{макс}}/2 < a < \lambda_{\text{мин}} \quad \text{и} \quad b < \lambda_{\text{мин}}/2 \quad (13.7.14)$$

На практике, однако, диапазон для выбора a сужают [13.4]:

$$0.6\lambda_{\text{макс}} < a < 0.9\lambda_{\text{мин}}, \quad (13.7.15)$$

чтобы избежать чрезмерного приближения λ к $\lambda_{\text{кр}}$, при котором затухание в волноводе резко увеличивается.

Для заданных размеров волновода полосу частот, в пределах которой сохраняется одноволновая передача, обычно характеризуют *коэффициентом широкополосности*:

$$\xi = \lambda_{\text{кр}1} / \lambda_{\text{кр}2}. \quad (13.7.16)$$

Если принять размеры волновода, исходя из (13.7.14), то теоретически можно добиться отношения крайних частот рабочего диапазона $\xi_{\text{раб}} = f_{\text{макс}} / f_{\text{мин}} = \xi = 2$. Однако, если воспользоваться

суженным условием, например, (13.7.15), реальное значение *широкополосности* составит $\xi_{\text{раб}} = 1.6 - 1.9$.

Допустимая передаваемая мощность

Используя формулы (13.5.1) и (13.7.8), находим выражение средней по периоду передаваемой мощности для основной волны H_{10} :

$$P_{\text{ср}}^{H_{10}} = \frac{E_0^2 ab}{4Z_c} \sqrt{1 - [\lambda/(2a)]^2}, \quad (13.7.17)$$

где $E_0 = (\omega \mu a / \pi) H_{0z}$ – амплитуда электрической напряженности волны H_{10} . Чтобы выразить предельную мощность через длину волны, примем $b \approx a/2$, $a \approx 0.75\lambda$, $E_{0\text{пред}} = 30$ кВ/см, тогда

$$P_{\text{пред}}^{H_{10}} \approx 125\lambda^2, \text{ кВт}, \quad (13.7.18)$$

где λ выражена в сантиметрах. Как видим, предельная (и, соответственно, допустимая) мощность быстро падает с уменьшением длины волны. Но даже на самых малых длинах волн, используемых в ВАЦ, допустимая передаваемая мощность волновода на несколько порядков больше той, что обычно используется в ВАЦ. Так, при $\lambda = 0.2$ см ($f = 150$ ГГц) предельная мощность основной волны $P_{\text{пред}}^{H_{10}} \approx 5$ кВт, а допустимая мощность $P_{\text{доп}}^{H_{10}} = 0.2 P_{\text{пред}}^{H_{10}} \approx 1$ кВт, в то время как в ВАЦ, в зависимости от вида тестируемого устройства и режима измерения, используется мощность примерно от 10 мВт до нескольких Вт.

Затухание в прямоугольном волноводе

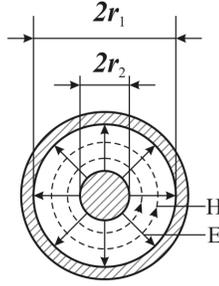
Хотя выше решение внутренней электродинамической задачи для прямоугольного волновода получено для простоты в предположении отсутствия потерь, в реальном волноводе потери, и следовательно, ослабление при передаче энергии, имеются, в первую очередь, за счет поглощения на стенках. Наименьшие потери имеют место при передаче энергии основной волной H_{10} . Из (13.7.8), (13.5.1) можно найти для этой волны [13.1]:

$$\alpha_m^{H_{10}} = \frac{2R_s}{bZ_c} \left[1 + \frac{2b}{a} \left(\frac{\lambda}{2a} \right)^2 \right] / \sqrt{1 - \left(\frac{\lambda}{2a} \right)^2}. \quad (13.7.19)$$

Аналогичным образом можно вывести формулы для коэффициента ослабления за счет потерь в металле для других типов волн.

Коаксиальная ЛП (коаксиал, основная волна, высшие волны и одноволновый режим, допустимая мощность, затухание, стандартное значение волнового сопротивления)

На практике используется почти исключительно *круглая коаксиальная линия*, или *коаксиал*, поперечное сечение которого показано на рис. 13.8.1. Это закрытая продольно-регулярная металло-диэлектрическая двухсвязная ЛП. Пространство между внешним и внутренним проводниками, в котором распространяются волны, может быть заполнено воздухом или другим диэлектриком с параметрами ϵ и μ . При анализе будем считать, что проводники имеют бесконечную проводимость, а диэлектрик идеальный. Согласно **13.4**, при этих условиях и вследствие двухсвязности, в коаксиале, кроме E -волн и H -волн, может распространяться и T -волна. Так как $\lambda_{\text{кр}}^T = \infty$, то в любой линии, в которой может распространяться T -волна, последняя является основной.



Поля в коаксиале

Конструктивно [13.7, 13.8] внутренний проводник коаксиала может быть сплошным, сплетенным из отдельных проволочек или трубчатым; выполнен из меди или биметаллической проволоки. Внешний проводник может быть в виде полый трубы (жесткий коаксиал), либо в виде оплетки из медной проволоки или ленты (гибкий коаксиал). В воздушных коаксиалах конструкция поддерживается с помощью диэлектрических шайб.

13.8.2. Структура T -волны, основной волны коаксиала

Введем цилиндрическую систему координат (r, φ, z) , совместив ось Z с осью внутреннего проводника (рис. 13.8.1). Сначала рассмотрим структуру T -волны. Ее векторы $\dot{\mathbf{E}}$ и $\dot{\mathbf{H}}$ представим в виде

$$\dot{\mathbf{E}}_m(r, \varphi, z) = \mathbf{E}^0(r, \varphi) \exp(-ikz), \quad \dot{\mathbf{H}}_m(r, \varphi, z) = \mathbf{H}^0(r, \varphi) \exp(-ikz) \quad (13.8.1)$$

где векторы $\mathbf{E}^0(r, \varphi)$, $\mathbf{H}^0(r, \varphi)$ не имеют продольных составляющих. Применяя справедливый для T -волны квазистатический принцип (см. 13.4), можно найти комплексные амплитуды электрического и магнитного полей в области $r_2 < r < r_1$ [13.1]:

$$\dot{\mathbf{E}}_m(r, \varphi, z) = \mathbf{r}_0 \frac{I^0 Z_c}{2\pi r} \exp(-ikz), \quad \dot{\mathbf{H}}_m(r, \varphi, z) = \varphi_0 \frac{I^0}{2\pi r} \exp(-ikz), \quad (13.8.2)$$

где I^0 – ток, текущий по внутреннему проводнику;

\mathbf{r}_0 – единичный вектор вдоль радиуса;

$Z_c = \sqrt{\mu/\varepsilon}$ – характеристическое сопротивление свободной волны в непоглощающей среде, имеющей те же ε, μ , что и диэлектрик данного коаксиала;

$k = \beta^T = \omega\sqrt{\varepsilon\mu}$ – коэффициент фазы (волновое число) в той же среде;

φ_0 – единичный вектор вдоль касательной к концентрической окружности.

В рамках данной физико-математической модели определить постоянную I^0 нельзя; для ее нахождения требуются дополнительные данные об источнике.

Как видно из (13.8.2), структура поля T -волны в сечении коаксиала соответствует рис. 13.8.1: векторы $\dot{\mathbf{E}}_m$ электрического поля направлены вдоль радиуса, векторы $\dot{\mathbf{H}}_m$ магнитного

поля направлены по касательным к концентрическим окружностям. Заметим, что модули электрической и магнитной напряженностей поля в точке сечения линии обратно пропорциональны расстоянию r до центра, откуда следует, что передаваемая мощность на единицу площади сечения (*плотность потока мощности*) убывает к периферии сечения.

Как показано в **13.4**, в T -волне любой ЛП коэффициент фазы β_T , фазовая скорость $v_{фТ}$, длина волны в линии Λ_T и характеристическое сопротивление Z_c^T те же самые, что в свободной волне, распространяющейся в безграничной однородной изотропной среде без потерь с теми же параметрами ϵ, μ , что в диэлектрике линии. Поэтому и для коаксиала остаются справедливыми формулы (13.4.3), (13.4.4), (13.4.5), (13.4.6). Что касается волнового сопротивления линии в режиме T -волны, то поскольку в этом режиме поле $\mathbf{E}^0, \mathbf{H}^0$ в поперечном сечении имеет *потенциальный характер*, физически существуют ток и напряжение в коаксиале. Комплексная амплитуда тока в центральном проводнике:

$$\dot{I}_m = I^0 \exp(-ikz), \quad (13.8.3)$$

а комплексная амплитуда разности потенциалов между центральным и внешним проводниками:

$$\dot{U}_m = \int_{r_2}^{r_1} \dot{E}_{mr} dr = \frac{I^0 Z_c}{2\pi} \ln \frac{r_1}{r_2} \cdot \exp(-ikz). \quad (13.8.4)$$

Теперь *волновое сопротивление коаксиала*:

$$Z_B = \frac{\dot{U}_m}{\dot{I}_m} = \frac{Z_c}{2\pi} \ln \frac{r_1}{r_2} = 60 \sqrt{\frac{\mu_r}{\epsilon_r}} \ln \frac{r_1}{r_2}, \quad (13.8.5)$$

где μ_r, ϵ_r – относительные магнитная и диэлектрическая проницаемости.

Высшие волны коаксиала и одноволновый режим

Перейдем к E - и H -волнам. В этих случаях квазистатический принцип не действует и необходимо решать уравнения Гельмгольца (13.2.2) в соответствующих граничных условиях. Зависимость от z дается равенствами (13.2.3), поперечные составляющие векторов $\dot{\mathbf{E}}_m$ и $\dot{\mathbf{H}}_m$ выражаются через продольные составляющие $\dot{E}_{mz}, \dot{H}_{mz}$, а двумерная краевая задача в сечении линии (13.2.4) в полярных координатах сечения принимает вид:

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial E_{mz}^0}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 E_{mz}^0}{\partial \varphi^2} + \gamma_{\perp}^2 E_{mz}^0 = 0, \quad (13.8.6)$$

и аналогичное уравнение для H_{mz}^0 . Эта задача решается методом разделения переменных при надлежащих граничных условиях на внешней поверхности внутреннего проводника и внутренней поверхности внешнего проводника. Решение выражается через функции Бесселя $J_m(\gamma_{\perp} r)$ и Неймана $N_m(\gamma_{\perp} r)$ m -го порядка и их первые производные.

Из граничных условий для E -волн: $E_{mz}^0(r_1, \varphi) = 0, E_{mz}^0(r_2, \varphi) = 0$ можно вывести уравнение относительно γ_{\perp}^E [13.1, 13.3]:

$$\frac{J_m(\gamma_{\perp} r_2)}{J_m(\gamma_{\perp} r_1)} = \frac{N_m(\gamma_{\perp} r_2)}{N_m(\gamma_{\perp} r_1)}, \quad (13.8.7)$$

а из граничных условий для H -волн – уравнение для γ_{\perp}^H :

$$\frac{J_m'(\gamma_{\perp} r_2)}{J_m'(\gamma_{\perp} r_1)} = \frac{N_m'(\gamma_{\perp} r_2)}{N_m'(\gamma_{\perp} r_1)}. \quad (13.8.8)$$

Эти уравнения трансцендентны и решаются численными методами. Анализ уравнений (13.8.7), (13.8.8) показывает, что при любом соотношении r_1 и r_2 первым высшим типом в коаксиале является волна H_{11} . Критическая длина волны для этого типа зависит от отношения r_2/r_1 . Асимптотический анализ уравнения (13.8.8) показывает, что при $r_2/r_1 \rightarrow 0$ $\lambda_{\text{кр}}^{H_{11}} \rightarrow 3.41r_1$, а при $r_2/r_1 \rightarrow 1$ $\lambda_{\text{кр}}^{H_{11}} \rightarrow \pi(r_1 + r_2)$. Обычно в качестве независимого от отношения r_2/r_1 значения критической длины волны приближенно принимают последний результат:

$$\lambda_{\text{кр}}^{H_{11}} = \pi(r_1 + r_2), \quad (13.8.9)$$

при этом ошибка этого равенства во всем диапазоне отношений r_2/r_1 не превышает 10%.

Таким образом, поскольку $\lambda_{\text{кр}}^T = \infty$, одноволновый режим в коаксиале будет для волн с длиной

$$\lambda > \lambda_{\text{кр}}^{H_{11}} = \pi(r_1 + r_2). \quad (13.8.10)$$

Допустимая передаваемая мощность

Средняя за период мощность, переносимая T -волной по коаксиалу найдется как Подставляя сюда первую формулу из (13.8.2), получаем:

$$P_{\text{ср}}^T = \frac{\pi E_0^2 r_2}{Z_c} \ln \frac{r_1}{r_2}, \quad (13.8.12)$$

где

$$E_0 = \frac{I^0 Z_c}{2\pi r_2} \quad (13.8.13)$$

– амплитуда напряженности электрического поля на поверхности внутреннего проводника, т. е. наибольшее значение этой амплитуды в сечении линии;

смысл I^0 и Z_c дан в пояснении к (13.8.2).

Поскольку электрический пробой зависит от E_0 , полезно найти такое соотношение между r_1 и r_2 , при котором, при заданной $P_{\text{ср}}^T$, величина E_0 будет минимальной. Это легко сделать с помощью (13.8.12), оптимальное с точки зрения пробоя соотношение между r_1 и r_2 получается таким: $\ln(r_1/r_2) = 0.5$, т. е.

$$r_1 = \sqrt{e} \cdot r_2 = 1.648 \cdot r_2 \quad (13.8.14)$$

При таком соотношении радиусов получается наибольшее значение предельной мощности, а волновое сопротивление коаксиала, согласно (13.8.5), получается

$$Z_B = 30\sqrt{\mu_r / \epsilon_r}, \text{ Ом.} \quad (13.8.15)$$

Оценку значения предельной мощности волны для коаксиала следует провести для двух случаев [13.1].

1) Если роль диэлектрика выполняет воздух (воздушный коаксиал), тогда пробой возникает при $E_0 = 30$ кВ/см, для воздуха $Z_c = 120\pi$, соотношение радиусов примем оптимальным, т. е. $\ln(r_1 / r_2) = 0.5$, в результате из (13.8.12) получаем:

$$P_{\text{пред}} = 3.75 \cdot 10^3 \cdot r_2^2, \text{ кВт,} \quad (13.8.16)$$

где r_1 – в сантиметрах.

2) Если пространство между центральным и внешним проводниками заполнено полностью или частично диэлектриком, то, во-первых, возникает новый фактор – возможность теплового пробоя диэлектрика, во-вторых, в технологически неизбежных воздушных зазорах, порядка (10 – 100) мкм, между диэлектриком и центральным проводником напряженность электрического поля примерно в ϵ_r выше, чем в самом диэлектрике. Поэтому во избежание пробоя воздушных зазоров следует принять

$$P_{\text{пред}} = 3.75 \cdot 10^3 \cdot r_1^2 / \epsilon_r^2. \quad (13.8.17)$$

Как аргументировано в **13.5**, допустимую мощность следует принимать в несколько раз меньшей, чем предельная, например, $P_{\text{доп}} = (0.2 \dots 0.3) P_{\text{пред}}$.

Затухание в коаксиале

Потери в коаксиале суть сумма потерь в диэлектрике и в металлических проводниках, так что коэффициент ослабления $\alpha = \alpha_d + \alpha_m$. Существуют методы расчета и измерения тех и других потерь. На практике можно пользоваться следующими оценками [13.4]:

$$\alpha_d = 27.3 \sqrt{\epsilon_r} \text{tg} \delta / \lambda, \quad (13.8.18)$$

$$\alpha_m = \frac{16.5 R_s \sqrt{\epsilon_r} f (1 + r_1 / r_2)}{r_1 \ln(r_1 / r_2)}, \quad (13.8.19)$$

где $\text{tg} \delta$ – тангенс угла диэлектрических потерь;

f – частота колебаний в гигагерцах;

R_s – активная часть поверхностного сопротивления проводника.

На волнах короче 10 см суммарный коэффициент ослабления коаксиала столь значителен, что применяют лишь короткие отрезки коаксиала.

Стандартное значение волнового сопротивления

Возвращаясь к выбору отношения $\psi = r_1 / r_2$ и волнового сопротивления Z_b коаксиала, замечаем, что этот выбор неоднозначен и зависит от критерия оптимальности.

1) Как показано в **13.8.4**, оптимальное с точки зрения пробы $\psi = 1.648$, при этом $Z_b = 30\sqrt{\mu_r / \epsilon_r}$, Ом (см. (13.8.14), (13.8.15)).

2) Оптимальное ψ , при котором ослабление за счет металлических проводников минимально, есть $\psi = 3.6$, при этом $Z_b = 77\sqrt{\mu_r / \epsilon_r}$.

3) Так как $r_2 < r_1$, то наибольшие потери энергии происходят в центральном проводнике, поэтому увеличение его радиуса приводит к уменьшению тока проводимости и уменьшению потерь. Если увеличение r_2 сопроводить таким же увеличением r_1 (пропорциональное увеличение радиусов), то вследствие условия одноволновой передачи (13.8.10) снизится максимальная частота доступного рабочего диапазона частот; если же r_2 увеличивать при неизменном r_1 , то снизится волновое сопротивление Z_b , что при данной передаваемой мощности P_{cp} приведет к увеличению тока в линии и соответствующему увеличению потерь. Оптимальное ψ в этом смысле указано в п. 2.

4) Иногда ставят задачу задать такое ψ , при котором минимальна разность потенциалов $\Delta U = |\dot{U}_m|$ между внутренним и внешним проводниками коаксиала. Оптимальное ψ в этом смысле $\psi = e = 2.716$, при этом $Z_b = 60\sqrt{\mu_r / \epsilon_r}$.

Принятый по рекомендации Международной электротехнической комиссии международный стандарт задает для коаксиала, предназначенного для передачи «значительной» мощности, $\psi = 2.3$, $Z_b = 50$ Ом и сетку пар радиусов с указанным ψ . В случае «малой» передаваемой мощности применяются коаксиалы с $Z_b = 75$ Ом, а также, иногда, с $Z_b = 100$ Ом и 150 Ом.

Полосковые ЛП (классификация полосковых ЛП, частотные ограничения, эффективная диэлектрическая проницаемость, волновое сопротивление и погонная емкость, длина волны в линии, коэффициент фазы, затухание в линии)

Модель эквивалентной ЛП

Двухпроводная симметричная модель

Радиосистемы, работающие в диапазоне СВЧ, обычно можно представить в виде *устройств СВЧ*, соединенных отрезками *линий передачи* (ЛП) и образующих *тракты СВЧ*. Компоненты трактов СВЧ не являются локальными радиотехническими элементами, такими как емкость, индуктивность, резистор и их комбинации, например, резонансный контур и т. п., по той причине, что в диапазоне СВЧ характерные размеры L устройств и ЛП не отвечают условию $L \ll \lambda$. Строгое рассмотрение распространения электромагнитных сигналов в трактах СВЧ на основе краевых задач электродинамики очень сложно и выполнимо только для простейших устройств и ЛП, например, отдельных отрезков ЛП. С другой стороны, если даже такое описание распространения сигналов может быть получено, оно часто оказывается излишне подробным [14.1]. Более простой альтернативой является переход к *эквивалентным схемам* устройств.

Для устройств СВЧ вообще (включая отрезки ЛП) этот переход происходит по двум направлениям:

1) Прделав предварительно, насколько это возможно, *декомпозицию* сложной цепи на ряд более простых *элементов*, которые можно анализировать независимо, пытаются для каждого элемента подобрать такую схему из сосредоточенных постоянных L , C , R , трансформаторов

и отрезков линий передачи, которая в смысле определенных количественных критериев дает достаточно близкие к действительности результаты своей работы. В необходимых случаях должна иметься возможность связать эквивалентные параметры L, C, R и интегральные параметры ЛП с волновыми параметрами строгого электродинамического описания

2) После такой же декомпозиции каждый линейный элемент представляют в виде $2n$ -полюсника, внешние параметры которого можно описать с помощью матриц того или иного вида (см. главу 10). В необходимых случаях должна иметься возможность связать элементы описывающей матрицы (например, S-параметры) с волновыми параметрами.

Возвращаясь теперь к более узкой задаче перехода от строгого электродинамического описания (см. главу 13) ЛП к описанию на уровне более простой модели, видим, что дело сводится к построению модели *эквивалентной ЛП* [14.2–14.4] (мы примем именно это наименование), или *эквивалентной цепи с распределенными параметрами* [14.5], или *длинной линии* [14.6], или *линии с распределенными параметрами*, или *эквивалентной схеме ЛП* и т. п. (различные названия одного и того же понятия).

Физически модель эквивалентной ЛП можно представлять как двухпроводную линию, в которой *первичными параметрами* являются погонные сопротивление, проводимость, емкость и индуктивность, а *вторичными* – напряжение и ток. Иногда вторичные параметры называют *интегральными*, поскольку они выражаются интегралами от первичных параметров (точнее, они суть решения дифференциальных уравнений, в которых первичные параметры задают коэффициенты). Электромагнитные процессы в эквивалентной ЛП описываются скалярными величинами (в общем случае – комплексными) напряжения и тока как функций лишь продольной координаты. Эти функции строятся на основе векторных напряженностей электрического и магнитного поля как функций пространственных и временной координат, получаемых для каждого типа ЛП из решения соответствующей электродинамической задачи.

Прежде чем задать эквивалентную ЛП как модель конкретной физической ЛП – волноводной, коаксиальной, полосковой – следует убедиться в возможной адекватности такого перехода, в частности, в возможности связать первичные параметры эквивалентной ЛП с волновыми параметрами строгого электродинамического описания, а также определить область использования этой упрощенной модели в пространстве характеристик ЛП как устройства СВЧ. Отметим очевидные ограничения области действия модели эквивалентной ЛП [14.2]:

1) Модель можно использовать лишь в условиях (например, в частотном диапазоне) одноволнового режима. К счастью, от этого недостатка можно избавиться, отнеся отдельную эквивалентную линию каждой моде поля при многоволновой передаче.

2) Модель не пригодна для определения предельной пропускаемой мощности $P_{\text{пред}}$.

3) Модель не пригодна для анализа степени взаимной связи между открытыми ЛП.

Вернемся к вопросу об адекватности модели. Если основной волной физической ЛП является T -волна (как в двухпроводной и коаксиальной линиях) и обеспечено условие одноволновой передачи, вопрос становится простым: поле в поперечнике линии потенциально, поэтому можно по распределению векторов \mathbf{E} электрического поля и \mathbf{H} магнитного поля T -волны в поперечнике однозначно, путем контурного интегрирования, определить напряжение \dot{U} и ток \dot{I} как функции продольной координаты z . В линии без потерь эти зависимости следуют из (13.2.3) и имеют вид:

$$\dot{U}(z) = U_0 \exp(-i\beta z), \quad \dot{I}(z) = I_0 \exp(-i\beta z) \quad (14.1.1)$$

Зная функции (14.1.1), можно вычислить волновое сопротивление $Z_{\text{в}} = \dot{U}(z) / \dot{I}(z)$ и среднюю за период переносимую мощность

$$P_{cp} = |\dot{U}(z)|^2 / (2Z_B). \quad (14.1.2)$$

Таким образом, в этом случае напряжение и ток в линии физически реальны и переход к модели эквивалентной линии адекватен.

Если основной волной физической ЛП является квази T -волна (как в полосковой линии) и обеспечено условие одноволновой передачи, то изложенное выше остается в силе, но уже приближенно (вопрос об ошибке модели в данном случае здесь не рассматривается).

Если в ЛП T -волна не распространяется (как в металлическом волноводе), то в этой ЛП нет физически реальных напряжения и тока, переносящих мощность [14.4]. В такой линии распространяются E -, H - или смешанные волны, и можно попытаться определить напряжение $\dot{U}_m(z)$ и ток $\dot{I}_m(z)$ эквивалентной линии формально через контурные интегралы от поперечных составляющих поля $\dot{E}_{m\perp}, \dot{H}_{m\perp}$. Однако, в этом случае поле, описываемое этими составляющими, не является потенциальным и функции $\dot{U}_m(z)$ и $\dot{I}_m(z)$ определяются неоднозначно: они зависят от выбора контуров интегрирования. Эту трудность можно обойти, заранее оговорив форму контуров.

Рассмотрим этот прием на примере [14.2]. Пусть имеется прямоугольный волновод с размерами стенок a (широкая) и b (узкая). Как показано в главе 13, основная волна в этом случае – волна H_{10} . Используя выражение (13.7.4) для составляющей \dot{E}_{my} волны H_{10} , находим комплексную амплитуду напряжения между точками, лежащими на средних линиях широких стенок при $x = a/2$:

$$\dot{U}_m(z) = \int_0^b \dot{E}_{my} dy \Big|_{x=a/2} = U_0 \exp(-i\beta z) \quad (14.1.3)$$

где

$$U_0 = -i(\omega\mu ab / \pi) H_{0z}, \quad (14.1.4)$$

и принимаем ее за величину, пропорциональную комплексной амплитуде $\dot{U}(z)$ напряжения в эквивалентной линии:

$$\dot{U}(z) = K_1 \dot{U}_m(z) = K_1 U_0 \exp(-i\beta z) = -K_1 i(\omega\mu ab / \pi) H_{0z} \exp(-i\beta z). \quad (14.1.5)$$

Как отмечено в **13.7**, $\dot{U}_m(z)$ определено с точностью до H_{0z} , такова же степень неопределенности формулы (14.1.5): с точностью до $K_1 H_{0z}$.

Используя выражение (13.7.4) для составляющей \dot{H}_{mx} волны H_{10} , находим комплексную амплитуду тока, текущего по нижней ($y = 0$) широкой стенке волновода:

$$\dot{I}_m(z) = \int_0^a \dot{H}_{mx} dx \Big|_{y=0} = I_0 \exp(-i\beta z), \quad (14.1.6)$$

где

$$I_0 = -i(2\beta a^2 / \pi) H_{0z}, \quad (14.1.7)$$

$\beta \equiv \beta_{H_{0z}}$, и принимаем ее за величину, пропорциональную комплексной амплитуде $\dot{I}(z)$ тока в эквивалентной линии:

$$\dot{I}(z) = K_2 \dot{I}_m(z) = K_2 I_0 \exp(-i\beta z) = -K_2 i (2\beta a^2 / \pi) H_{0z} \exp(-i\beta z). \quad (14.1.8)$$

Как отмечено в **13.7**, $\dot{I}_m(z)$ определено с точностью до H_{0z} , такова же степень неопределенности формулы (14.1.8): с точностью до $K_2 H_{0z}$.

Учитывая (14.1.3) и (14.1.6), можно определить волновое сопротивление для волны H_{10} :

$$Z_b^{H_{10}} = \dot{U}_m(z) / \dot{I}_m(z) = \pi b Z_c^{H_{10}} / (2a) = \pi b Z_c / (2a \sqrt{1 - (\lambda / \lambda_{кр})^2}), \quad (14.1.9)$$

и принять его за волновое сопротивление для эквивалентной линии. Поскольку (14.1.5) определено с точностью до коэффициента $K_1 H_{0z}$, а (14.1.8) – с точностью до коэффициента $K_2 H_{0z}$, то (14.1.9) определено с точностью до коэффициента

$$A = \frac{K_1 H_{0z}}{K_2 H_{0z}} = \frac{K_1}{K_2}. \quad (14.1.10)$$

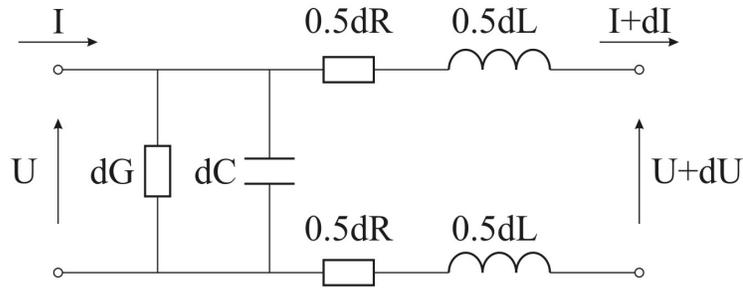
Это значит, что, изменив форму контура интегрирования или даже саму методику определения напряжения $\dot{U}_m(z)$ и тока $\dot{I}_m(z)$ в ЛП, мы получим другое выражение для Z_b , но во всех случаях формула для Z_b будет иметь вид:

$$Z_b^{H_{10}} = B \frac{b}{a} Z_c, \quad (14.1.11)$$

где B – коэффициент, зависящий только от способа вычисления $\dot{U}_m(z)$ и $\dot{I}_m(z)$. Неопределенность в выборе этого коэффициента существенного значения не имеет [14.2], т. к. при проектировании цепей СВЧ важно знать отношение волновых сопротивлений соединяемых отрезков ЛП, а не конкретные значения каждого из них.

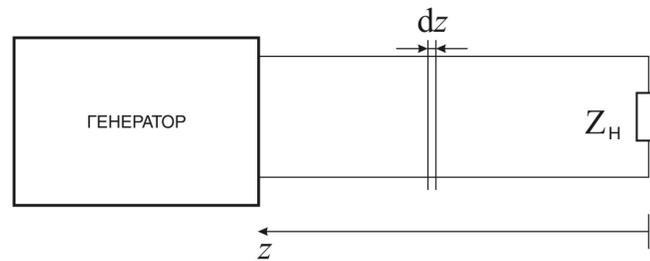
Таким образом, любую из рассмотренных в главе 13 линий передачи (волноводную, коаксиальную, полосковую) в интересах анализа можно заменить эквивалентной ЛП, в которой распространяются соответствующие волны напряжения и тока. И хотя обоснование правомерности этой замены для разных типов ЛП имеет разную степень убедительности, модель эквивалентной ЛП чрезвычайно полезна при инженерном проектировании трактов СВЧ, оптимальна по сложности и информативности описания и многократно проверена экспериментально.

Для анализа процессов, происходящих в анализаторе цепей, концепция эквивалентных ЛП имеет особое значение: анализатор нацелен на измерение характеристик тестируемого устройства на уровне подробности S -параметров (ВАЦ измеряет комплексные значения S -параметров, САЦ – модули S -параметров), этому уровню подробности точно соответствует описание ЛП в модели эквивалентных линий.



Двухпроводная модель эквивалентной ЛП

Первичные параметры



Координата в линии передачи

В качестве модели эквивалентной ЛП рассмотрим двухпроводную длинную линию (рис. 12.1.1), к которой подключен генератор монохроматического излучения круговой частоты ω , с нагрузкой $Z_{\text{н}} = R_{\text{н}} + iX_{\text{н}}$. Координату z отсчитываем от нагрузки в сторону генератора. Пусть $\dot{U}(z), \dot{I}(z)$ – комплексные амплитуды напряжения и тока в линии, $R_1(z)$, Ом/м – погонное сопротивление, $G_1(z)$, 1/Ом·м – погонная проводимость, $C_1(z)$, Ф/м – погонная емкость, $L_1(z)$, Гн/м – погонная индуктивность линии. Далее рассматриваем *регулярную эквивалентную линию*, в которой погонные параметры не зависят от z . Электрическая эквивалентная схема дифференциально малого участка $(z, z + dz)$ линии показана на рис. 14.2.1 (для удобства схема симметризована), стрелками показаны выбранные направления напряжения и тока.

Телеграфные уравнения

Сопротивление, проводимость, емкость и индуктивность участка $(z, z + dz)$ линии определяются равенствами

$$dR = R_1 dz, \quad dG = G_1 dz, \quad dC = C_1 dz, \quad dL = L_1 dz, \quad (14.2.1)$$

отсюда приращения напряжения и тока на этом участке:

$$d\dot{U} = \dot{I}(dR + i\omega dL) = \dot{I}Z_1 dz; \quad d\dot{I} = \dot{U}(dG + i\omega dC) = \dot{U}Y_1 dz, \quad (14.2.2)$$

где $Z_1 = R_1 + i\omega L_1$ – комплексное погонное сопротивление линии;

$Y_1 = G_1 + i\omega C_1$ – комплексная погонная проводимость линии.

Поделив равенства (14.2.2) на dz , получим *телеграфные уравнения*:

$$\frac{d\dot{U}}{dz} = iZ_1 \dot{I}; \quad \frac{d\dot{I}}{dz} = \dot{U} Y_1, \quad (14.2.3)$$

определяющие связь между напряжением и током в любом сечении эквивалентной ЛП.

Уравнения Гельмгольца (волнового уравнения), коэффициент распространения

Исключим из первого уравнения ток, а из второго – напряжение. Для этого продифференцируем их по z и учтем, что в силу регулярности линии

$$\frac{dZ_1}{dz} = 0; \quad \frac{dY_1}{dz} = 0, \quad (14.2.4)$$

тогда

$$\frac{d^2\dot{U}}{dz^2} = \frac{d\dot{I}}{dz} Z_1; \quad \frac{d^2\dot{I}}{dz^2} = \frac{d\dot{U}}{dz} Y_1, \quad (14.2.5)$$

и подставляя в эти равенства соотношения (14.2.3), получаем *волновые уравнения (уравнения Гельмгольца)*:

$$\frac{d^2\dot{U}}{dz^2} - \gamma^2 \dot{U} = 0; \quad \frac{d^2\dot{I}}{dz^2} - \gamma^2 \dot{I} = 0, \quad (14.2.6)$$

где

$$\gamma = \sqrt{Z_1 Y_1}, \quad (14.2.7)$$

важный частотно-зависимый параметр эквивалентной линии – *коэффициент распространения волны* в линии.

Решение любого из уравнений (14.2.6) вместе с граничными условиями на концах линии полностью описывает распространение волн напряжения и тока в линии, причем достаточно решить одно из уравнений, например, для напряжения, а вторую функцию координаты z можно получить из (14.2.3).

Падающие и отраженные волны, полное напряжение

Хотя достаточно решить одно из уравнений (14.2.6), для удобства дальнейшей интерпретации выпишем решения обоих уравнений. Общее решение однородных линейных уравнений второго порядка типа (14.2.6) хорошо известно и имеет вид:

$$\dot{U}(z) = \dot{A}_U e^{\gamma z} + \dot{B}_U e^{-\gamma z}; \quad \dot{I}(z) = \dot{A}_I e^{\gamma z} + \dot{B}_I e^{-\gamma z}, \quad (14.3.1)$$

где \dot{A}_U, \dot{B}_U – не зависящие друг от друга произвольные постоянные размерности напряжения;

\dot{A}_I, \dot{B}_I – не зависящие друг от друга произвольные постоянные размерности тока.

Решения волновых уравнений в виде (14.3.1) имеют весьма характерный вид: первое слагаемое в каждом решении есть *падающая волна* напряжения (тока), распространяющаяся от генератора к нагрузке, второе слагаемое – *отраженная волна*, распространяющаяся от нагрузки к

генератору. Положив $z = 0$, выясняем смысл коэффициентов: \dot{A}_U, \dot{A}_I – комплексные амплитуды падающих волн напряжения и тока соответственно на (в сечении) нагрузке, \dot{B}_U, \dot{B}_I – комплексные амплитуды отраженных волн напряжения и тока соответственно на (в сечении) нагрузке. Полное напряжение (ток) на нагрузке, как следует из (14.3.1), равно сумме падающего и отраженного напряжений (токов):

$$\dot{U}_{\text{пол нагр}} = \dot{U}(0) = \dot{A}_U + \dot{B}_U; \quad \dot{I}_{\text{пол нагр}} = \dot{I}(0) = \dot{A}_I + \dot{B}_I, \quad (14.3.2)$$

а равенства (14.3.1) показывают, что то же самое справедливо для любого сечения z линии: полное напряжение $\dot{U}(z)$ (ток $\dot{I}(z)$) в любом сечении равно сумме напряжений (токов) падающей и отраженной волн напряжения (тока). Мы пришли к важному выводу: характерной особенностью эквивалентных ЛП как «длинных линий» является возможность существования в них двух волн, распространяющихся навстречу друг другу: падающей, образованной подключенным к линии генератором, и отраженной, обязанной отражению падающей волны от нагрузки; все разнообразие процессов, происходящих в ЛП, определяется амплитудно-фазовыми соотношениями между падающей и отраженной волнами.

Поскольку нагрузка суть пассивный элемент, амплитуды отраженных волн напряжения и тока не могут превышать амплитуд падающих волн:

$$|\dot{B}_U| \leq |\dot{A}_U|; \quad |\dot{B}_I| \leq |\dot{A}_I|. \quad (14.3.3)$$

Чтобы не возникало путаницы, заметим, что направление распространения волн определяется знаком в показателях экспонент в (14.3.1): при выборе направления оси Z как на рис. 12.1.1 знак «плюс» – волна распространяется в отрицательном направлении оси z (падающая); «минус» – в положительном направлении оси z (отраженная). В частности, падающие волны напряжения и тока:

$$\dot{U}_{\text{пад}}(z) = \dot{A}_U e^{\gamma z}; \quad \dot{I}_{\text{пад}}(z) = \dot{A}_I e^{\gamma z}, \quad (14.3.4)$$

а отраженные волны:

$$\dot{U}_{\text{отр}}(z) = \dot{B}_U e^{-\gamma z}; \quad \dot{I}_{\text{отр}}(z) = \dot{B}_I e^{-\gamma z}. \quad (14.3.5)$$

Вторичные параметры ЭЛП (коэффициент затухания, коэффициент фазы, погонное затухание, линия без потерь, длина волны в линии, полная фаза, фазовая скорость, частотная дисперсия, волновое сопротивление)

Локальные параметры эквивалентной ЛП можно считать *первичными параметрами* линии. Рассмотрим теперь *вторичные параметры* эквивалентной ЛП, описывающие свойства распространяющихся в линии волн в целом и выражающиеся через первичные параметры. Начнем с введенного в (14.2.7) коэффициента распространения γ . Поскольку в общем случае γ суть комплексная величина, ее можно записать в виде:

$$\gamma(\omega) = \sqrt{Z_1 Y_1} = \sqrt{(R_1 + i\omega L_1)(G_1 + i\omega C_1)} = \alpha(\omega) + i\beta(\omega). \quad (14.4.1)$$

Действительная часть коэффициента распространения α , $1/\text{м}$, называется *коэффициентом затухания (ослабления)* волны в линии, мнимая часть β , $1/\text{м}$, – *коэффициентом фазы* волны. Смысл этих частотно-зависимых параметров будет понятен, если записать, например, комплексную амплитуду падающей волны напряжения по (14.3.4) в виде:

$$\dot{U}_{\text{пад}}(z) = \dot{A}_U e^{\gamma z} = \dot{A}_U e^{\alpha z} e^{i\beta z}, \quad (14.4.2)$$

т. е. α это логарифмический декремент затухания волны за счет потерь в металле и диэлектрике и, возможно, за счет потерь на излучение в сторону; β это сдвиг фазы на единицу длины линии. Если z_1, z_2 – два сечения линии, отстоящие друг от друга на 1 м, то

$$\alpha = \ln \frac{|\dot{U}_{\text{пад}}(z_1)|}{|\dot{U}_{\text{пад}}(z_2)|}.$$

Обычно вводят *погонное затухание (ослабление)*:

$$\Delta = 20 \lg \frac{|\dot{U}_{\text{пад}}(z_1)|}{|\dot{U}_{\text{пад}}(z_2)|}, \text{ дБ/м,}$$

связанное с α соотношением

$$\Delta = 20\alpha \cdot \lg e \approx 8686\alpha,$$

Записав $\dot{A}_U = |\dot{A}_U| e^{i\varphi_0}$ и перейдя от комплексной амплитуды к мгновенному значению, имеем:

$$u(z, t) = \text{Re}(\dot{U}_{\text{пад}} e^{i\omega t}) = |\dot{A}_U| e^{\alpha z} \cos(\omega t + \beta z + \varphi_0), \quad (14.4.3)$$

откуда смысл β виден еще более ясно.

Из (14.4.1) легко получить выражения α и β через первичные параметры линии [14.5]:

$$\alpha = \sqrt{\frac{1}{2}(R_1 G_1 - \omega^2 L_1 C_1) + \frac{1}{2}\sqrt{(R_1^2 + \omega^2 L_1^2)(G_1^2 + \omega^2 C_1^2)}}; \quad (14.4.4)$$

$$\beta = \sqrt{\frac{1}{2}(\omega^2 L_1 C_1 - R_1 G_1) + \frac{1}{2}\sqrt{(R_1^2 + \omega^2 L_1^2)(G_1^2 + \omega^2 C_1^2)}}. \quad (14.4.5)$$

В важном идеализированном частном случае отсутствия потерь ($R_1 = G_1 = 0$) из этих равенств следует:

$$\alpha = 0, \quad \beta = \omega \sqrt{L_1 C_1}. \quad (14.4.6)$$

Выражение (14.4.3) дает распределение мгновенного напряжения во времени и вдоль линии. Зафиксируем время, например, $t = 0$, и найдем распределение вдоль линии в случае отсутствия потерь ($\alpha = 0$):

$$u(z, 0)_{\alpha=0} = |\dot{A}_U| \cos(\beta z + \varphi_0). \quad (14.4.7)$$

Пространственный период этой косинусоиды

$$\lambda_n = 2\pi/\beta \quad (14.4.8)$$

называется *длиной волны в линии*.

Аргумент косинусоиды в (14.4.3) называется *полной фазой* волны в линии:

$$\Phi(z, t) = \omega t + \beta z + \varphi_0. \quad (14.4.9)$$

Отсюда уравнение точек равных полных фаз $\Phi = const$:

$$\omega t + \beta z + \varphi_0 = const. \quad (14.4.10)$$

Выражая отсюда z и дифференцируя по времени, получаем *фазовую скорость* волны частотой ω :

$$V_\phi = \omega / \beta. \quad (14.4.11)$$

В общем случае β зависит от ω нелинейно (см. (14.4.5)) и V_ϕ зависит от ω , т. е. имеет место *частотная дисперсия*. Однако, в идеализированном случае отсутствия потерь, как следует из (14.4.6), (14.4.11),

$$V_\phi = 1 / \sqrt{L_1 C_1}. \quad (14.4.12)$$

Из электродинамики известно [14.1, 14.5], что $1 / \sqrt{L_1 C_1}$ равно скорости света в той диэлектрической среде без потерь с относительными проницаемостями ϵ, μ , которая заполняет линию передачи. В свою очередь, эта скорость равна $c / \sqrt{\epsilon \mu}$. Таким образом, в линии без потерь скорость распространения волны V_ϕ такая же, как скорость распространения в неограниченном диэлектрике с теми же ϵ, μ :

$$V_\phi = c / \sqrt{\epsilon \mu}, \quad (14.4.13)$$

а ограничивающие идеальные токонесущие поверхности не влияют на скорость распространения в линии [14.5]. В частности, в воздушной линии ($\epsilon \approx 1, \mu \approx 1$), как и следовало ожидать, $V_\phi \approx c$.

Из (14.3.4) следует, что комплексные амплитуды напряжения и тока в падающей волне отличаются лишь коэффициентами \dot{A}_U, \dot{A}_I . Найдем связь между ними, подставив оба равенства (14.3.4) в одно из телеграфных уравнений (14.2.3), например, в первое, и учитывая (14.2.7), тогда получим

$$Z_b = \frac{\dot{U}_{пад}}{\dot{I}_{пад}} = \sqrt{\frac{Z_1}{Y_1}} = \sqrt{\frac{R_1 + i\omega L_1}{G_1 + i\omega C_1}} = |Z_b| e^{i\psi}, \text{ Ом.} \quad (14.4.14)$$

Поскольку Z_b есть отношение комплексных амплитуд напряжения и тока в падающей волне, оно называется *волновым сопротивлением* эквивалентной ЛП. Это один из важнейших вторичных параметров ЛП. Как видно из (14.4.14), в общем случае волновое сопротивление – комплекснозначная функция ω . Однако, в идеализированном случае отсутствия потерь ($R_1 = G_1 = 0$) Z_b действительно, не зависит от ω и равно

$$Z_b|_{\alpha=0} = \sqrt{L_1 / C_1}. \quad (14.4.15)$$

Аналогичным образом находим:

$$\frac{\dot{U}_{\text{отр}}}{\dot{I}_{\text{отр}}} = -Z_{\text{в}}. \quad (14.4.16)$$

Различие в знаках между (14.4.14) и (14.4.16) связано с тем, ток отраженной волны направлен противоположно току падающей волны.

Из (14.4.14) легко выразить модуль $|Z_{\text{в}}|$ и аргумент ψ волнового сопротивления:

$$|Z_{\text{в}}| = \sqrt[4]{\frac{R_1^2 + \omega^2 L_1^2}{G_1^2 + \omega^2 C_1^2}}, \quad (14.4.17)$$

$$\psi = \frac{1}{2} \arctg \frac{\frac{G_1}{\omega C_1} - \frac{R_1}{\omega L_1}}{1 + \frac{G_1 R_1}{\omega^2 L_1 C_1}} \quad (14.4.18)$$

Как и следовало ожидать, в идеализированном случае отсутствия потерь ($R_1 = G_1 = 0$) (14.4.17) переходит в (14.4.15), а $\psi = 0$, т. е. напряжение и ток синфазны.

Частные решения волнового уравнения

Решение граничной задачи

Вернемся к общим решениям (14.3.1). Для нахождения частных решений необходимо задавать граничные условия и определять коэффициенты $\dot{A}_U, \dot{B}_U, \dot{A}_I, \dot{B}_I$. Как отмечено выше, решения для напряжения и тока в (14.3.1) не являются независимыми; действительно, подставив в (14.4.14) и (14.4.16) $z = 0$, получаем

$$\dot{A}_I = \dot{A}_U / Z_{\text{в}}, \quad \dot{B}_I = -\dot{B}_U / Z_{\text{в}}, \quad (14.5.1)$$

и решения (14.3.1) можно переписать в виде

$$\dot{U}(z) = \dot{A}_U e^{\gamma z} + \dot{B}_U e^{-\gamma z}, \quad \dot{I}(z) = (\dot{A}_U e^{\gamma z} - \dot{B}_U e^{-\gamma z}) / Z_{\text{в}}. \quad (14.5.2)$$

Таким образом, для получения частного решения необходимо определить только два коэффициента \dot{A}_U, \dot{B}_U . Удобно и физично в качестве граничных условий взять напряжение и ток на нагрузке ($z = 0$):

$$\dot{U}(0) = \dot{U}_{\text{н}}; \quad \dot{I}(0) = \dot{I}_{\text{н}}. \quad (14.5.3)$$

Тогда из (14.5.2) при $z = 0$ найдем:

$$\dot{A}_U = \frac{1}{2}(\dot{U}_{\text{н}} + \dot{I}_{\text{н}} Z_{\text{в}}); \quad \dot{B}_U = \frac{1}{2}(\dot{U}_{\text{н}} - \dot{I}_{\text{н}} Z_{\text{в}}). \quad (14.5.4)$$

Подставив (14.5.4) в (14.5.2) и учитывая равенства

$$\text{sh}(\gamma z) = \frac{1}{2}(e^{\gamma z} - e^{-\gamma z}); \quad \text{ch}(\gamma z) = \frac{1}{2}(e^{\gamma z} + e^{-\gamma z}), \quad (14.5.5)$$

получим:

$$\dot{U}(z) = \dot{U}_n \operatorname{ch}(\gamma z) + \dot{I}_n Z_b \operatorname{sh}(\gamma z); \quad \dot{I}(z) = \dot{I}_n \operatorname{ch}(\gamma z) + (\dot{U}_n / Z_b) \operatorname{sh}(\gamma z). \quad (14.5.6)$$

Если в (14.3.4) и (14.3.5) положить $z = 0$, то окажется, что комплексные коэффициенты суть комплексные амплитуды падающих и отраженных волн напряжения и тока на нагрузке:

$$\dot{A}_U = \dot{U}_{\text{пад}}(0); \quad \dot{A}_I = \dot{I}_{\text{пад}}(0); \quad \dot{B}_U = \dot{U}_{\text{отр}}(0); \quad \dot{B}_I = \dot{I}_{\text{отр}}(0). \quad (14.5.7)$$

Если эти комплексные амплитуды считать заданными, то равенства (14.3.1) также можно считать частными решениями волновых уравнений (14.2.6). Разница между частными решениями (14.3.1) и частными решениями (14.5.6) состоит в том, что первые выражены через комплексные амплитуды падающих и отраженных волн на нагрузке, а вторые – через комплексные амплитуды полных напряжения и тока на нагрузке.

В идеализированном случае отсутствия потерь ($R_l = G_l = 0$) $\gamma = i\beta$ и с учетом равенств

$$\operatorname{ch}(i\beta z) = \cos(\beta z); \quad \operatorname{sh}(i\beta z) = i \sin(\beta z) \quad (14.5.8)$$

решения (14.5.6) можно записать в виде

$$\dot{U}(z) = \dot{U}_n \cos(\beta z) + i \dot{I}_n Z_b \sin(\beta z); \quad \dot{I}(z) = \dot{I}_n \cos(\beta z) + i(\dot{U}_n / Z_b) \sin(\beta z). \quad (14.5.9)$$

Это же выражение можно получить и из (14.3.1), используя (14.3.2) и (14.5.1). С другой стороны, учтя (14.5.7), можно в случае отсутствия потерь записать (14.3.1) в виде:

$$\dot{U}(z) = \dot{U}_{\text{пад}}(0)e^{i\beta z} + \dot{U}_{\text{отр}}(0)e^{-i\beta z}; \quad \dot{I}(z) = \dot{I}_{\text{пад}}(0)e^{i\beta z} + \dot{I}_{\text{отр}}(0)e^{-i\beta z}. \quad (14.5.10)$$

Коэффициент отражения и входное сопротивление

Обозначим отношение комплексных амплитуд напряжений падающей и отраженной волн на нагрузке:

$$\Gamma = \frac{\dot{U}_{\text{отр}}(0)}{\dot{U}_{\text{пад}}(0)} = \frac{\dot{B}_U}{\dot{A}_U}. \quad (14.6.1)$$

Комплексная величина $\Gamma = |\Gamma|e^{i\psi_\Gamma}$ называется *коэффициентом отражения по напряжению*, является важным параметром системы «ЛП+нагрузка» и, как показано ниже, зависит от соотношения нагрузки Z_n и волнового сопротивления ЛП Z_b . Аналогичное отношение токов, которое можно назвать *коэффициентом отражения по току*, в силу (14.5.1), равно:

$$\Gamma_I = \frac{\dot{I}_{\text{отр}}(0)}{\dot{I}_{\text{пад}}(0)} = \frac{\dot{B}_I}{\dot{A}_I} = -\frac{\dot{B}_U}{\dot{A}_U} = -\Gamma, \quad (14.6.2)$$

и на практике не применяется, поэтому слова «по напряжению» в названии Γ можно отбросить.

Теперь решения (14.5.10) можно переписать в виде:

$$\dot{U}(z) = \dot{U}_{\text{пад}}(0) \cdot (e^{i\beta z} + \Gamma e^{-i\beta z}); \quad \dot{I}(z) = \dot{I}_{\text{пад}}(0) \cdot (e^{i\beta z} - \Gamma e^{-i\beta z}). \quad (14.6.3)$$

Другим важным параметром системы «ЛП+нагрузка» является *входное сопротивление* нагруженной линии

$$Z_{\text{вх}}(z) = \dot{U}(z) / \dot{I}(z), \quad (14.6.4)$$

связанное с трансформирующими свойствами отрезка линии.

ЛП с малыми потерями

Как показано выше, такие вторичные параметры ЛП, как коэффициент ослабления α , коэффициент фазы β , фазовая скорость V_ϕ , длина волны в линии λ_λ , волновое сопротивление Z_λ , и такие функции, как $\dot{U}_{\text{пад}}(z)$, $\dot{U}_{\text{отр}}(z)$ и др., выражаются через первичные параметры R_1 , G_1 , L_1 , C_1 . В общем случае произвольных первичных параметров эти выражения сложны и громоздки (см. (14.4.4), (14.4.5)), что мешает практическому использованию теории эквивалентных ЛП. С другой стороны, в предположении отсутствия потерь (и в металле, и в диэлектрике) эти соотношения резко упрощаются, но само это предположение является теоретической идеализацией, и использование получаемых простых соотношений создает ошибки, исследование величин которых представляет собой непростую задачу.

Оказывается, однако, что современные ЛП часто, если не сказать всегда, принадлежат к классу *линий с малыми потерями*, для которых приближенные соотношения вторичных и первичных параметров компромиссны по сложности, а оценки ошибок очевидны. Примем в качестве критерия малости потерь линии условия [14.6]:

$$\frac{R_1}{\omega L_1} \ll 1; \quad \frac{G_1}{\omega C_1} \ll 1. \quad (14.7.1)$$

Представив постоянную распространения γ в виде:

$$\gamma = \sqrt{(R_1 + i\omega L_1)(G_1 + i\omega C_1)} = i\omega \sqrt{L_1 C_1} \sqrt{[1 + R_1/(i\omega L_1)][1 + G_1/(i\omega C_1)]}, \quad (14.7.2)$$

разложим ее по степеням отношений $R_1/\omega L_1$, $G_1/\omega C_1$, удерживая члены второго порядка этих отношений, тогда получим:

$$\alpha \approx \frac{1}{2} \left(\frac{R_1}{Z_{\text{в0}}} + G_1 Z_{\text{в0}} \right); \quad (14.7.3)$$

$$\beta \approx \beta_0 \left[1 + \frac{1}{8} \left(\frac{R_1}{\omega L_1} - \frac{G_1}{\omega C_1} \right)^2 \right], \quad (14.7.4)$$

где $Z_{\text{в0}} = \sqrt{L_1/C_1}$ – волновое сопротивление линии без потерь (см. (14.4.15)),

$\beta_0 = \omega \sqrt{L_1 C_1}$ – коэффициент фазы линии без потерь (см. (14.4.6)).

Относительная ошибка при использовании этих формул определяется модулем первого отброшенного члена в разложении.

В современных ЛП потери в диэлектрике пренебрежимо малы по сравнению с потерями в металле, это означает, что члены с G_1 в обеих формулах пренебрежимы, и формулы упрощаются:

$$\alpha \approx R_1 / 2Z_{в0} ; \quad (14.7.5)$$

$$\beta \approx \beta_0 [1 + \frac{1}{8} (\frac{R_1}{\omega L_1})^2] . \quad (14.7.6)$$

Для конкретной ЛП, т. е. при заданных первичных параметрах, условия (14.7.1) выполняются тем лучше, чем больше частота; начиная с некоторой частоты и для более высоких частот становится приемлемым нулевое приближение в используемом разложении ($\alpha = 0$, $\beta = \beta_0$), т. е. предположение об отсутствии потерь.

Связь коэффициента отражения с нагрузкой

Пусть эквивалентная ЛП нагружена на, в общем случае – комплексное, сопротивление Z_n (рис. 12.1.1). Комбинируя (14.3.2), (14.4.14), (14.4.16), (14.5.1) и очевидное равенство

$$\dot{U}_n = \dot{I}_n Z_n , \quad (14.8.1)$$

получим [14.5]:

$$Z_n = Z_b \frac{\dot{U}_{пад}(0) + \dot{U}_{отр}(0)}{\dot{U}_{пад}(0) - \dot{U}_{отр}(0)} . \quad (14.8.2)$$

Поделив числитель и знаменатель на $\dot{U}_{пад}$ и учтя определение (14.6.1) коэффициента отражения Γ , найдем:

$$\frac{Z_n}{Z_b} = \xi = \frac{1 + \Gamma}{1 - \Gamma} , \quad (14.8.3)$$

откуда

$$\Gamma = \frac{Z_n - Z_b}{Z_n + Z_b} = \frac{\xi - 1}{\xi + 1} , \quad (14.8.4)$$

т. е. коэффициент отражения Γ полностью определяется (комплексным) отношением ξ сопротивления нагрузки и волнового сопротивления.

В большинстве случаев при конструировании радиотехнических устройств СВЧ – диапазона требуется добиваться, насколько возможно, согласования ЛП и нагрузки. Так, «...со знанием параметров рассеяния транзистора конструирование усилительных, осцилляторных и смесительных цепей сводится к задаче импедансного согласования» [14.3]. В измерительных приборах СВЧ-диапазона проблема согласования стоит особенно остро, т. к. именно рассогласованиям обязан ряд ошибок измерения. Из формулы (14.8.4) видно, что единственный способ избежать отражения от нагрузки, т. е. добиться равенства $\Gamma = 0$, – это выбрать сопротивление нагрузки равным волновому сопротивлению линии: $Z_n = Z_b$. Если волновое сопротивление ЛП чисто активно: $Z_b = R_b$, то согласующее сопротивление должно быть также активным и равным волновому: $R_n = R_b$.

Подставляя в (14.8.4) выражения $Z_n = R_n + iX_n$; $Z_b = R_b + iX_b$, получим:

$$\Gamma = \rho e^{i\psi_\Gamma} = \frac{(R_n - R_b) + i(X_n - X_b)}{(R_n + R_b) + i(X_n + X_b)}. \quad (14.8.5)$$

На практике ЛП рассчитываются и изготавливаются так, чтобы волновое сопротивление было чисто активным, т. е. $X_b = 0$. Учитывая это, перепишем предыдущее выражение:

$$\Gamma = \frac{(R_n - R_b) + iX_n}{(R_n + R_b) + iX_n}. \quad (14.8.6)$$

Отсюда найдем модуль коэффициента отражения:

$$\rho = \sqrt{\frac{(R_n - R_b)^2 + X_n^2}{(R_n + R_b)^2 + X_n^2}}. \quad (14.8.7)$$

Поскольку $R_n \geq 0$ и $R_b \geq 0$ (вследствие пассивности нагрузки и ЛП), знаменатель подкоренного выражения не больше числителя, т. е.

$$\rho \leq 1. \quad (14.8.8)$$

Это условие эквивалентно (14.3.3) и означает, что амплитуда отраженной волны не превосходит амплитуду падающей волны.

Распределение амплитуд напряжения и тока вдоль нагруженной линии

Интерференция волн

Из вышеизложенного следует, что в ЛП, нагруженной на сопротивление Z_n и возбуждаемой монохроматическим сигналом частоты ω , существует две волны – падающая и отраженная, распространяющиеся в противоположные стороны. В каждой точке (сечении) линии каждая из этих волн имеет свою амплитуду и свою фазу, и происходит *интерференция* этих волн. Рассмотрим результаты интерференции в идеализированном случае отсутствия потерь.

Относительные распределения амплитуд, их свойства

Формулы (14.6.3) дают распределения комплексных амплитуд полного напряжения и тока вдоль линии, коэффициент отражения имеет вид (см. (14.8.5)) $\Gamma = \rho e^{i\psi_\Gamma}$, при действительном волновом сопротивлении ($Z_b = R_b$) $\psi_\Gamma = \psi_n$, где ψ_n – аргумент комплексного сопротивления нагрузки: $Z_n = |Z_n| e^{i\psi_n}$. Находя из (14.6.3) модули (амплитуды) полного напряжения и тока и деля их на те же модули в точке $z = 0$ соответственно, находим относительные распределения амплитуд напряжения и тока вдоль линии:

$$\frac{|\dot{U}(z)|}{|\dot{U}_{\text{пад}}(0)|} = \sqrt{1 + 2\rho \cos(2\beta z - \psi_n) + \rho^2}; \quad (14.9.1)$$

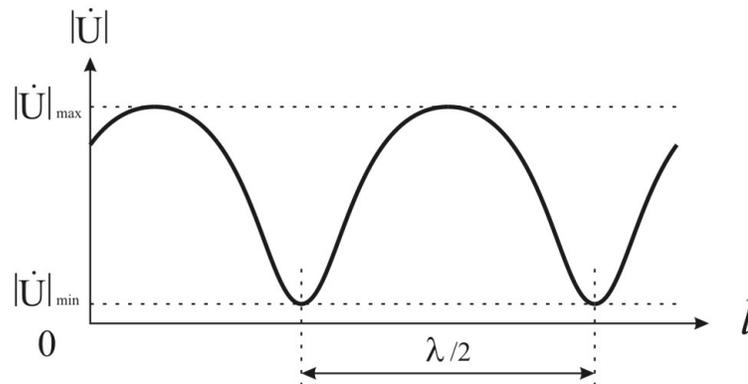
$$\frac{|\dot{I}(z)|}{|\dot{I}_{\text{пад}}(0)|} = \sqrt{1 - 2\rho \cos(2\beta z - \psi_n) + \rho^2}. \quad (14.9.2)$$

Непосредственно из (14.9.1), (14.9.2) следует:

1) Амплитуды напряжения и тока в линии – периодические функции продольной координаты z . Период L удовлетворяет условию $2\beta L = 2\pi$, а поскольку в случае отсутствия потерь (см. (14.4.8)) $\beta = 2\pi/\lambda_d$, то $L = \lambda_d$.

2) Относительные амплитуды напряжений и токов колеблются между $1+\rho$ (максимум) и $1-\rho$ (минимум). Максимум относительной амплитуды напряжения достигается при

$z = (n + \frac{\Psi_H}{2\pi}) \frac{\lambda_d}{2}$, $n = 0, 1, 2, \dots$; минимум – при $z = (n + \frac{1}{2} + \frac{\Psi_H}{2\pi}) \frac{\lambda_d}{2}$, $n = 0, 1, 2, \dots$; максимум тока находится в точке минимума напряжения, и наоборот.



Пучности и узлы в ЛП в режиме смешанных волн

3) При $\rho \neq 0$ интерференция падающей и отраженной волн образует *стоячую волну*, в которой максимумы называют *пучностями*, а минимумы – *узлами* (рис. 14.9.1). Степень *глубины* стоячей волны определяется лишь величиной ρ : при $\rho=1$ глубина максимальна, в пучности относительная амплитуда равна 2, в узле – равна 0 (такую картину иногда называют *чисто стоячей волной*); при $0 < \rho < 1$ пучность ниже двух, а узел выше нуля; при $\rho = 0$ стоячей волны нет: отраженная волна отсутствует, относительные амплитуды напряжения и тока при всех z равны 1, имеет место *режим бегущей волны*.

4) Фазовый угол ψ влияет лишь на координаты узлов и пучностей, но не на значения амплитуды в них.

5) Как видно из рис. 14.9.1, минимум амплитуды вблизи узла выражен более резко, чем максимум амплитуды вблизи пучности. Это свойство стоячей волны важно в измерениях с помощью *длинных линий*: измерение положения пучности менее чувствительно, чем положения узла, зато проводится при большем «отношении сигнал-шум».

6) Для исчерпывающего описания распределения амплитуд вдоль линии достаточно изучить поведение напряжения по (14.9.1), поскольку распределение тока (14.9.2) повторяет распределение напряжения, но со смещением вдоль линии на $\lambda_d/2$.

7) Для количественного описания глубины стоячей волны вводят *коэффициент стоячей волны*, равный отношению максимальной и минимальной амплитуд:

$$k_{CB} = \frac{|\dot{U}|_{\max}}{|\dot{U}|_{\min}} = \frac{|i|_{\max}}{|i|_{\min}} = \frac{1+\rho}{1-\rho}, \quad (14.9.3)$$

откуда

$$\rho = \frac{k_{CB} - 1}{k_{CB} + 1}. \quad (14.9.4)$$

Коэффициент стоячей волны – скалярная величина, принимающая значения в диапазоне

$$1 \leq k_{CT} \leq \infty. \quad (14.9.5)$$

Левая граница этого диапазона соответствует полному согласованию, $\rho = 0$, т. е. режиму бегущей волны, правая – полному рассогласованию, $\rho = 1$, т. е. режиму чисто стоячей волны. И тот, и другой режимы на практике не достижимы в полной мере и являются математической абстракцией, приближение к которой возможно в той или иной степени. Идеально согласованная нагрузка $R_H = R_B$ невозможна из-за тепловых потерь, различных нерегулярностей и неоднородностей, обусловленных конечной точностью изготовления линии, наличием элементов крепления, несовершенством разъемов и других факторов, вызывающих появление отраженной волны. Идеальная полностью отражающая нагрузка невозможна из-за паразитных емкости, индуктивности и утечек.

Режимы работы ЛП без потерь

Обычно режимы в нагруженной ЛП разделяют на три класса:

- 1) *Режим бегущей волны*: линия согласована, $R_H = R_B$, $\Gamma = 0$, отраженная волна отсутствует, распространяется только падающая волна, которая полностью поглощается нагрузкой, амплитуды напряжения и тока постоянны вдоль линии.
- 2) *Режим чисто стоячей волны*: $\Gamma = \pm 1$, в сечении $z = 0$ амплитуда отраженной волны равна амплитуде падающей волны, энергия падающей волны полностью отражается от нагрузки и возвращается обратно в генератор, $\rho = 1$, $k_{CB} = \infty$.
- 3) *Режим смешанных волн*: $0 < \rho < 1$, $1 < k_{CB} < \infty$, часть энергии падающей волны поглощается нагрузкой, остальная ее часть в виде отраженной волны возвращается обратно в генератор, стоячая волна имеет неполную глубину, т. е. в пучности относительная глубина меньше двух, в узле относительная глубина больше нуля.

Канонические нагрузки

Рассмотрим распределения вдоль линии без потерь амплитуд напряжения и тока и входного сопротивления для нагрузок: согласованной резистивной нагрузки, холостого хода, короткого замыкания, емкости, индуктивности [14.6]. Эти нагрузки можно назвать *каноническими*; примеры, связанные с ними, помогают лучше понять структуру волн в нагруженной линии, а также имеют самостоятельный интерес.

Согласованная нагрузка

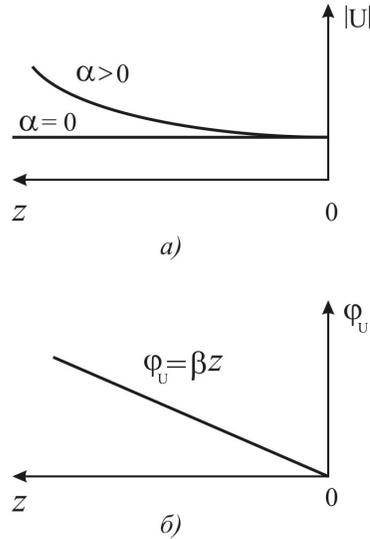
Далее считаем волновое сопротивление чисто активным ($Z_B = R_B$), тогда согласованная нагрузка – сопротивление, равное волновому ($R_H = R_B$). Как указано выше, при такой нагрузке получается режим бегущей волны с постоянными вдоль линии амплитудами напряжения и тока и линейной фазой (рис. 14.11.1). Напряжение и ток в этом режиме (см. (14.5.10)) равны:

$$\dot{U}(z) = \dot{U}_{\text{пад}}(0)e^{i\beta z}; \quad \dot{I}(z) = \dot{I}_{\text{пад}}(0)e^{i\beta z}, \quad (14.11.1)$$

поэтому

$$Z_{\text{вх}}(z) = \frac{\dot{U}(z)}{\dot{I}(z)} = \frac{\dot{U}_{\text{пад}}(0)}{\dot{I}_{\text{пад}}(0)} = R_{\text{н}} = R_{\text{в}}. \quad (14.11.2)$$

Таким образом, в режиме бегущей волны входное сопротивление любого отрезка нагруженной линии независимо от его длины в точности равно волновому сопротивлению линии.



Амплитуда напряжения (а) и фаза (б) в ЛП при согласованной нагрузке

Холостой ход

В этом случае $Z_{\text{н}} = \infty$, $\dot{I}_{\text{н}} = 0$, из (14.8.4) следует, что $\Gamma = 1$, а это означает, что $\dot{U}_{\text{отр}}(0) = \dot{U}_{\text{пад}}(0)$, $\psi_{\text{н}} = 0$, $|\dot{U}(0)| = 2|\dot{U}_{\text{пад}}(0)|$, имеет место режим чисто стоячей волны, при $z = 0$ наблюдается пучность напряжения и узел тока: $\dot{I}_{\text{отр}}(0) = -\dot{I}_{\text{пад}}(0)$. Из (14.5.9), также как из (14.9.1), (14.9.2), следует:

$$\dot{U}(z) = 2\dot{U}_{\text{пад}}(0)\cos(\beta z); \quad \dot{I}(z) = i[2\dot{U}_{\text{пад}}(0)/Z_{\text{в}}]\sin(\beta z), \quad (14.11.3)$$

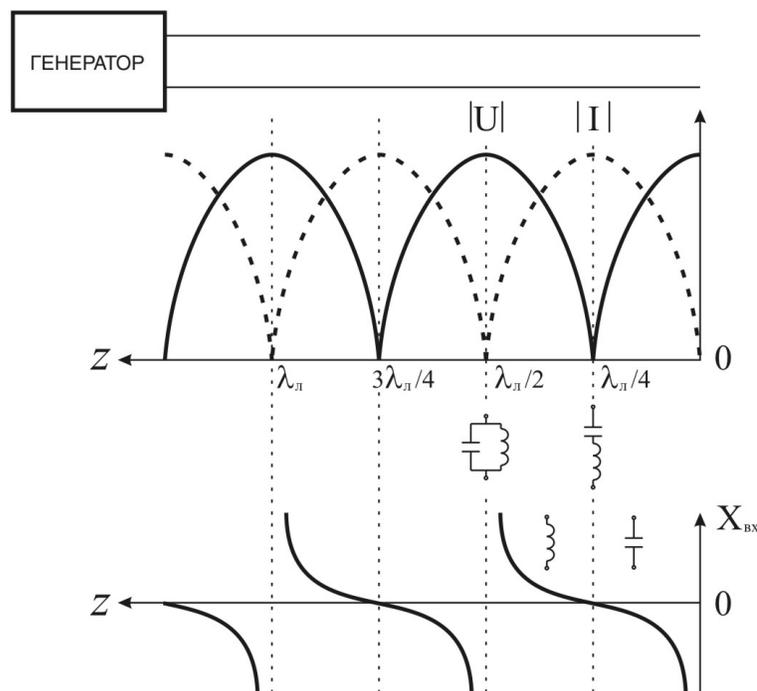
откуда

$$Z_{\text{вх}} = \dot{U}(z)/\dot{I}(z) = -iZ_{\text{в}}\text{ctg}(\beta z) = iX_{\text{вх}}. \quad (14.11.4)$$

Рис. 14.11.2 иллюстрирует соотношения (14.11.3), (14.11.4). Из рисунка и соотношений следует:

- в линии устанавливается чисто стоячая волна с пучностью напряжения и узлом тока на конце линии;

- напряжение, ток и входное сопротивление периодичны с периодом $\lambda_d / 2$;
- входное сопротивление чисто мнимое всюду, кроме точек с координатами $z = n\lambda_d / 4$, $n = 0, 1, 2, \dots$;
- если длина разомкнутой линии меньше $\lambda_d / 4$, то такая линия эквивалентна емкости;
- если длина такой линии равна $\lambda_d / 4$, то она эквивалентна последовательному резонансному на данной частоте контуру и имеет нулевое входное сопротивление;
- если длина такой линии находится в интервале $\lambda_d / 4 < L < \lambda_d / 2$, то она эквивалентна индуктивности;
- если длина такой линии равна $\lambda_d / 2$, то она эквивалентна параллельному резонансному на данной частоте контуру и имеет бесконечно большое входное сопротивление.



Амплитуды напряжения и тока (верхний рисунок) и реактивное входное сопротивление (нижний рисунок) в ЛП, нагруженной на холостой ход

Короткое замыкание

В этом случае $Z_n = 0$, $\dot{U}_n = 0$, из (14.8.4) следует, что $\Gamma = -1$, а это означает, что $\dot{I}_{отр}(0) = \dot{I}_{пад}(0)$, $\psi_n = \pi$, $|\dot{I}(0)| = 2|\dot{I}_{пад}(0)|$, имеет место режим чисто стоячей волны, при $z = 0$ наблюдается пучность тока и узел напряжения: $\dot{U}_{отр}(0) = -\dot{U}_{пад}(0)$. Из (14.5.9), также как из (14.9.1), (14.9.2), следует:

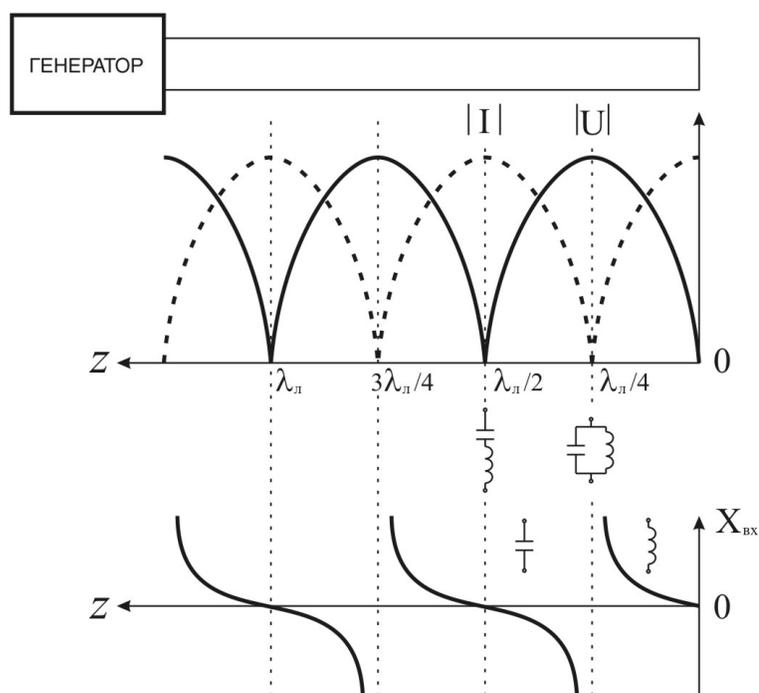
$$\dot{U}(z) = 2i\dot{I}_{пад}(0)Z_n \sin(\beta z); \quad \dot{I}(z) = 2\dot{I}_{пад}(0) \cos(\beta z), \quad (14.11.5)$$

откуда

$$Z_{\text{вх}} = iZ_{\text{в}} \operatorname{tg}(\beta z) = iX_{\text{вх}}. \quad (14.11.6)$$

Рис. 14.11.3 иллюстрирует соотношения (14.11.5), (14.11.6). Из рисунка и соотношений следует:

- в линии устанавливается чисто стоячая волна с пучностью тока и узлом напряжения на конце линии;
- напряжение, ток и входное сопротивление периодичны с периодом $\lambda_{\text{л}}/2$;
- входное сопротивление чисто мнимое всюду, кроме точек с координатами $z = n\lambda_{\text{л}}/4$, $n = 0, 1, 2, \dots$;
- при длинах замкнутой линии $L < \lambda_{\text{л}}/4$, $L = \lambda_{\text{л}}/4$, $\lambda_{\text{л}}/4 < L < \lambda_{\text{л}}/2$, $L = \lambda_{\text{л}}/2$ она эквивалентна индуктивности, параллельному на данной частоте контуру, емкости, последовательному на данной частоте контуру, соответственно.



Амплитуды напряжения и тока (верхний рисунок) и реактивное входное сопротивление (нижний рисунок) в ЛП, нагруженной на короткое замыкание

Емкость

Пусть линия нагружена на емкость C . В этом случае $Z_{\text{н}} = iX_{\text{н}}$, $X_{\text{н}} < 0$. Из (14.8.4) следует, что

$$\Gamma = -\frac{Z_{\text{в}} - iX_{\text{н}}}{Z_{\text{в}} + iX_{\text{н}}}, \quad (14.11.7)$$

откуда

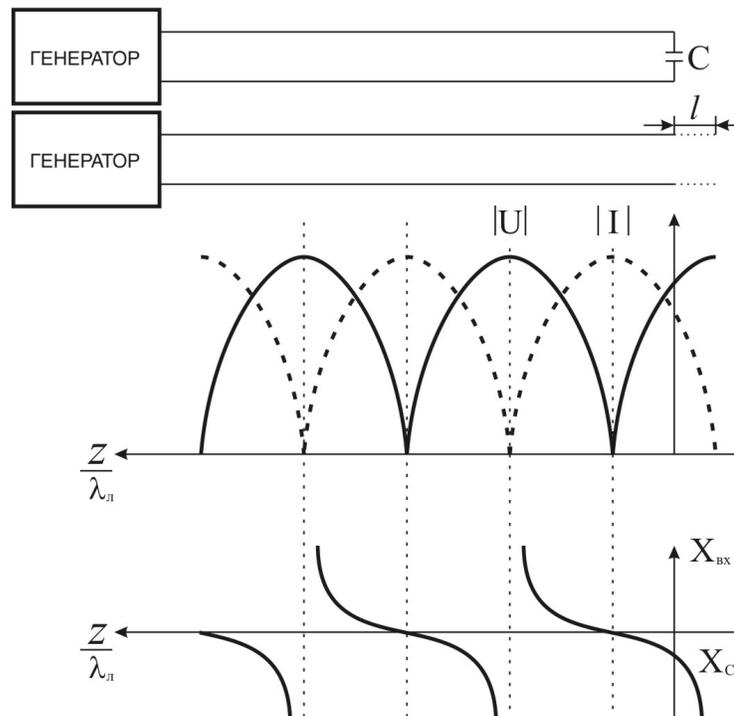
$$\rho = 1, \quad \psi_n = \pi - 2 \operatorname{arctg} \frac{X_n}{Z_b}. \quad (14.11.8)$$

Подставляя эти формулы в (14.9.1), получаем распределение модуля напряжения вдоль линии:

$$|\dot{U}(z)| = 2|\dot{U}_{\text{пад}}(0)| \left| \sin\left(\beta z + \operatorname{arctg} \frac{X_n}{Z_b}\right) \right|, \quad (14.11.9)$$

а распределение модуля тока сдвинуто на $\lambda_n/4$ (рис. 14.11.4). Таким образом, в этом случае также имеет место режим чисто стоячей волны, но в конце линии нет ни пучности, ни узла. Ближайшими к нагрузке экстремумами будут узел напряжения и пучность тока, они отстоят от конца линии на $l = (1/\beta) \operatorname{arctg}(\omega C Z_b)$.

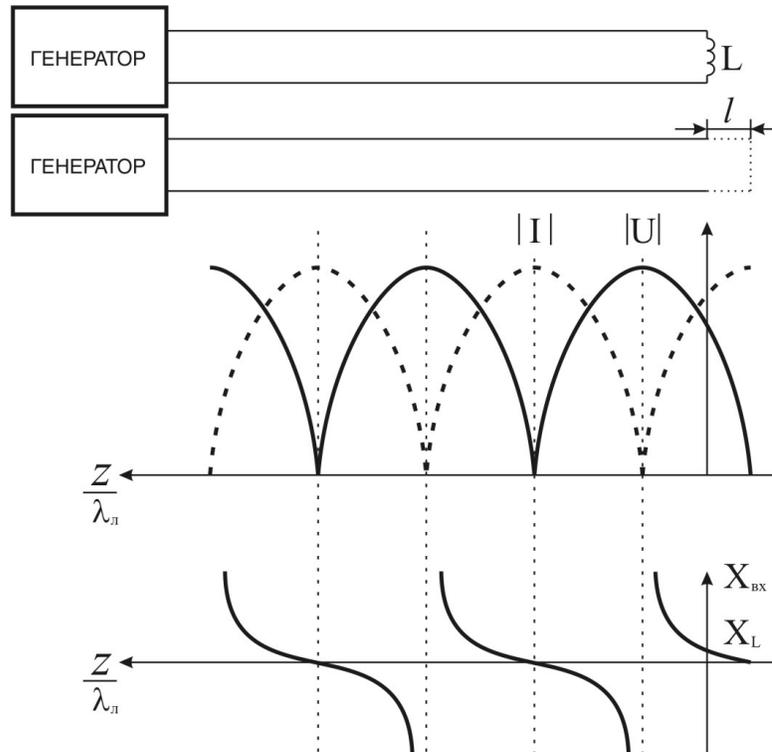
Амплитуды напряжения и тока (верхний рисунок) и реактивное входное сопротивление (нижний рисунок) в ЛП, нагруженной на емкость



Амплитуды напряжения и тока (верхний рисунок) и реактивное входное сопротивление (нижний рисунок) в ЛП, нагруженной на емкость

Индуктивность

Пусть линия нагружена на индуктивность L . В этом случае $Z_n = iX_n$, $X_n > 0$. Очевидно, что распределение модулей напряжения и тока, а также входного сопротивления (рис. 14.11.5), будет повторять картину в случае емкостной нагрузки, но со сдвигом в противоположную сторону на $l = (1/\beta) \operatorname{arctg}(Z_b / \omega L)$, так что ближайшими к нагрузке экстремумами будут пучность напряжения и узел тока.



Амплитуды напряжения и тока (верхний рисунок) и реактивное входное сопротивление (нижний рисунок) в ЛП, нагруженной на индуктивность

Распределение в линии, нагруженной на активное сопротивление, не равное волновому

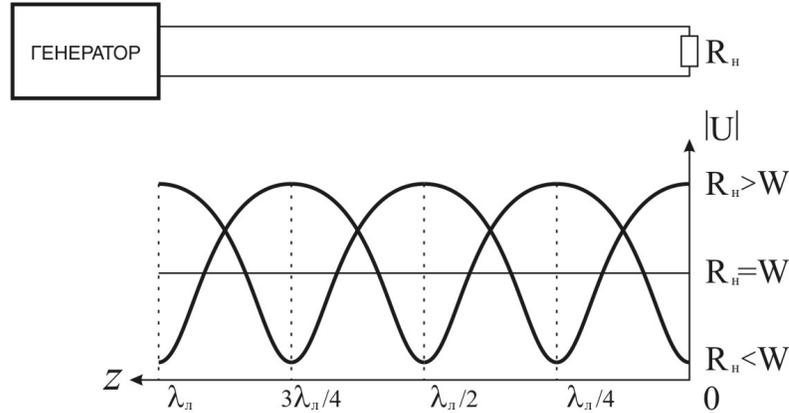
Пусть $Z_b = R_b$, $Z_n = R_n \neq R_b$. Как следует из (14.8.4), при $R_n > R_b$ коэффициент отражения Γ будет вещественным положительным числом и, как следует из (14.9.1), распределение модуля напряжения вдоль линии:

$$|\dot{U}(z)| = |\dot{U}_{\text{пад}}(0)| \cdot \sqrt{1 + 2\rho \cos 2\beta z + \rho^2}, \quad (14.12.1)$$

а при $R_n < R_b$ Γ будет вещественным отрицательным числом и

$$|\dot{U}(z)| = |\dot{U}_{\text{пад}}(0)| \cdot \sqrt{1 - 2\rho \cos 2\beta z + \rho^2}. \quad (14.12.2)$$

Назначим какое-либо $0 < \rho < 1$ и построим зависимость $|\dot{U}(z)|$ (рис. 14.12.1). Имеет место режим смешанных волн, на конце линии достигаются экстремумы: пучность напряжения и узел тока при $R_n > R_b$, пучность тока и узел напряжения – при $R_n < R_b$.



Амплитуды напряжения в ЛП, нагруженной на активное сопротивление

Распределение фазы вдоль нагруженной линии

В стоячей волне той или иной глубины, образованной интерференцией падающей и отраженной волн, вдоль линии меняется не только амплитуда, но и фаза. Рассмотрим этот вопрос на примере напряжения в ЛП без потерь. Запишем первую формулу в (14.6.3) в другом виде:

$$\begin{aligned} \dot{U}(z) &= \dot{U}_{\text{пад}}(0)(e^{i\beta z} + \Gamma e^{-i\beta z}) = \dot{U}_{\text{пад}}(0)e^{i\beta z}(1 + \Gamma e^{-2i\beta z}) = \\ &= \dot{U}_{\text{пад}}(0)e^{i\beta z}[1 + \rho \cos(2\beta z - \psi_{\text{н}}) - i\rho \sin(2\beta z - \psi_{\text{н}})]. \end{aligned} \quad (14.13.1)$$

Отсюда получаем *фазу* полного напряжения как аргумент комплексного числа:

$$\varphi(z) = \beta z - \operatorname{arctg} \frac{\rho \sin(2\beta z - \psi_{\text{н}})}{1 + \rho \cos(2\beta z - \psi_{\text{н}})}. \quad (14.13.2)$$

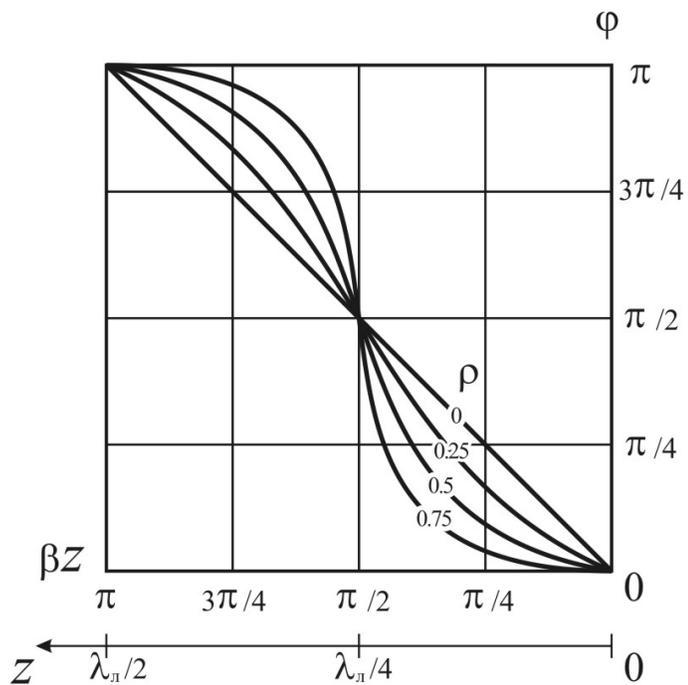
Учитывая соотношение $\lambda_{\text{н}} = 2\pi/\beta$, для случая резистивной нагрузки ($\psi_{\text{н}} = 0$) найдем:

$$\varphi(z) = \frac{2\pi z}{\lambda_{\text{н}}} - \operatorname{arctg} \frac{\rho \sin(4\pi z/\lambda_{\text{н}})}{1 + \rho \cos(4\pi z/\lambda_{\text{н}})}. \quad (14.13.3)$$

Рис. 14.13.1 иллюстрирует это соотношение для нескольких значений ρ . При $\rho = 0$ отраженной волны нет и из (14.13.3), как и следовало ожидать, получаем $\varphi(z; \rho = 0) = 2\pi z/\lambda_{\text{н}}$. По мере роста ρ кривые становятся все более нелинейными, при $\rho = 1$ график переходит в «ступеньку»:

$$\varphi(z) = \begin{cases} 0, & 0 \leq z \leq \lambda_{\text{н}}/4; \\ \pi, & \lambda_{\text{н}}/4 < z \leq \lambda_{\text{н}}/2, \end{cases} \quad (14.13.4)$$

т. е. при переходе через узел фаза скачком изменяется на π .



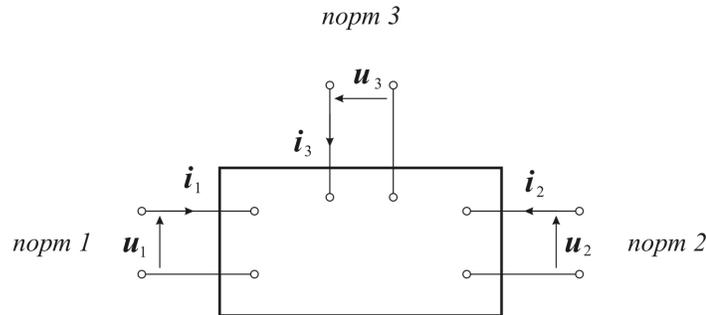
Фаза полного напряжения в ЛП при различных значениях модуля коэффициента отражения ρ

ПАРАМЕТРЫ РАССЕЙНИЯ

Классические матричные описания линейных многополюсников

Линейный многополюсник

Концепция $2n$ -полюсников хорошо известна [10.1, 10.2] и изложена в любом стандартном учебнике по радиотехническим цепям или радиотехнике. Этот подход позволяет абстрагироваться от конкретной схемы и внутреннего устройства цепи и использовать некие общие параметры $2n$ -полюсника, связь между которыми и есть искомые характеристики.



Токи и напряжения в трехпортовой цепи

Многополюсники (n -портовые цепи) разделяются на линейные и нелинейные. Далее рассматриваются только *линейные многополюсники*. Это понятие синонимично понятию *линейная цепь*. В традиционных представлениях многополюсников первичными параметрами возбужденного $2n$ -полюсника являются n выходных токов $i_k, k = 1, \dots, n$, и n выходных напряжений $u_k, k = 1, \dots, n$; ниже специально для диапазона СВЧ будет введена другая система первичных параметров. Пока нет необходимости конкретизировать вид математического описания токов и напряжений как первичных параметров (действительные величины, комплексные величины), их зависимость от времени и т. п.

Разбиение первичных параметров на группы А и В

На рис. 10.1.1 показан 6-полюсник (3-портовая цепь) с произвольно выбранными положительными направлениями токов и напряжений. Разделим $2n$ первичных параметров на две группы по n параметров в каждой (необязательно, чтобы в одной группе были только токи, а в другой – только напряжения, а поэтому необязательно, чтобы параметры одной группы имели одинаковую размерность). Обозначим одну из групп буквой А, другую – буквой В.

Матричное соотношение групп параметров

В силу линейности многополюсника, первичные параметры связаны только линейными соотношениями друг с другом, а из законов Кирхгофа следует, что независимых из этих соотношений ровно n . Удобно эти соотношения записать в виде выражений каждого параметра группы В со всеми параметрами группы А. Единственный линейный оператор в конечномерном n -мерном пространстве – квадратная матрица порядка n . Поэтому систему искомых соотношений между первичными параметрами $2n$ -полюсника можно записать в виде:

$$\mathbf{b} = \mathbf{M}\mathbf{a}, \quad (10.1.1)$$

где $\mathbf{a} = \{a_i\}$ – n -мерный вектор-столбец, составленный из параметров группы А;

$\mathbf{b} = \{b_i\}$ – n -мерный вектор-столбец, составленный из параметров группы В;

$\mathbf{M} = \{M_{ij}\}$ – матрица ($n \times n$) с элементами необязательно одинаковой физической размерности.

Будем называть эту матрицу *характеристической*. Подчеркнем, что матрица \mathbf{M} получилась квадратной вследствие равенства числа n параметров в каждой группе числу n независимых линейных соотношений между параметрами.

В расписанном виде равенство (10.1.1) имеет вид:

$$\begin{aligned} b_1 &= M_{11}a_1 + M_{12}a_2 + \dots + M_{1n}a_n \\ b_2 &= M_{21}a_1 + M_{22}a_2 + \dots + M_{2n}a_n \\ &\dots\dots\dots \\ &\dots\dots\dots \\ b_n &= M_{n1}a_1 + M_{n2}a_2 + \dots + M_{nn}a_n \end{aligned}$$

Уравнения этой системы называются *основными уравнениями* теории многополюсников. Коэффициенты системы основных уравнений, т. е. элементы M_{ij} , $i, j = 1 \dots n$ матрицы \mathbf{M} составляют *систему вторичных параметров* многополюсника, описывающую его свойства в рамках концепции многополюсников.

В зависимости от разбиения первичных параметров на группы А и В, физический смысл (и размерности) систем вторичных параметров получаются разными, и в этом смысле системы вторичных параметров имеют разные представления. Всего, без учета перестановок элементов в группах, поскольку они не имеют принципиального значения, возможны

$$C_{2n}^2 = \frac{(2n)!}{2 \cdot (2n-2)!} = n(2n-1)$$

разбиения на группы А и В (C_k^s – число сочетаний из k по s , знак «!» обозначает факториал), столько же возможно разных по физическому смыслу систем вторичных параметров и соответствующих им характеристических матриц. Как видим, уже при $n = 2$ (четыреполюсник) число таких матриц равно 6 (см. табл. 10.1.1 [10.1]), и с ростом n быстро возрастает.

Система вторичных параметров многополюсника

Таблица 10.1.1

Четыреполюсник ($n = 2$). Варианты разбиения первичных параметров на группы А и В и матричного представления вторичных параметров

Варианты	1	2	3	4	5	6
Множ-во А	u_1, u_2	i_1, i_2	u_1, i_1	u_2, i_2	i_1, u_2	u_1, i_2
Множ-во В	i_1, i_2	u_1, u_2	u_2, i_2	u_1, i_1	u_1, i_2	i_1, u_2

Обозн-ние мат-цы	Y	Z	B	A	H	F
------------------	----------	----------	----------	----------	----------	----------

Однако на практике используются не все варианты характеристических матриц, а лишь те, что предоставляют некоторые удобства в тех или иных ситуациях. Понятно, что характеристические матрицы разных видов взаимно-однозначно связаны друг с другом.

Наиболее популярные классические системы параметров (z, y, a, h)



Комплексные амплитуды токов и напряжений в двухпортовой цепи

Мотивы выбора классических систем

Обратимся для простоты рассмотрения к случаю четырехполюсника (двухпортовой цепи), $n = 2$ (рис. 10.1.2). Классическая система первичных параметров в этом случае состоит из двух напряжений и двух токов, матрица \mathbf{M} – (2×2) , система вторичных параметров также включает 4 элемента, возможны 6 вариантов характеристической матрицы (табл. 10.1.1), наиболее часто в практике конструирования и измерения используются 4 из них:

Система z-параметров составляет матрицу $\mathbf{Z} = \{z_{ij}\}, i, j = 1, 2$, все параметры имеют размерность сопротивления (в общем случае – комплексного), система основных уравнений имеет вид:

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix}.$$

Система y-параметров составляет матрицу $\mathbf{Y} = \{y_{ij}\}, i, j = 1, 2$, все параметры имеют размерность проводимости (в общем случае – комплексной), система основных уравнений имеет вид:

$$\begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Система a-параметров составляет матрицу $\mathbf{A} = \{a_{ij}\}, i, j = 1, 2$, система гибридна, т. е. параметры имеют разную размерность: a_{11} и a_{22} безразмерны, a_{12} имеет размерность сопротивления, a_{21} – проводимости. Система основных уравнений имеет вид:

$$\begin{bmatrix} u_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} u_2 \\ -i_2 \end{bmatrix}.$$

Система h -параметров составляет матрицу $\mathbf{H} = \{h_{ij}\}$, $i, j = 1, 2$, система гибридна: h_{12} и h_{21} – безразмерны, h_{11} имеет размерность сопротивления, h_{22} – проводимости. Система основных уравнений имеет вид:

$$\begin{bmatrix} u_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} i_1 \\ u_2 \end{bmatrix}.$$

Выбор для использования той или иной характеристической матрицы мотивируется удобством решения поставленной задачи (например, конструирование усилителей, генераторов или смесителей, и т. п.), в частности, простоты определения матрицы «сложной» схемы, составленной из более простых схем. Например [10.2], y -параметры «сложного» четырехполюсника, полученного параллельно-параллельным соединением двух «простых» четырехполюсников, просто равны суммам соответствующих параметров «простых» четырехполюсников:

$$\begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} i_1^a + i_1^b \\ i_2^a + i_2^b \end{bmatrix} = \begin{bmatrix} y_{11}^a + y_{11}^b & y_{12}^a + y_{12}^b \\ y_{21}^a + y_{21}^b & y_{22}^a + y_{22}^b \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},$$

т. е. \mathbf{Y} -матрица «сложного» четырехполюсника равна сумме \mathbf{Y} -матриц «простых» четырехполюсников:

$$\mathbf{Y} = \mathbf{Y}_a + \mathbf{Y}_b.$$

Аналогичным образом, при последовательно-последовательном соединении четырехполюсников складываются z -параметры, при последовательно-параллельном – h -параметры. При каскадном соединении четырехполюсников \mathbf{A} -матрицы каскадов перемножаются, т. е.

$$\mathbf{A} = \mathbf{A}_a \mathbf{A}_b,$$

этот факт часто является весомым аргументом применения a -параметров на определенном этапе конструирования.

Вследствие взаимно-однозначной связи одной характеристической матрицы с другой, в каждой ситуации достаточно измерить только элементы одной матрицы. Например, z -параметры и y -параметры связаны матричным соотношением

$$\mathbf{Y} = \mathbf{Z}^{-1}.$$

Смысл и измерение классических вторичных параметров многополюсников

Соотношения комплексных амплитуд в классических системах

Обратимся к важному вопросу о принципиальной возможности измерения вторичных параметров цепи. Для упрощения рассуждений, не ограничивая общности ответов, в качестве многополюсников рассматриваем четырехполюсники, а первичные параметры (токи и напряжения) считаем монохроматическими, в комплексном представлении. Сократив члены вышеприведенных основных уравнений на множитель $\exp(i\omega t)$, получаем выражения через комплексные амплитуды токов и напряжений:

- для системы z -параметров:

$$\begin{aligned}\dot{U}_1 &= z_{11}\dot{I}_1 + z_{12}\dot{I}_2 \\ \dot{U}_2 &= z_{21}\dot{I}_1 + z_{22}\dot{I}_2\end{aligned}\quad (10.2.1)$$

- для системы y -параметров:

$$\begin{aligned}\dot{I}_1 &= y_{11}\dot{U}_1 + y_{12}\dot{U}_2 \\ \dot{I}_2 &= y_{21}\dot{U}_1 + y_{22}\dot{U}_2\end{aligned}\quad (10.2.2)$$

- для системы a -параметров:

$$\begin{aligned}\dot{U}_1 &= a_{11}\dot{U}_2 - a_{12}\dot{I}_2 \\ \dot{I}_1 &= a_{21}\dot{U}_2 - a_{22}\dot{I}_2\end{aligned}\quad (10.2.3)$$

- для системы h -параметров:

$$\begin{aligned}\dot{U}_1 &= h_{11}\dot{I}_1 + h_{12}\dot{U}_2 \\ \dot{I}_2 &= h_{21}\dot{I}_1 + h_{22}\dot{U}_2\end{aligned}\quad (10.2.4)$$

Необходимость измерений в режимах ХХ и КЗ

Обратимся к z -параметрам (см. систему (10.2.1)). Пусть порт 2 работает в режиме холостого хода (ХХ), тогда $\dot{I}_2 = 0$ и из первого уравнения системы (10.2.1) следует:

$$z_{11} = \left(\frac{\dot{U}_1}{\dot{I}_1} \right)_{\text{ХХ на порте 2}},$$

а из второго уравнения той же системы:

$$z_{21} = \left(\frac{\dot{U}_2}{\dot{I}_1} \right)_{\text{ХХ на порте 2}}.$$

Пусть теперь порт 1 работает в режиме холостого хода (ХХ), тогда $\dot{I}_1 = 0$ и из первого уравнения системы (10.2.1) следует:

$$z_{12} = \left(\frac{\dot{U}_1}{\dot{I}_2} \right)_{\text{ХХ на порте 1}},$$

а из второго уравнения той же системы:

$$z_{22} = \left(\frac{\dot{U}_2}{\dot{I}_2} \right)_{\text{ХХ на порте 1}}.$$

Полученные формулы для z -параметров можно рассматривать как выражения их физического смысла, и одновременно – как способ их измерения. Как видим, для измерения z -параметров необходимо последовательно реализовать режимы ХХ на входе и на выходе четырехполюсника.

Обратимся к y -параметрам (см. систему (10.2.2)). Пусть порт 2 работает в режиме короткого замыкания (КЗ), тогда $\dot{U}_2 = 0$ и из первого уравнения системы (10.2.2) следует:

$$y_{11} = \left(\frac{\dot{I}_1}{\dot{U}_1} \right)_{\text{КЗ на порте 2}},$$

а из второго уравнения той же системы:

$$y_{21} = \left(\frac{\dot{I}_2}{\dot{U}_1} \right)_{\text{КЗ на порте 2}}.$$

Пусть теперь порт 1 работает в режиме КЗ, тогда $\dot{U}_1 = 0$ и из первого уравнения системы (10.2.2) следует:

$$y_{12} = \left(\frac{\dot{I}_1}{\dot{U}_2} \right)_{\text{КЗ на порте 1}},$$

а из второго уравнения той же системы:

$$y_{22} = \left(\frac{\dot{I}_2}{\dot{U}_2} \right)_{\text{КЗ на порте 1}}.$$

Полученные формулы для y -параметров можно рассматривать как выражения их физического смысла, и одновременно – как способ их измерения. Как видим, для измерения y -параметров необходимо последовательно реализовать режимы КЗ на входе и на выходе четырехполюсника.

Аналогичным образом находим для системы a -параметров:

$$a_{11} = \left(\frac{\dot{U}_1}{\dot{U}_2} \right)_{\text{XX на порте 2}}, \quad a_{21} = \left(\frac{\dot{I}_1}{\dot{U}_2} \right)_{\text{XX на порте 2}}, \quad a_{12} = - \left(\frac{\dot{U}_1}{\dot{I}_2} \right)_{\text{КЗ на порте 1}}, \quad a_{22} = - \left(\frac{\dot{I}_1}{\dot{I}_2} \right)_{\text{КЗ на порте 1}},$$

и для системы h -параметров:

$$h_{11} = \left(\frac{\dot{U}_1}{\dot{I}_1} \right)_{\text{КЗ на порте 2}}, \quad h_{21} = \left(\frac{\dot{I}_2}{\dot{I}_1} \right)_{\text{КЗ на порте 2}}, \quad h_{12} = \left(\frac{\dot{U}_1}{\dot{U}_2} \right)_{\text{XX на порте 1}}, \quad h_{22} = \left(\frac{\dot{I}_2}{\dot{U}_2} \right)_{\text{XX на порте 1}}.$$

Подведем итог, справедливый для всех классических систем вторичных параметров четырехполюсников, находящихся под воздействием монохроматических колебаний: вторичный параметр находится как отношение комплексных амплитуд токов и (или) напряжений в режиме ХХ или КЗ одного из портов. Этот вывод распространяется и на вторичные параметры $2n$ -полюсников при $n > 1$: вторичный параметр находится как отношение комплексных амплитуд токов и (или) напряжений в режиме ХХ и (или) КЗ на $n - 1$ портах.

Мотивы введения параметров рассеяния

Трудности измерения классических параметров на СВЧ

На сверхвысоких частотах классические вторичные параметры очень трудно измерять по следующим причинам:

А) Для СВЧ-сигналов очень трудно реализовать режимы короткого замыкания и холостого хода из-за паразитных емкостей и индуктивностей.

Б) Измерения токов и напряжений в схеме на СВЧ требуют подстроечных отводов, отдельно отрегулированных на каждой частоте (что исключает измерения в режиме автоматического сканирования частоты), чтобы отобразить необходимые режимы.

В) В активных приборах, таких как транзистор или диод с отрицательным сопротивлением, режимы ХХ и КЗ часто вызывают неустойчивую работу и даже осцилляции.

Принципиальные отличия параметров рассеяния от классических систем

Для решения этих проблем, на сверхвысоких частотах вместо классических систем вторичных параметров используются *параметры рассеяния* (*s-параметры*), которые связаны не с режимами холостого хода и короткого замыкания, а с *падающей* и *отраженной* мощностями. Преимущество использования параметров рассеяния в том, что поскольку они связаны с падающей и отраженной мощностями, они не меняют своего значения вдоль передающей линии без потерь. Это означает, что параметры рассеяния могут быть измерены на приборе, находящемся на некотором расстоянии от точки измерения. К тому же, параметры рассеяния измеряются в условиях импедансного согласования, поэтому обходят осцилляции неустойчивости в измерениях на активных приборах.

Сначала заметим, что параметры рассеяния как система вторичных параметров многополюсника принципиально отличаются от классических систем вторичных параметров тем, что для первых первичными параметрами являются уже не токи и напряжения на выходных контактах, как для вторых, а *падающие (входящие)* и *отраженные (выходящие)* волны на этих контактах (разъемах).

Параметры рассеяния первоначально пришли из теории линий передачи, и наиболее простой и прозрачный способ их введения основан на представлении цепи СВЧ как *сочленения* линий передачи. Такое представление цепи удобно при рассмотрении многих вопросов векторного анализа цепей. Однако на самом деле параметры рассеяния имеют более общий смысл некоторых параметров мощности, более конкретно – коэффициентов линейных соотношений между *волнами* мощности, концепцию которых и связанное с ними обобщенное представление параметров рассеяния развил Курокава [10.3].

Параметры рассеяния (коэффициенты рассеяния) были вначале упомянуты Кэмбеллом и Фостером [10.4] в 1920 г., применительно еще к телефонным цепям; затем уже были использованы в задачах микроволновой техники и микроволновых линий передачи [10.5, 10.6], а также в общей теории цепей [10.7]. Они естественно возникали в физической интерпретации решений стандартных дифференциальных уравнений линии передачи для напряжения и тока как функций дистанции.

Вначале введем эти параметры наиболее простым способом, затем несколько обобщим трактовку.

Смысл и измерение параметров рассеяния

Нормализация токов и напряжений

Многополюсник почти всегда можно рассматривать как *сочленение линий передачи* (вопрос, собственно, в том, можно ли рассматривать связи многополюсника с остальной частью устройства как длинные линии, например, достаточно ли они длинные для затухания «лишних» мод, и т. п.). Для данного типа монохроматической волны распределение поперечных составляющих напряженности электрического поля в любом сечении однородной линии передачи одинаково (с точностью до амплитуды) [10.8]. Мощность P , переносимая падающей волной в однородной линии без потерь равна

$$P = \frac{1}{2} \int_S [\dot{\mathbf{E}}_{\perp}, \dot{\mathbf{H}}_{\perp}] \mathbf{e}_z dS = \frac{1}{2Z_c} \int_S |\dot{\mathbf{E}}_{\perp}|^2 dS = \frac{Z_c}{2} \int_S |\dot{\mathbf{H}}_{\perp}|^2 dS,$$

где $\dot{\mathbf{E}}_{\perp}, \dot{\mathbf{H}}_{\perp}$ – поперечные (к оси OZ линии передачи) составляющие комплексных амплитуд векторов электрического и магнитного полей;

Z_c – характеристическое сопротивление линии передачи для данного типа волны;

$[\mathbf{A}, \mathbf{B}]$ – векторное произведение векторов \mathbf{A} и \mathbf{B} ;

\mathbf{e}_z – единичный вектор в направлении оси OZ ;

интегрирование ведется по поперечному сечению S линии.

Нормированные электрическая и магнитная напряженности $\dot{\mathbf{E}}_{\perp}^n, \dot{\mathbf{H}}_{\perp}^n$ вводятся так, чтобы

$$\int_S |\dot{\mathbf{E}}_{\perp}^n|^2 dS = \int_S |\dot{\mathbf{H}}_{\perp}^n|^2 dS = 1,$$

тогда, в силу одинаковости распределения $\dot{\mathbf{E}}_{\perp}$ по сечению S при разных z , можно записать

$$\dot{\mathbf{E}}_{\perp} = a_U \dot{U}_{\text{эк}} \dot{\mathbf{E}}_{\perp}^n, \quad \dot{\mathbf{H}}_{\perp} = a_I \dot{I}_{\text{эк}} \dot{\mathbf{H}}_{\perp}^n,$$

где a_U, a_I – размерные коэффициенты нормировки;

$\dot{U}_{\text{эк}}, \dot{I}_{\text{эк}}$ – эквивалентные напряжение и ток в линии передачи, пропорциональные, как видим, амплитудам напряженностей электрического и магнитного полей, соответственно, и называемые *интегральными параметрами* линии передачи.

Введем также эквивалентные *волновые сопротивления* по напряжению и току:

$$Z_{\text{вУ}} = Z_c / |a_U|^2, \quad Z_{\text{вИ}} = Z_c |a_I|^2.$$

Пары коэффициентов нормировки a_U, a_I можно выбрать по-разному, задавая тем самым различные нормировки полей и эквивалентных напряжения и тока. Из соображений удобства записи формул наиболее часто выбирают

$$|a_U| = \sqrt{Z_c}, \quad |a_I| = 1/\sqrt{Z_c}$$

(предполагается, что характеристическое сопротивление линии передачи действительно). В получаемой при таком выборе коэффициентов нормировки эквивалентные напряжение и ток называются *нормализованными* и обозначаются $\dot{U}_{\text{эк}}^n, \dot{I}_{\text{эк}}^n$, а волновые сопротивления $Z_{\text{вУ}} = Z_{\text{вИ}} = 1$.

Из вышеприведенных соотношений легко получить формулу для переносимой мощности:

$$P = \frac{1}{2} |\dot{U}_{\text{эк}}^n|^2 = \frac{1}{2} |\dot{I}_{\text{эк}}^n|^2 = \frac{1}{2} \dot{U}_{\text{эк}}^n \dot{I}_{\text{эк}}^{n*} \quad (10.4.1)$$

откуда следует, что нормализованные напряжение и ток имеют одинаковую размерность $\sqrt{\text{Вт}}$. Все это дает право рассматривать нормализованную волну с комплексными амплитудами напряжения и тока $\dot{U}_{\text{эк}}^n, \dot{I}_{\text{эк}}^n$ как *волну мощности* [10.3].

Введение параметров рассеяния (илл. Осн анал цепей, с. 24, с. 28, с. 29, 30(таб))

Рассмотрим $2n$ -полнолюстик (n -портовое устройство, n -плечное устройство) как сочленение n линий передачи. Внутренняя структура многополнолюстика может быть как угодно сложной, при возбуждении его со стороны одной или нескольких линий внутри сочленения возникает как угодно сложное электромагнитное поле, но нас интересуют лишь соотношения между первичными и вторичными параметрами многополнолюстика. В общем случае подводящие к сочленению линии передачи могут быть разного типа (волноводного, коаксиального, полоскового), разной конструкции, разных характерных размеров, иметь разные постоянные распространения, в них могут распространяться различные наборы мод и т. д. Для большей ясности изложения будем считать, что подводящие линии передачи однородны, без потерь, одной конструкции, с одномодовым переносом волны и одинаковым коэффициентом фазы β на данной частоте. От некоторых из этих ограничений можно отказаться, например, при переносе по линии нескольких независимых мод следует заменить эту линию на соответствующее число линий со своими постоянными распространения. Однако нам подобные обобщения не понадобятся.

Выберем и зафиксируем на каждой подводящей линии передачи некую плоскость поперечного сечения (*плоскость отсчета фаз, опорная плоскость*), находящуюся на своем электрическом расстоянии $z_m = \beta l_m$, $m = 1, 2, \dots, n$ от одной условной точки (*центр*) сочленения. Величины l_m должны быть достаточно большими, чтобы нежелательные моды волн на их протяжении сильно затухали. Выбранные n плоскостей отсчета фаз являются одновременно входами и выходами (*портами*) $2n$ -полнолюстика. Условимся называть *падающими (входящими)* волны, распространяющиеся в сторону сочленения, и *отраженными (выходящими)* – распространяющиеся от сочленения. В общем случае на каждом из n портов есть одна падающая и одна отраженная волны (в частном случае некоторые падающие волны могут отсутствовать).

В соответствии с ранее принятыми обозначениями, комплексные амплитуды нормализованных падающих и отраженных волн напряжений и токов пришлось бы теперь записывать так:

$$\dot{U}_{m \text{ пад}}^n(l_m), \quad \dot{I}_{m \text{ пад}}^n(l_m) \text{ и т. д.}$$

где учтено, что они зависят от параметров l_m , $m = 1, 2, \dots, n$. Для упрощения записи введем переобозначения:

$$\dot{i}_m^+ = \dot{U}_{m \text{ пад}}^n(l_m), \quad \dot{i}_m^- = \dot{I}_{m \text{ пад}}^n(l_m), \quad \dot{u}_m^- = \dot{U}_{m \text{ отр}}^n(l_m), \quad \dot{u}_m^+ = \dot{I}_{m \text{ отр}}^n(l_m).$$

В новых обозначениях зависимости от l_m не показаны явно, однако следует помнить, что комплексные амплитуды нормализованных волн относятся к выбранным опорным плоскостям, которые далее называются *портами*. Согласно (10.4.1), мощности падающей и отраженной волн в плече m :

$$P_m^+ = \frac{1}{2} |\dot{i}_m^+|^2 = \frac{1}{2} |\dot{i}_m^+|^2, \quad P_m^- = \frac{1}{2} |\dot{u}_m^-|^2 = \frac{1}{2} |\dot{i}_m^-|^2,$$

поэтому могут называться *волнами мощности*.

Как и в изложенных выше классических матричных описаниях n -портовой линейной цепи, мы нуждаемся в $2n$ параметрах возбужденной цепи, которые можно разделить на две группы по n параметров в каждой, и установить n независимых линейных соотношений между элементами этих групп. Возьмем в качестве первичных параметров возбужденной цепи падающие волны, а в качестве вторичных – отраженные волны (в этом принципиальное отличие нового описания от классических систем, где в качестве пара-

метров фигурировали токи и напряжения). Отметим принципиальный момент: поскольку полное напряжение в каждом порте равно (векторной) сумме падающей и отраженной волн напряжения, то совокупность всех волн напряжения возбужденного многополюсника представляет собой линейную систему; аналогично обстоит дело и с волнами тока.

Образует n -мерные *нормализованные векторы-столбцы* падающих и отраженных волн напряжения:

$$\dot{\mathbf{U}}^+ = \{\dot{u}_k^+\}, \quad k = 1, 2, \dots, n; \quad \dot{\mathbf{U}}^- = \{\dot{u}_k^-\}, \quad k = 1, 2, \dots, n,$$

тогда линейное соотношение между этими векторами (основная система равенств $2n$ -полюсника) запишется в виде:

$$\dot{\mathbf{U}}^- = \mathbf{S}\dot{\mathbf{U}}^+, \tag{10.5.1}$$

или в расписанном виде

$$\begin{aligned} \dot{u}_1^- &= S_{11}\dot{u}_1^+ + S_{12}\dot{u}_2^+ + \dots + S_{1n}\dot{u}_n^+ \\ \dot{u}_2^- &= S_{21}\dot{u}_1^+ + S_{22}\dot{u}_2^+ + \dots + S_{2n}\dot{u}_n^+ \\ &\dots\dots\dots \\ &\dots\dots\dots \\ \dot{u}_n^- &= S_{n1}\dot{u}_1^+ + S_{n2}\dot{u}_2^+ + \dots + S_{nn}\dot{u}_n^+ \end{aligned} \tag{10.5.2}$$

В равенствах (10.5.1), (10.5.2): $\mathbf{S} = \{S_{ij}\}$, $i, j = 1, 2, \dots, n$ – квадратная комплекснозначная матрица ($n \times n$), называемая *матрицей рассеяния* (*S-матрицей, волновой матрицей*); комплексные безразмерные величины S_{ij} , $i, j = 1, 2, \dots, n$ – элементы матрицы рассеяния, называемые *параметрами рассеяния* (*s-параметрами*). Заметим, что, поскольку нормализованные волны относятся к опорным плоскостям, то и связывающие их параметры рассеяния также относятся к этим плоскостям, и при изменении положения этих плоскостей параметры рассеяния, вообще говоря, меняются.

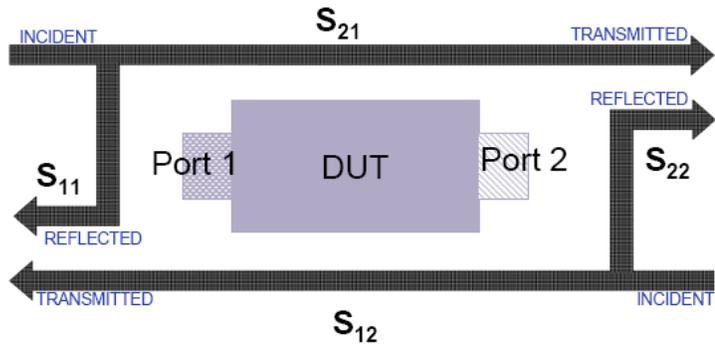
Можно, наоборот, выразить вектор $\dot{\mathbf{U}}^+$ падающих волн через вектор $\dot{\mathbf{U}}^-$ отраженных волн:

$$\dot{\mathbf{U}}^+ = \mathbf{S}^{-1}\dot{\mathbf{U}}^- \tag{10.5.3}$$

при условии, что матрица рассеяния \mathbf{S} не особенная, т. е. что $\det \mathbf{S} \neq 0$.

Так же как каждая классическая матрица линейного многополюсника, матрица рассеяния полностью описывает свойства последнего на данной частоте возбуждения. Рассмотрев произвольную m -ую строку системы, видим, что нормализованная отраженная волна на m -ом порте равна взвешенной векторной сумме падающих волн на всех портах, а элементы m -ой строки матрицы рассеяния \mathbf{S} суть комплексные веса в этой сумме.

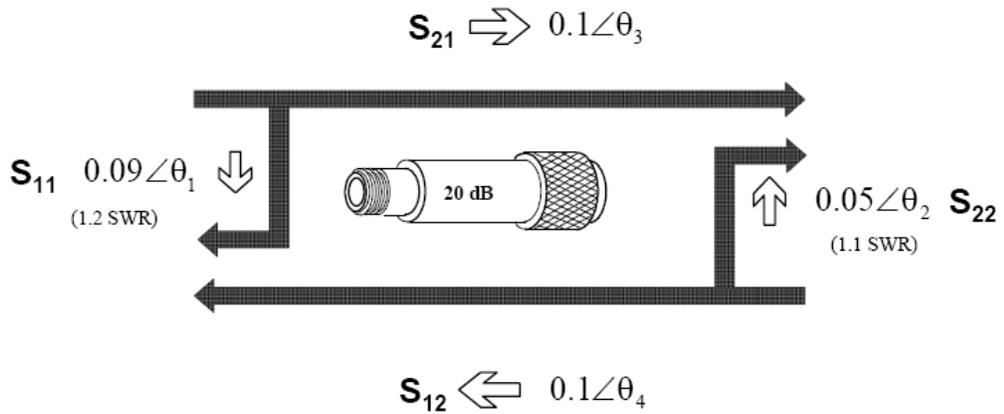
S Parameters



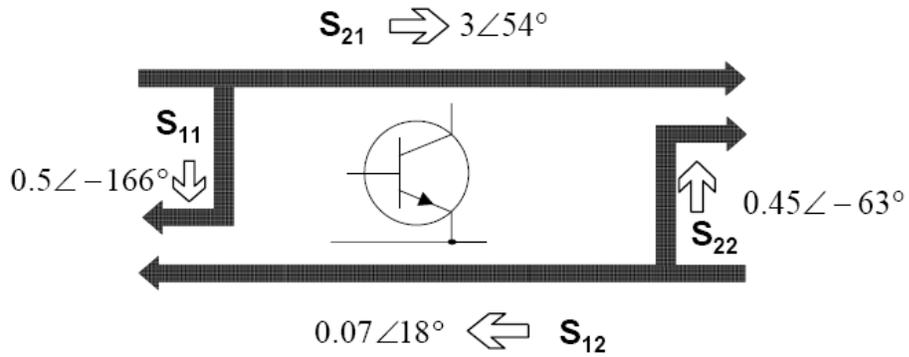
$$b_1 = S_{11}a_1 + S_{12}a_2$$

$$b_2 = S_{21}a_1 + S_{22}a_2$$

Attenuator S-Parameters



Transistor S-parameters



2N5103 @ 1 GHz

Transistor Data

AT- 41486 (hp)



!AT-41486!

S AND NOISE PARAMETERS at Vce=8V Ic=10mA. LAST UPDATED 07-21-92

ghz s ma r 50

F GHz	S11		S12		S21		S22	
	Mag.	Angle	Mag	Angle	Mag	Angle	Mag	Angle
0.1	.74	-38	25.46	157	.011	68	.94	-12
0.5	.59	-127	12.63	107	.031	47	.60	-29
1.0	.56	-168	6.92	84	.041	46	.49	-29
2.0	.62	152	3.61	56	.058	43	.42	-39
3.0	.64	130	2.41	37	.078	52	.39	-50
4.0	.71	113	1.80	16	.106	48	.35	-70
5.0	.77	99	1.42	-4	.139	43	.35	-98
6.0	.81	87	1.13	-22	.170	34	.35	-131

Смысл параметров рассеяния

Пусть $2n$ -полнолюсок возбуждается только из m -го порта ($m=1, 2, \dots, n$) падающей волной \dot{u}_m^+ , а остальные порты нагружены на согласованные нагрузки. Так как коэффициент отражения от согласованной нагрузки равен нулю, то в остальные плечи многополосника энергия не поступает, т. е. все падающие волны, кроме m -ой, равны нулю. Тогда от системы (10.5.2) остается только «столбцовая» система

$$\dot{u}_k^- = S_{km} \dot{u}_m^+, \quad k = 1, 2, \dots, n \quad (10.6.1)$$

показывающая, что под действием такого возбуждения в сочленении возникает электромагнитное поле, в свою очередь возбуждающее выходящие из сочленения (отраженные) волны, все пропорциональные одной и той же падающей волне \dot{u}_m^+ , но с разными (в общем случае – комплексными) коэффициентами пропорциональности S_{km} , $k = 1, 2, \dots, n$. С математической точки зрения коэффициенты S_{km} суть элементы m -го столбца матрицы рассеяния \mathbf{S} , с физической – они зависят от внутренней структуры многополосника и положения плоскостей отсчета.

Поскольку аналогичные соотношения можно записать при возбуждении многополосника со стороны других плеч (выбор любого m в пределах от 1 до n), при согласовании остальных плеч, соотношение (10.6.1) дает возможность одновременно выяснить физический смысл параметров рассеяния и найти принципиальный метод их измерения. При $m = k$ из (10.6.1) получаем:

$$S_{kk} = \dot{u}_k^- / \dot{u}_k^+ = \Gamma_k, \quad k = 1, 2, \dots, n, \quad (10.6.2)$$

т. е. диагональные элементы матрицы рассеяния суть (комплексные) коэффициенты отражения в плоскостях отсчета соответственных плеч, и принципиальный способ измерения диагональных элементов – измерения соответствующих коэффициентов отражения в условиях согласования всех плеч, кроме измеряемого.

При $m \neq k$ из (10.6.1) получаем:

$$S_{km} = \dot{u}_k^- / \dot{u}_m^+, \quad k, m = 1, 2, \dots, n; \quad m \neq k, \quad (10.6.3)$$

т. е. недиагональный элемент S_{km} матрицы рассеяния суть (комплексный) коэффициент передачи из плеча m в плечо k , при условии, что к плечу m подключен генератор, а все остальные плечи согласованы; принципиальный способ измерения этого элемента – измерение названного коэффициента передачи при условии всех названных согласований. Во избежание путаницы подчеркнем еще раз: первый индекс недиагонального матричного элемента – номер плеча, в который направлен коэффициент передачи, второй индекс – номер возбуждаемого плеча.

Измерение параметров рассеяния

Рассмотрим частный случай четырехполюсника (двухпортовой цепи), $n=2$ (рис. 10.6.1). Основная система равенств в терминах параметров рассеяния имеет вид:

$$\dot{u}_1^- = S_{11}\dot{u}_1^+ + S_{12}\dot{u}_2^+, \quad (10.6.4)$$

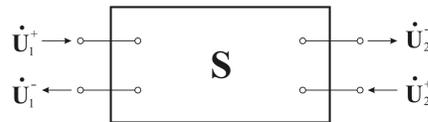
$$\dot{u}_2^- = S_{21}\dot{u}_1^+ + S_{22}\dot{u}_2^+, \quad (10.6.5)$$

S_{11} – (комплексный) коэффициент отражения на входе (в порте 1) при согласовании на выходе (в порте 2);

S_{22} – (комплексный) коэффициент отражения на выходе (в порте 2) при согласовании на входе (в порте 1);

S_{21} – (комплексный) коэффициент передачи из порта 1 в порт 2 при согласовании на выходе (в порте 2);

S_{12} – (комплексный) коэффициент передачи из порта 2 в порт 1 при согласовании на входе (в порте 1).



Нормализованные волны четырехполюсника

Каскадно-специфическая матрица рассеяния

Не выходя из рамок концепции падающих и отраженных волн, т. е. оставляя последние в качестве первичных и вторичных параметров, можно построить несколько иное матричное описание [10.2], более удобное в случае каскадного соединения четырехполюсников. Возьмем в качестве первичных параметров нормализованные волны \dot{u}_1^+, \dot{i}_1^- на входе, а в качестве вторичных – нормализованные волны \dot{u}_2^+, \dot{i}_2^- на выходе. Образовав, в соответствии с рис. 10.6.1, векторы-столбцы

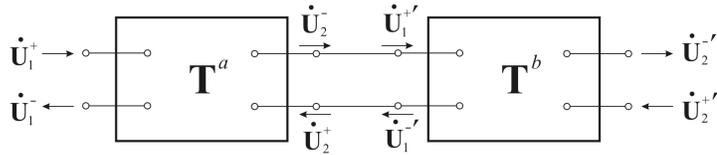
$$\begin{bmatrix} \dot{u}_1^+ \\ \dot{u}_1^- \end{bmatrix}, \quad \begin{bmatrix} \dot{u}_2^- \\ \dot{u}_2^+ \end{bmatrix},$$

(обратим внимание на неодинаковый порядок компонент в векторах), свяжем их линейно:

$$\begin{bmatrix} \dot{u}_1^+ \\ \dot{u}_1^- \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} \dot{u}_2^- \\ \dot{u}_2^+ \end{bmatrix},$$

с помощью матрицы $\mathbf{T} = \{T_{ij}\}$, $i, j = 1, 2$, которую назовем *каскадно-специфической*. То же соотношение в распisanном виде:

$$\begin{aligned} \dot{u}_1^+ &= T_{11}\dot{u}_2^- + T_{12}\dot{u}_2^+ \\ \dot{u}_1^- &= T_{21}\dot{u}_2^- + T_{22}\dot{u}_2^+ \end{aligned}$$



Каскадное соединение двухпортовых цепей

Из рассмотрения рис. 10.8.1 ясно, что T -матрица каскадного соединения двух четырехполюсников равна произведению их T -матриц, т. е.

$$\begin{bmatrix} \dot{u}_1^+ \\ \dot{u}_1^- \end{bmatrix} = \begin{bmatrix} T_{11}^a & T_{12}^a \\ T_{21}^a & T_{22}^a \end{bmatrix} \begin{bmatrix} T_{11}^b & T_{12}^b \\ T_{21}^b & T_{22}^b \end{bmatrix} \begin{bmatrix} \dot{u}_2^- \\ \dot{u}_2^+ \end{bmatrix}.$$

Это свойство справедливо для любого числа каскадов.

Матрицы \mathbf{S} и \mathbf{T} связаны взаимно-однозначно:

$$\text{при } T_{11} \neq 0 \quad S_{11} = T_{21}/T_{11}; \quad S_{21} = 1/T_{11}; \quad S_{12} = T_{22} - T_{21}T_{12}/T_{11}; \quad S_{22} = -T_{12}/T_{11};$$

$$\text{при } S_{21} \neq 0 \quad T_{11} = 1/S_{21}; \quad T_{12} = -S_{22}/S_{21}; \quad T_{21} = S_{11}/S_{21}; \quad T_{22} = S_{12} - S_{11}S_{22}/S_{21}$$

Преимущества параметров рассеяния на СВЧ

Описание линейного многополюсника с помощью параметров рассеяния свободно от недостатков, присущих классическим системам параметров на СВЧ

А) В технике СВЧ основные непосредственно измеряемые величины – частота, коэффициент отражения и мощность (источником рассматриваются как генераторы с постоянной мощностью). Поэтому в принципе элементы матрицы рассеяния можно измерять

непосредственно, в то время как элементы матриц сопротивлений и проводимостей являются абстрактными величинами.

Б) Измерение параметров рассеяния n -портовой цепи требует не организации трудно осуществимых на СВЧ режимов ХХ и КЗ, а согласования $n-1$ порта.

В) Поскольку мощности падающей и отраженной волн не меняются при распространении по однородной линии передачи без потерь (меняются, причем простым образом, только фазы волн), нет необходимости проводить их измерения или подключать генераторы в точках трудного подключения на СВЧ, например, на выводах микроволновых приборов, а возможно это делать в любых удобных сечениях линий, уже подключенных, или специально добавленных; необходимо лишь учитывать фазовую разность между опорным сечением и сечением измерения или возбуждения

Г) В отличие от режимов ХХ и КЗ, согласования портов активных приборов не представляет опасности их неустойчивой работы и осцилляций.

Д) Если конкретный многополюсник обладает геометрической симметрией, теоретическое определение его матрицы рассеяния намного проще, чем определение матриц сопротивлений или проводимостей.

Е) Измерение именно параметров рассеяния не означает недоступность на СВЧ параметров классических систем, часто полезных для конструирования и оптимизации устройств: система параметров рассеяния связана взаимно-однозначно с каждой из классических систем. Поэтому, измерив параметры рассеяния устройства, можно затем рассчитать параметры любой другой системы.

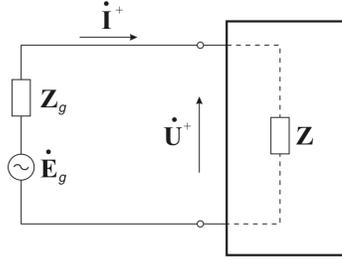
Ж) В отличие от классических матричных описаний, которые не существуют для некоторых устройств СВЧ (например, матрица \mathbf{Z} не определена для короткозамкнутого четвертьволнового и холостогодного полуволнового отрезков линии передачи, матрица \mathbf{Y} – для короткозамкнутого полуволнового и холостогодного четвертьволнового отрезков), матрица рассеяния \mathbf{S} определена для всех пассивных линейных устройств СВЧ.

З) Изменение опорных плоскостей (плоскостей отсчета) изменяет матрицу рассеяния \mathbf{S} гораздо более простым образом, чем классические матрицы.

И) Соотношения между элементами матрицы рассеяния, являющиеся следствиями симметрии описываемого ею устройства, устанавливаются гораздо проще, чем для классических характеристических матриц того же устройства.

Обобщенная трактовка параметров рассеяния

Выше параметры рассеяния цепи СВЧ введены на основе ее трактовки как сочленения линий передачи, имеющих одинаковое действительное характеристическое сопротивление. Теперь изложим более общую трактовку [10.2, 10.3], опирающуюся на концепцию падающих и отраженных нормализованных волн тока и напряжения, не нуждающуюся в подключенных к портам цепи линиях передачи, справедливую и при комплексных волновых сопротивлениях, и проясняющую роль генераторов, возбуждающих описываемую цепь и имеющих в общем случае комплексный внутренний импеданс. Для простоты изложения сначала рассмотрим частный случай двухполюсника (однопортовой цепи), затем частный случай четырехполюсника (двухпортовой цепи) и, наконец, общий случай n -портовой цепи.



Генератор и двухполюсник

На рис. 10.9.1 показан двухполюсник (однопортовая цепь) с комплексным входным импедансом Z , на который подается монохроматический сигнал от генератора с комплексной амплитудой эквивалентного напряжения \dot{E}_g и внутренним импедансом Z_g . Приняты обычные условные направления комплексных амплитуд входного тока \dot{I}^+ и входного напряжения \dot{U}^+ . Если импеданс нагрузки равен сопряженному генераторному импедансу:

$$Z = Z_g^*,$$

то однопортовая цепь называется *сопряжено-согласованной с генератором*. При этом условии отражения от нагрузки нет, ток на выводе есть *падающий ток*:

$$\dot{I}^+ = \frac{\dot{E}_g}{Z_g^* + Z_g} = \frac{\dot{E}_g}{2\operatorname{Re}(Z_g)},$$

напряжение между выводами – *падающее напряжение*:

$$\dot{U}^+ = \frac{Z_g^* \dot{E}_g}{2\operatorname{Re}(Z_g)}$$

а соотношение между ними:

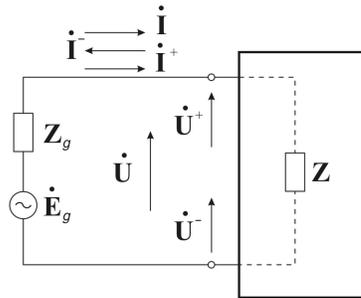
$$\dot{U}^+ = Z_g^* \dot{I}^+.$$

Сопряженно-согласованная с генератором нагрузка принимает максимально возможную мощность от генератора, равную

$$P_{\max} = \frac{1}{2} \operatorname{Re}(\dot{U} \dot{I}^*) = \frac{|\dot{E}_g|^2}{8\operatorname{Re}(Z_g)}.$$

Важно заметить, что падающие напряжение \dot{U}^+ и ток \dot{I}^+ не зависят от импеданса однопортовой цепи, а выражаются через генераторный импеданс, который поэтому называется *опорным импедансом цепи*. Это означает, что для данного генератора комплексные амплитуды падающих напряжения и тока постоянны, вне зависимости от подключенной нагрузки, и могут использоваться в качестве *измерительной опоры* в практических измерениях.

Формально-математически при $\operatorname{Re} Z_g = 0$ падающие напряжение и ток не могут быть однозначно определены, однако этот режим физически не реализуем.



Падающие и отраженные напряжения и токи

В общем случае, когда нагрузка не обязательно комплексно-сопряжена с генератором, кроме падающих напряжения и тока на выводах могут быть и *отраженные напряжение и ток*, которые определяются как

$$\dot{U}^- = \dot{U} - \dot{U}^+, \quad -\dot{I}^- = \dot{I} - \dot{I}^+,$$

(см. рис. 10.9.2), \dot{U} называется *полным* напряжением, а \dot{I} – *полным* током, знак минус перед \dot{I}^- означает, что направление тока \dot{I}^- выбрано противоположным направлению тока \dot{I}^+ .

Комплексный коэффициент отражения напряжения вводится как

$$\Gamma_u = \frac{\dot{U}^-}{\dot{U}^+} = \frac{Z_g(Z - Z_g^*)}{Z_g^*(Z + Z_g)},$$

а комплексный коэффициент отражения тока – как

$$\Gamma_i = \frac{\dot{I}^-}{\dot{I}^+} = \frac{Z - Z_g^*}{Z + Z_g},$$

соотношение между ними:

$$\frac{\Gamma_u}{\Gamma_i} = \frac{Z_g}{Z_g^*} = e^{2i \cdot \arg Z_g},$$

т. е. в общем случае эти коэффициенты не равны, а в случае действительного генераторного импеданса они одинаковы.

Только при сопряженно-генераторном согласовании нагрузки, т. е. при $Z = Z_g^*$, оба коэффициента равны нулю, т. е. отраженное напряжение и ток отсутствуют, полное напряжение (ток) равно падающему напряжению (току), падающая мощность полностью поглощается нагрузкой.

Комплексная амплитуда *падающей* нормализованной волны напряжения вводится как

$$\dot{i}^+ = \frac{\dot{U}^+ \sqrt{\operatorname{Re}(Z_g)}}{Z_g^*} = \frac{\dot{E}_g}{2 \operatorname{Re}(Z_g)},$$

а комплексная амплитуда *отраженной* нормализованной волны напряжения – как

$$\dot{i}^- = \frac{\dot{U}^- \sqrt{\operatorname{Re}(Z_g)}}{Z_g}.$$

Если ввести комплексные амплитуды *падающей* и *отраженной* нормализованных волн тока:

$$i^+ = I^+ \sqrt{\operatorname{Re}(Z_g)}, \quad i^- = I^- \sqrt{\operatorname{Re}(Z_g)},$$

то окажется, что $i^+ = \dot{i}^+$, $i^- = \dot{i}^-$, т. е. нормализованные волны тока и нормализованные волны напряжения суть одно и то же, поэтому слово «напряжение» («ток») можно отбросить и говорить о *падающей* и *отраженной* нормализованных волнах.

Смысл нормализованных волн выяснится, если найти, что

$$\frac{1}{2} |\dot{i}^+|^2 = \frac{|\dot{U}^+|^2 \operatorname{Re}(Z_g)}{2 |Z_g^*|^2} = P^+, \quad \frac{1}{2} |\dot{i}^-|^2 = \frac{|\dot{U}^-|^2 \operatorname{Re}(Z_g)}{2 |Z_g|^2} = P^-,$$

где P^+ , P^- – падающая и отраженная мощности, т. е. нормализованные волны это *волны мощности* [10.3], а размерность нормализованных волн – $\sqrt{\text{Вт}}$. Заметим, что падающая нормализованная волна \dot{i}^+ связана только с генераторным импедансом, в то время как отраженная нормализованная волна связана и с генераторным импедансом, и с импедансом цепи (через \dot{U}^-).

Полезно выразить нормализованные волны через полное напряжение \dot{U} и ток I . Поскольку

$$\begin{aligned} \dot{U} &= \dot{U}^+ + \dot{U}^- = \frac{Z_g^*}{\sqrt{\operatorname{Re}(Z_g)}} \dot{i}^+ + \frac{Z_g}{\sqrt{\operatorname{Re}(Z_g)}} \dot{i}^- \\ I &= I^+ - I^- = \frac{1}{\sqrt{\operatorname{Re}(Z_g)}} \dot{i}^+ - \frac{1}{\sqrt{\operatorname{Re}(Z_g)}} \dot{i}^-, \end{aligned}$$

то

$$\dot{i}^+ = \frac{\dot{U} + Z_g I}{2 \sqrt{\operatorname{Re}(Z_g)}}, \quad \dot{i}^- = \frac{\dot{U} - Z_g^* I}{2 \sqrt{\operatorname{Re}(Z_g)}}.$$

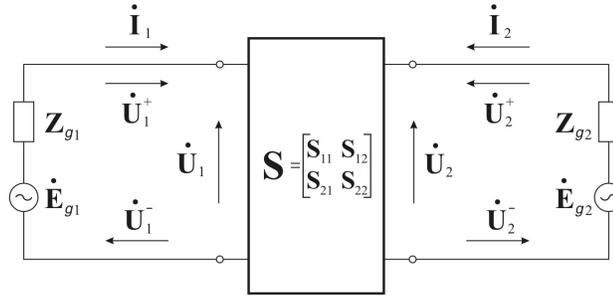
Как и раньше, комплексный коэффициент отражения рассматриваемой однопортовой цепи (нагрузки)

$$\Gamma = \dot{u}^- / \dot{u}^+,$$

или

$$\Gamma = \frac{\dot{U} - Z_g^* \dot{I}}{\dot{U} + Z_g \dot{I}},$$

суть ее единственный параметр рассеяния S .



Генераторы, нормализованные волны и полные токи и напряжения двухпортовой цепи.

Перейдем теперь к двухпортовой цепи (четырёхполюснику). В общем случае она возбуждается с обоих портов (рис. 10.9.3), по аналогии со случаем однопортовой цепи можно записать:

$$\dot{u}_1^+ = \frac{\dot{U}_1 + Z_{g1} \dot{I}_1}{2\sqrt{\text{Re}(Z_{g1})}}, \quad \dot{u}_2^+ = \frac{\dot{U}_2 + Z_{g2} \dot{I}_2}{2\sqrt{\text{Re}(Z_{g2})}},$$

и основная система в терминах параметров рассеяния повторяет уравнения:

$$\dot{u}_1^- = S_{11} \dot{u}_1^+ + S_{12} \dot{u}_2^+, \quad (10.6.4)$$

$$\dot{u}_2^- = S_{21} \dot{u}_1^+ + S_{22} \dot{u}_2^+. \quad (10.6.5)$$

Наконец, в общем случае n -портовой цепи падающие нормализованные волны:

$$\dot{u}_k^+ = \frac{\dot{U}_k + Z_{gk} \dot{I}_k}{2\sqrt{\text{Re}(Z_{gk})}}, \quad k = 1, 2, \dots, n,$$

и основная система в терминах параметров рассеяния повторяет уравнения:

$$\dot{\mathbf{U}}^- = \mathbf{S} \dot{\mathbf{U}}^+, \quad (10.5.1)$$

или в расписанном виде:

$$\dot{u}_1^- = S_{11} \dot{u}_1^+ + S_{12} \dot{u}_2^+ + \dots + S_{1n} \dot{u}_n^+ \quad (10.5.2)$$

$$\dot{u}_2^- = S_{21}u_1^+ + S_{22}\dot{u}_2^+ + \dots + S_{2n}\dot{u}_n^+$$

.....

.....

$$\dot{u}_n^- = S_{n1}u_1^+ + S_{n2}\dot{u}_2^+ + \dots + S_{nn}\dot{u}_n^+$$

Изменение опорных плоскостей

Как отмечено выше, комплексные амплитуды нормализованных волн относятся к опорным плоскостям. Следовательно, и связывающие эти волны параметры рассеяния также относятся к опорным плоскостям; при изменении положения хотя бы одной опорной плоскости параметры рассеяния, вообще говоря, изменяются. Выявим закономерность этих изменений. Пусть характеризуемое параметрами рассеяния устройство является сочленением n однородных линий передачи без потерь с коэффициентами фазы β_k , $k = 1, 2, \dots, n$; опорные плоскости изменили свое положение на Δl_k , $k = 1, 2, \dots, n$, не обязательно все отличные от нуля (положительное направление сдвига опорной плоскости – от сочленения). Из (10.6.2), (10.6.3) следует, что новое значение S'_{km} параметра рассеяния выражается через старое значение S_{km} как

$$S'_{km} = S_{km} \exp\{-i(\beta_k \Delta l_k + \beta_m \Delta l_m)\}, \quad k, m = 1, 2, \dots, n. \quad (10.10.1)$$

В соответствии с этой системой соотношений, новая матрица \mathbf{S}' рассеяния выражается через «старую» \mathbf{S} как

$$\mathbf{S}' = \mathbf{L} \mathbf{S} \mathbf{L},$$

где \mathbf{L} – диагональная матрица порядка n , диагональные элементы которой имеют значения $l_{kk} = \exp\{-i\beta_k \Delta l_k\}$. В рассматриваемом частном случае линий без потерь изменение положения опорных плоскостей влияет только на аргументы (фазы) параметров рассеяния, оставляя неизменными их модули (это полезное свойство не выполняется для классических характеристических матриц: при сдвиге опорных плоскостей элементы этих матриц меняются как по фазе, так и по амплитуде). В общем случае линий передачи с потерями приведенные соотношения остаются в силе, если заменить в них $i\beta_k$ на постоянные распространения $\gamma_k = \alpha_k + i\beta_k$, где α_k – коэффициенты затухания линий. Из (10.10.1) следует, что любой комплексный параметр рассеяния можно сделать действительным, изменив определенным образом положение соответствующей опорной плоскости. Так как всего опорных плоскостей n , то максимальное число таких превращений также равно n , в то время как всего параметров рассеяния n^2 .

Математические свойства матриц рассеяния

В зависимости от внутреннего устройства многополюсника его матрица рассеяния имеет те или иные математические свойства [10.9]. Чем более подробно известны внутренние свойства многополюсника, тем больше, в принципе, свойств можно обнаружить у его матрицы рассеяния. Однако существуют важные классы устройств, которым соответствуют довольно общие свойства матриц рассеяния, важные в приложениях. Приведем эти свойства без вывода. Как показано выше, матрица \mathbf{S} рассеяния существует для любого пассивного линейного n -портового устройства, имеет размер $(n \times n)$, ее элементы в общем случае комплексны и безразмерны.

1) Устройство СВЧ, в котором использованы только изотропные материалы, является взаимным. Матрица \mathbf{S} рассеяния взаимного устройства симметрична, т. е. $\mathbf{S}^T = \mathbf{S}$, или $S_{ij} = S_{ji}$, $i, j = 1, 2, \dots, n$ (значок « T » означает транспонирование). Справедливо и обратное: если матрица рассеяния устройства симметрична, то оно взаимно.

2) Если многополюсник не имеет потерь, он называется недиссипативным. Матрица рассеяния недиссипативного устройства унитарна, т. е.

$$\mathbf{S}^* \mathbf{S} = \mathbf{I},$$

где \mathbf{I} – единичная матрица порядка n , «звездочка» означает эрмитово сопряжение, т. е. последовательные транспонирование и комплексное сопряжение элементов. Справедливо и обратное: если матрица рассеяния устройства унитарна, то оно недиссипативно. Из теории матриц известно, что столбцы (а также строки) унитарной матрицы ортонормированы:

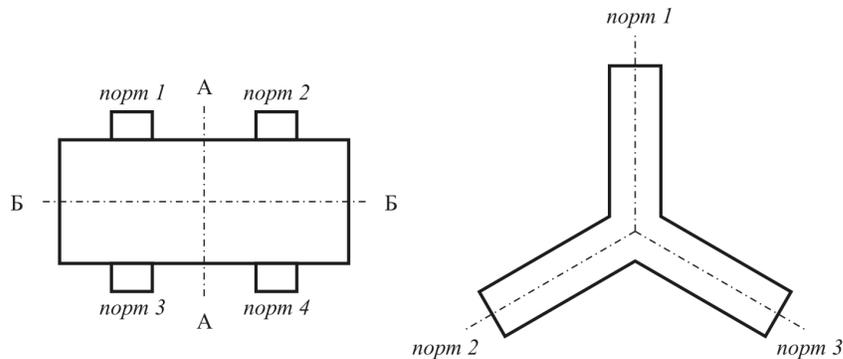
$$\sum_{k=1}^n S_{kr}^* S_{kl} = \begin{cases} 0, & r \neq l \\ 1, & r = l, \quad r, l = 1, 2, \dots, n \end{cases},$$

т. е. модуль вектора-столбца (вектора-строки) унитарной матрицы равен единице:

$$\sum_{k=1}^n |S_{kl}|^2 = 1, \quad l = 1, 2, \dots, n,$$

а скалярное произведение двух разных векторов-столбцов (векторов-строк) равно нулю.

3) Устройство СВЧ называется симметричным, если оно не изменяется (инвариантно) при преобразовании симметрии, т. е. повороте устройства вокруг некоторой оси на какой-то угол, зеркальном отображении относительно какой-либо плоскости и т. п. Существуют различные типы симметрии, им соответствуют различные преобразования симметрии. Любое преобразование симметрии есть перенумерация портов без изменения самого устройства. На рис. 10.11.1 приведены простые примеры симметричных многополюсников: четырехпортовое устройство с зеркальной симметрией относительно плоскостей А-А и Б-Б (рис. 10.11.1(а)) (устройство не меняется при замене номеров портов $1 \rightarrow 3$, $2 \rightarrow 4$ и наоборот, и (или) номеров портов $1 \rightarrow 2$, $3 \rightarrow 4$ и наоборот); симметричный Y-тройник с поворотной симметрией на 120° (рис. 10.11.1(б)) (устройство не меняется при замене номеров портов $1 \rightarrow 2$, $2 \rightarrow 3$, $3 \rightarrow 1$ и наоборот).



(слева, справа)

Слева: четырехпортовое устройство с зеркальной симметрией относительно плоскостей А-А и Б-Б. Справа: симметричный тройник с поворотной симметрией на 120°

Общий метод учета симметрии данного устройства и формулировки следующих из нее соотношений элементов матрицы рассеяния таков:

А) Для каждой отдельной симметрии составляется *матрица симметрии*. Это квадратная матрица \mathbf{A} того же порядка n , что и матрица \mathbf{S} . В каждом столбце (и в каждой строке) имеется один элемент, равный 1, остальные равны нулю (поэтому \mathbf{A} унитарна); для единичного элемента, стоящего на пересечении j -го столбца и i -ой строки, j – номер порта до преобразования симметрии, i – номер того же порта после преобразования симметрии. Например, для симметричного тройника (рис. 10.11.1 (справа)) матрица симметрии имеет вид:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix};$$

четырёхпортовое устройство с двумя плоскостями симметрии (рис. 10.11.1 (слева)) имеет две симметрии: относительно плоскости А-А и относительно плоскости Б-Б, им соответствуют матрицы симметрии:

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

(заметим, что \mathbf{A}_1 симметрична относительно главной диагонали, а \mathbf{A}_2 – относительно побочной диагонали). Обратное преобразование симметрии описывается матрицей симметрии \mathbf{A}^{-1} , а поскольку \mathbf{A} унитарна, то $\mathbf{A}^{-1} = \mathbf{A}^T$.

Б) Следующие из симметрии устройства свойства элементов матрицы \mathbf{S} получаются из системы равенств (одного матричного равенства):

$$\mathbf{AS} = \mathbf{SA},$$

означающей, что матрицы \mathbf{A} и \mathbf{S} коммутируют. Если у устройства несколько симметрий, которым соответствует несколько матриц симметрии, то \mathbf{S} коммутирует с каждой из них. Например, для приведенного на рис. 10.11.1 (слева) устройства одновременно справедливы равенства

$$\mathbf{A}_1\mathbf{S} = \mathbf{SA}_1, \quad \mathbf{A}_2\mathbf{S} = \mathbf{SA}_2.$$

Несмотря на наличие общего метода учета симметрий, часто следующие из симметрии устройства свойства элементов матрицы \mathbf{S} находят непосредственно, глядя на рисунок и не прибегая к построению матрицы симметрии.

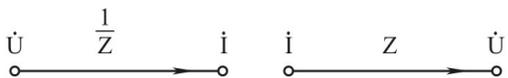
Главное значение рассмотренных общих математических свойства матриц рассеяния состоит в том, что использование этих свойств (если к этому есть основания: взаимность, отсутствие потерь и (или) симметрии) позволяет уменьшить число независимых элементов матрицы рассеяния и, тем самым, число параметров рассеяния, которые необходимо непосредственно измерять.

ПОТОКОВЫЕ СИГНАЛЬНЫЕ ГРАФЫ

Топологические и сигнальные графы

Применительно к линейным электрическим цепям *графы* являются наглядными геометрическими моделями, упрощающими и упорядочивающими математическое рассмотрение свойств и расчет цепей и сигналов в них. Существует два вида таких графов: *топологический граф цепи* и *сигнальный граф цепи*. Топологический граф [11.1, 11.2] отражает *геометрическую структуру* (топологию) сложной цепи, характеризуемую совокупностью узлов и соединяющих их *ветвей*, независимо от конкретных особенностей элементов, образующих ветви. Обычно топологический граф используется наряду с электрической схемой цепи или эквивалентной электрической схемой цепи. В сравнении с электрической схемой топологический граф сохраняет узлы схемы, а ветви заменяются линиями, форма которых может быть произвольной. Обычно источники энергии на топологическом графе не указываются; генераторы э.д.с. исходной схемы заменяются на графе короткозамкнутыми участками, а генераторы тока – разрывами. Если на топологическом графе указаны направления токов, то он называется *направленным*, в противном случае – *ненаправленным*. Топологические графы используются, например, для выбора независимых контуров и узлов при записи системы уравнений Кирхгофа, или для непосредственного расчета цепей [11.1, 11.2]. Далее топологические графы не рассматриваются и не используются.

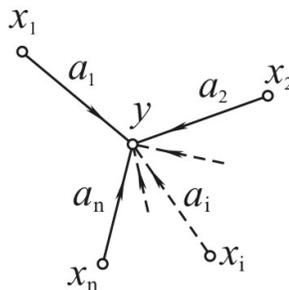
Сигнальный граф отражает не геометрическую структуру цепи, а функциональные зависимости между процессами, происходящими в цепи, т. е. между токами и напряжениями, или падающими и отраженными волнами и т. д. Поскольку алгебраически эти зависимости описываются системой линейных алгебраических уравнений, можно сказать, что сигнальные графы геометрически отражают соотношения между токами и напряжениями, падающими и отраженными волнами и т. п. в цепи. При этом сигнальный граф определяется, как правило, не конфигурацией цепи, а структурой соответствующей системы уравнений. Сигнальные графы обладают рядом особенностей. В частности, узлы графа отображают *заданные (независимые)* и *искомые (зависимые)* переменные; узлы соединяются друг с другом ветвями; с каждой ветвью связываются направление (значит, сигнальный граф всегда направленный) и *вес (усиление)*; вес количественно определяет влияние одной переменной (причины) на другую (следствие).



Слева: сигнальный граф для закона Ома при \dot{U} как независимой переменной, справа: то же при \dot{I} как независимой переменной

Дадим простейшие иллюстрации сказанному. Пусть \dot{U} – комплексная амплитуда напряжения на некотором комплексном сопротивлении Z , а \dot{I} – комплексная амплитуда тока на нем. Величины \dot{U} , \dot{I} связаны линейным соотношением $\dot{I} = \dot{U} / Z$, называемым законом Ома. Если считать \dot{U} независимой переменной, а \dot{I} – зависимой переменной, то сигнальный граф для закона Ома имеет вид рис. 11.1.1 (слева); если считать \dot{I} независимой, а \dot{U} – зависимой переменной, то граф имеет вид рис. 11.1.1 (справа). Другой пример: пусть зависимая переменная Y суть линейная функция независимых переменных

$X_i, i=1, \dots, n: Y = \sum_{i=1}^n a_i X_i$. Тогда соответствующий сигнальный граф имеет вид, показанный на рис. 11.1.2.



Граф для линейной функции независимых переменных

Сигнальный граф и матричное представление линейной цепи. Поточковый граф

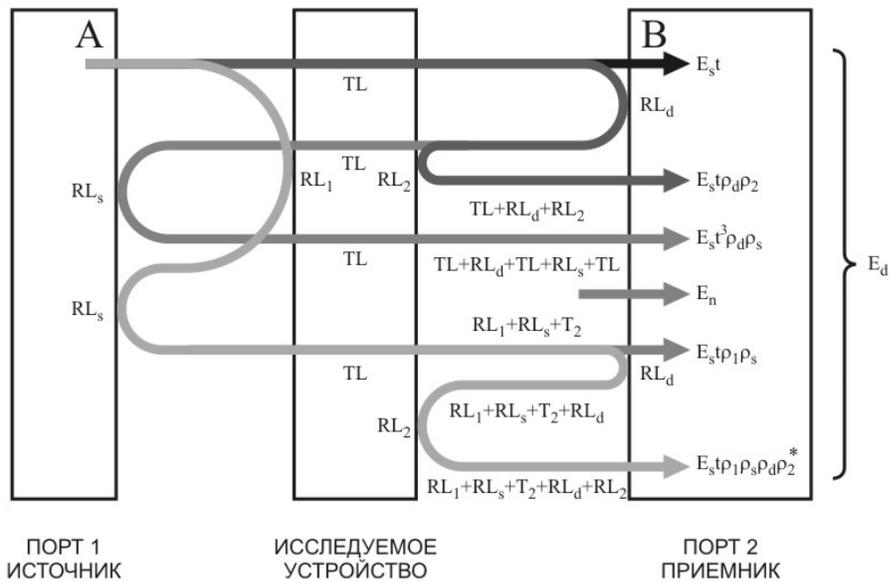
Решить сигнальный граф значит по заданным независимым переменным и весам ветвей найти некоторые или все зависимые переменные. При этом используются определенные правила, важнейшим из которых является рассмотренное ниже «правило непересекающихся петель» Мэйсона [11.3, 11.4]. Поскольку, как указано выше, сигнальный граф отображает линейные соотношения между процессами в реальной или эквивалентной цепи, то решение графа соответствует частичному или полному решению системы этих соотношений.

Узел, соответствующий независимой переменной, иногда называют *истоком*; такой узел имеет только *исходящие ветви*. Узел, имеющий только входящие ветви, тогда называется *стоком*, такой узел имеет только *входящие ветви*. Сигнальный граф может содержать также узлы третьего вида, имеющие как входящие, так и исходящие ветви; такой узел называется *смешанным*. Если переменная, соответствующая смешанному узлу, не подлежит определению при решении графа, то такой узел может быть устранен, т. е. построен новый, более простой эквивалентный граф. Эквивалентность означает [11.5], что соотношения между $n - 1$ оставшимися переменными в новом графе точно такие же, как в первоначальном графе.

Поскольку сигнальный граф адекватен некоторой системе линейных уравнений для процессов цепи, возникает вопрос, в чем преимущества графического представления и решения графа в сравнении с непосредственным решением соответствующей системы уравнений. Действительно, любую систему линейных уравнений можно решить регулярными алгебраическими методами, например, методом исключения (методом Гаусса). Часто, однако, матрица системы уравнений имеет высокий порядок, что обуславливает большую громоздкость и малую обозримость решений, т. е. трудность интерпретации результатов. Преимущества графического подхода возникают, во-первых, из того, что на практике часто не требуется находить все зависимые переменные системы, а лишь некоторые из них (иногда даже одну переменную), т. е. полного решения системы не требуется. Во-вторых, обычно матрица системы имеет большое число нулей, что позволяет эффективно исключать узлы графа, причем само рассматривание графа позволяет выявить оптимальный порядок исключения узлов. В-третьих, решение с помощью графа обладает большей наглядностью и облегчает интерпретацию результатов решения.

Как известно [11.5], для любой заданной линейной системы (не обязательно радиотехнической) имеется множество способов выбора переменных и записей уравнений. С каждым набором уравнений связывается свой сигнальный граф, и таким образом данная линейная система может быть представлена с помощью множества различных сигнальных графов. Сказанное полностью относится к радиотехнической цепи, работающей в режиме переменных сигналов. Даже если ограничиться лишь эквивалентными радиотехническими цепями в виде линейных $2n$ -полюсников (n -портовых цепей) и переменными только в виде токов и напряжений, то, как показано в главе 10, для одной n -портовой цепи можно составить $C_{2n}^2 = n(2n-1)$ различных систем линейных уравнений и, следовательно, различных сигнальных графов этой цепи с различными по физическому смыслу узлами. Как видим, число различных графов одной n -портовой цепи быстро нарастает с увеличением n : при $n=2$ их 6, включая графы, соответствующие \mathbf{Y} , \mathbf{Z} , \mathbf{B} , \mathbf{A} , \mathbf{H} , \mathbf{F} -матричным представлениям, а при $n=3$ различных графов уже 15. Возникает вопрос: какое из представлений системой уравнений относительно токов и напряжений и соответствующих сигнальных графов наиболее удобно для анализа цепей, создаваемых при тестировании устройств СВЧ с помощью анализаторов цепей, для спектрального анализа и создаваемых при этом цепей и т.д., а также в измерениях с целью калибровки приборов СВЧ? Ответ состоит в том, что все варианты выбора переменных в виде токов и напряжений неудобны и неперспективны для достижения этих целей. По причинам, указанным в главе 10, наилучшим является выбор переменных не в виде токов и напряжений, а в виде нормированных падающих и отраженных волн на разъемах цепи (волн мощности). В этом случае система линейных уравнений (10.5.1), описывающая состояние линейной радиотехнической цепи в выбранном базисе плоскостей отсчета, имеет в качестве коэффициентов элементы матрицы \mathbf{S} рассеяния, т. е. реализует \mathbf{S} -матричное представление. Этому представлению соответствует специальный вид сигнальных графов – *сигнальный потоковый граф* (далее прилагательное «сигнальный» опускаем).

В дополнение к высказанным выше аргументам в пользу применения вообще сигнальных графов, можно добавить соображения о полезности именно потоковых графов. Поскольку эти графы соответствуют представлению сигналов через падающие и отраженные волны, для интерпретации сигналов приходится учитывать отражения всех порядков от неоднородностей цепи. При этом уже цепь с более чем одной внутренней неоднородностью становится трудной для анализа путем бесконечного суммирования отраженных волн. На рис.11.2.1 [11.9] показаны лишь отражения невысокого порядка в системе «анализатор цепей – тестируемое устройство».



Отражения невысокого порядка в системе «анализатор – тестируемая цепь»

Хотя S-матричное представление формально решает эту проблему, проникновение в физический смысл соотношений в цепи уменьшается с увеличением ее сложности. Поточковый граф цепи, во-первых, помогает уяснить этот физический смысл, во-вторых, позволяет учесть отсутствующие на принципиальной схеме паразитные потоки, связанные с нежелательными отражениями и протечками, и в-третьих, позволяет проводить упрощения цепи путем хорошо обоснованных физически аппроксимаций и приближений.

Зафиксируем более точный и более узкий смысл, в котором далее будем использовать понятие потокового графа.

1) Поточковый граф будет пониматься как сигнальный граф, соответствующий S-матричному описанию линейной эквивалентной цепи, состоящей из соединенных линиями передачи без потерь $2n$ -полюсников.

2) *Переменными*, которым соответствуют узлы потокового графа, будут комплексные амплитуды монохроматических волн мощности, нормированные соответствующим образом (см. главу 10).

3) Описываемая потоковым графом цепь может включать как распределенные, так и сосредоточенные элементы.

4) Поточковый граф может включать узлы и ветви, отражающие существование и распространение паразитных волн, обязанных протечкам, отражениям от паразитных элементов и другим явлениям, не учтенным электрической эквивалентной схемой цепи.

5) Каждому порту эквивалентной принципиальной схемы соответствует два узла графа: узел типа *исток*, из которого исходят направленные ветви, соответствующий входящей (падающей извне, например, от генератора) волне, при необходимости обозначается буквой *a* с соответствующим индексом; узел типа *сток*, к которому подходят направленные ветви, соответствующий выходящей наружу (отраженной) волне, при необходи-

мости обозначается буквой b с соответствующим индексом. Например, в случае двухпортовой цепи появляются узлы a_1, a_2, b_1, b_2 и, возможно, узел b_s , связанный с генератором. Разумеется, как и во всяком сигнальном графе, в потоковом графе могут быть смешанные (внутренние) узлы.

6) С каждой ветвью связан и указывается рядом с ней *множитель ветви* (в общем случае комплексный).

7) Ветви начинаются в узлах независимых переменных и заканчиваются на узлах зависимых переменных (зависимость и независимость устанавливаются соответствующими уравнениями). На каждой ветви указывается *стрелкой* направление потока и числом или буквой – множителем.

8) Если поток от узла независимой переменной к узлу зависимой переменной проходит без изменения, то на связывающей их ветви указывается множитель 1.

9) Величина каждой зависимой переменной определяется множителями и независимыми переменными, связанными с ветвями от соответствующих узлов.

10) Хотя сказанное выше о структуре сигнального графа, естественно, остается в силе и для потокового графа, в построении, виде и интерпретации последнего имеется своя специфика, диктуемая, в основном, удобством работы с ним. Эта специфика будет проиллюстрирована примерами.

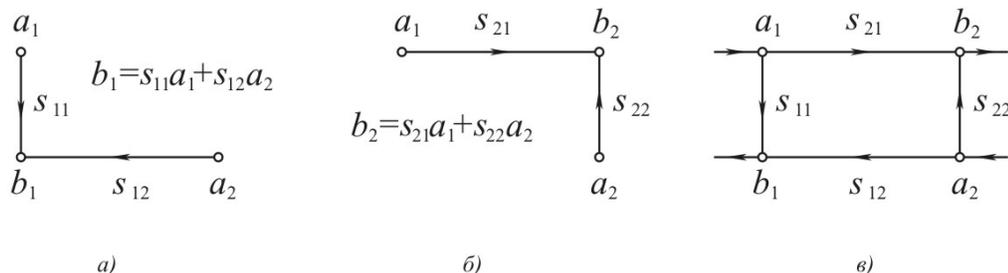
Декларированное выше удобство потоковых графов для анализа процессов измерения параметров устройств СВЧ с помощью анализаторов цепей было осознано давно и объясняется, во-первых, тем, что описываемая этими графами цепь может включать как распределенные, так и сосредоточенные элементы, во-вторых, простотой учета с помощью этих графов паразитных волн, что необходимо в рассмотрении систематических и случайных ошибок измерения и методов калибровки анализаторов цепей. Другим объектом стандартного применения потоковых графов, восходящего к работам [11.3, 11.4], являются цепи с обратными связями.

Построение потокового графа

Рис. 11.3.1, 11.3.2 иллюстрируют построение потоковых графов по уравнениям в S-параметрах или по принципиальным схемам в рамках ненагруженного и нагруженного четырехполюсников. Рис. 11.3.1 поясняет составление графа ненагруженного четырехполюсника, соответствующего S-параметрической системе уравнений

(11.3.1)

$$\begin{aligned} b_1 &= S_{11}a_1 + S_{12}a_2 \\ b_2 &= S_{21}a_1 + S_{22}a_2 \end{aligned}$$



Составление графа ненагруженной двухпортовой цепи

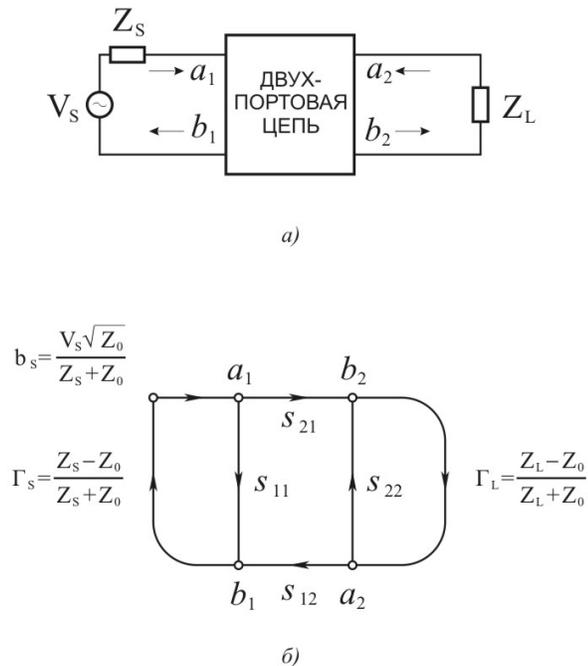


Рис. 11.3.2. Схема и граф нагруженной двухпортовой цепи.

Узлы a_1, a_2 соответствуют независимым переменным и являются истоками, узлы b_1, b_2 – зависимым переменным и являются стоками. Напряжения от генераторов могут подаваться на вход четырехполюсника, т. е. на графе – на узел a_1 , или на выход, т. е. на графе – на узел a_2 , или на оба узла сразу. Направления потоков указаны *стрелками*, весами являются S-параметры, т. е. элементы S-матрицы четырехполюсника. Обратим внимание на генезис графа четырехполюсника (рис. 11.3.1)): первое уравнение системы представлено в виде графа на рис. 11.3.1а, наглядно показывающего, как зависимая переменная b_1 представляет линейную комбинацию независимых переменных a_1, a_2 ; аналогично устроен подграф на рис. 11.3.1б, представляющий второе уравнение системы. Однако, для получения удобного вида объединения этих подграфов в общий граф, второй подграф «перевернут».

На рис. 11.3.2б показан граф нагруженного четырехполюсника с характеристическим сопротивлением Z_0 ; эквивалентная схема соответствующей цепи приведена на рис. 11.3.2а: слева подключен генератор с комплексной амплитудой напряжения V_s и с внутренним (комплексным) сопротивлением Z_s , справа – нагрузка с комплексным сопротивлением Z_L . Нормированная комплексная амплитуда падающей волны от генератора

$$b_s = \frac{V_s \sqrt{Z_0}}{Z_s + Z_0}, \quad (11.3.2)$$

формально отображена на графе узлом b_s ; комплексный коэффициент отражения генератора

$$\Gamma_s = \frac{Z_s - Z_0}{Z_s + Z_0}, \quad (11.3.3)$$

формально отображен на графе ветвью от b_1 к a_1 с весом Γ_s ; комплексный коэффициент отражения нагрузки

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0}, \quad (11.3.4)$$

формально отображен на графе ветвью от b_2 к a_2 с весом Γ_L .

В результате нагружения четырехполюсника слева и справа все четыре узла a_1, a_2, b_1, b_2 становятся смешанными узлами, а это означает, что при желании граф может быть упрощен (далее техника упрощения графа путем исключения смешанных узлов не излагается, а применяется общий метод решения графа).

Решение потокового графа

В общем смысле, решение графа есть вычисление зависимых переменных по независимым переменным, или, на специфическом языке теории графов, вычисление одних узлов по другим узлам. Излагаемое ниже правило решения графа позволяет вычислять зависимые переменные, или просто узлы-стоки по одному, т. е. общее решение графа включает набор решений для индивидуальных узлов-стоков (назовем их *частными решениями*). Иногда нет необходимости во всех частных решениях, а нужно найти только некоторые из них, а иногда даже только одно из них. Таким образом, необходимо правило нахождения частного решения для конкретного узла-стока. Определим некоторые дополнительные термины.

Путь от одного узла к другому есть непрерывная последовательность ветвей, связывающая эти узлы, такая, что ни один узел не проходится более одного раза; непрерывность здесь означает, что в этой последовательности конец одной ветви является началом следующей.

Прямой путь есть путь, в котором каждая ветвь проходится только в направлении ее стрелки. В дальнейшем нам понадобятся только прямые пути. Два узла могут связывать и несколько прямых путей, отличающихся одной или несколькими ветвями.

Путевой множитель (путевое усиление) есть произведение всех множителей ветвей вдоль прямого пути. Каждый прямой путь имеет, вообще говоря, свой путевой множитель. Например, в графе на рис. 11.3.2б есть один прямой путь от узла a_1 (или, что то же самое, от узла b_s) к узлу b_2 с путевым множителем S_{21} , два прямых пути от узла a_1 к узлу b_1 с путевыми множителями S_{11} и $S_{21}\Gamma_L S_{12}$ соответственно, один прямой путь от узла a_2 к узлу b_1 с путевым множителем S_{12} и два прямых пути от узла a_2 к узлу b_2 с путевыми множителями S_{22} и $S_{12}\Gamma_s S_{21}$ соответственно.

Петля (прямая петля) есть прямой путь, который начинается и заканчивается на одном и том же узле. Один и тот же узел может состоять в нескольких петлях.

Петлевой множитель (петлевое усиление) есть произведение множителей ветвей вдоль петли. Например, в графе на рис. 11.3.2б есть две петли, начинающихся и заканчи-

вающихся на узле a_1 с петлевыми множителями $S_{11}\Gamma_S$ и $S_{21}\Gamma_L S_{12}\Gamma_S$ соответственно, две петли, начинающихся и заканчивающихся на узле a_2 с петлевыми множителями $S_{22}\Gamma_L$ и $S_{12}\Gamma_S S_{21}\Gamma_L$ соответственно.

Две или более петель называются *непересекающимися* (*несоприкасающимися*), если они не имеют не только общих ветвей, но даже общих узлов. Например, в графе на рис. 11.3.2б не пересекаются петля в левой части с петлевым множителем $S_{11}\Gamma_S$ и петля в правой части с петлевым множителем $S_{22}\Gamma_L$.

Петля первого порядка – понятие тождественное понятию *петля*. Например, петли с петлевыми множителями $S_{11}\Gamma_S$, $S_{12}\Gamma_S S_{21}\Gamma_L$ и $S_{22}\Gamma_L$ суть петли первого порядка.

Петля второго порядка есть набор из двух непересекающихся петель первого порядка. *Множитель петли второго порядка* равен произведению множителей составляющих ее петель. Например, в графе на рис. 11.3.2б есть одна петля второго порядка, она объединяет две вышеупомянутых непересекающихся петли и имеет множитель $S_{11}\Gamma_S S_{22}\Gamma_L$.

Петля n-го порядка ($n=2, 3, \dots$) есть набор из n непересекающихся петель первого порядка. *Множитель петли n-го порядка* равен произведению множителей составляющих ее петель. Например, в графе на рис. 11.3.2б нет петель третьего и более высоких порядков.

На языке теории графов узел, ветвь, путь, петля могут пониматься как геометрические объекты или как величины (комплексные числа). *Узел как величина* есть значение соответствующей этому узлу переменной; *ветвь, путь, петля как величины* есть значения множителей этих объектов.

Переходя к технике отыскания частных решений потокового графа, сделаем три замечания.

1) Величина узла-стока (т. е. величина зависимой переменной, соответствующей этому узлу) равна сумме вкладов в этот узел от всех узлов-истоков, от которых имеется прямой путь к этому узлу. Это становится очевидным, если вспомнить, что потоковый граф есть графическое отображение соответствующей S -параметрической системы уравнений, в которой каждое уравнение представляет одну зависимую переменную как линейную комбинацию независимых переменных.

2) Вклад в зависимую переменную от одной независимой переменной, т. е. вклад в узел-сток от одного узла-истока находится по «правилу непересекающихся петель» Мэйсона (S. J. Mason), который предложил его в 50-е годы прошлого столетия и систематически его применял (см., например, [11.3, 11.4]). Строгий вывод этого правила на основе свойств определителей дан в [11.5].

3) Согласно этому правилу, вклад в зависимую переменную от одной независимой переменной равен произведению последней на некоторый множитель (*переносная функция* [11.9]), представляющий собой некоторое выражение от множителей ветвей графа. Обычно в качестве правила используют формулу для этого множителя, т. е. для отношения конкретной зависимой переменной к конкретной независимой переменной. Чтобы найти полное частное решение для конкретной зависимой переменной, находим вклад от каждой независимой переменной путем умножения переносной функции на величину последней и сложения всех таких произведений по независимым узлам, дающим вклад.

Формула для переносной функции от независимой переменной a_i к зависимой переменной b_j имеет вид:

$$T(a_i, b_j) = \frac{b_j}{a_i} = \frac{\Delta_{ij}}{\Delta}, \quad (11.4.1)$$

где

$$\Delta = 1 - \sum L(1) + \sum L(2) - \sum L(3) + \dots, \quad (11.4.2)$$

$$\Delta_{ij} = P_1[1 - \sum L(1)^{(1)} + \sum L(2)^{(1)} - \dots] + P_2[1 - \sum L(1)^{(2)} + \sum L(2)^{(2)} - \dots] + P_3[\dots], \quad (11.4.3)$$

P_1, P_2, \dots – прямые пути для данной пары узлов;

$\sum L(1)$ – сумма всех петель первого порядка графа;

$\sum L(2)$ – сумма всех петель второго порядка графа; ...

$\sum L(1)^{(n)}$, ($n = 1, 2, \dots$) – сумма петель первого порядка, не касающихся ни в одной точке прямого пути P_n ;

$\sum L(2)^{(n)}$ ($n = 1, 2, \dots$) – сумма петель второго порядка, не касающихся ни в одной точке прямого пути P_n ; ...

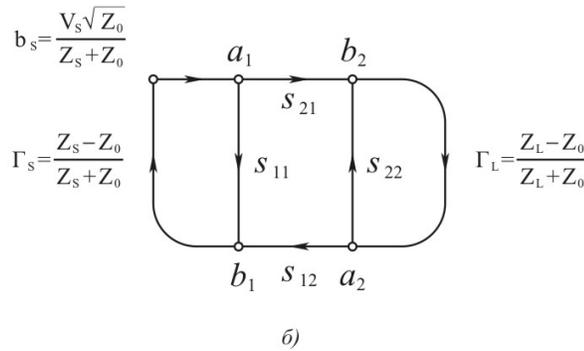
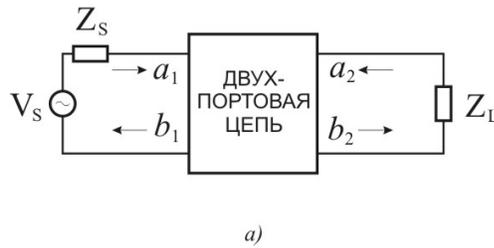


Рис. 11.3.2. Схема и граф нагруженной двухпортовой цепи.

Заметим, что знаменатель Δ формулы (11.4.1) одинаков для всех пар «независимый узел – зависимый узел» и зависит только от топологии графа. В графе на рис. 11.3.2б три петли первого порядка и их сумма

$$\sum L(1) = S_{11}\Gamma_s + S_{21}S_{12}\Gamma_s\Gamma_L + S_{22}\Gamma_L, \quad (11.4.4)$$

одна петля второго порядка и

$$\sum L(2) = S_{11}S_{22}\Gamma_S\Gamma_L, \quad (11.4.5)$$

нет петель третьего и более высоких порядков, в итоге

$$\Delta = 1 - (S_{11}\Gamma_S + S_{21}S_{12}\Gamma_S\Gamma_L + S_{22}\Gamma_L) + S_{11}S_{22}\Gamma_S\Gamma_L. \quad (11.4.6)$$

Вычисляем $T(a_1, b_2) = T(b_s, b_2)$. Для этой пары узлов имеется только один прямой путь: $P_1 = S_{21}$ и нет петель, не соприкасающихся с этим путем, поэтому

$$T(a_1, b_2) = \Delta_{12} / \Delta = S_{11} / [1 - (S_{11}\Gamma_S + S_{21}S_{12}\Gamma_S\Gamma_L + S_{22}\Gamma_L) + S_{11}S_{22}\Gamma_S\Gamma_L]. \quad (11.4.7)$$

Вычисляем $T(a_1, b_1)$. Для этой пары узлов имеется два прямых пути: $P_1 = S_{11}$ и $P_2 = S_{21}\Gamma_L S_{12}$. С путем P_1 не соприкасается только петля первого порядка $L(1)^{(1)} = S_{22}\Gamma_L$; петля, не соприкасающихся с путем P_2 нет. В итоге

$$T(a_1, b_1) = \frac{\Delta_{11}}{\Delta}, \quad (11.4.8)$$

где

$$\Delta_{11} = P_1[1 - L(1)^{(1)}] + P_2 = S_{21}[1 - S_{22}\Gamma_L] + S_{21}S_{12}\Gamma_L. \quad (11.4.9)$$

Вычисляем $T(a_2, b_1)$. Для этой пары узлов имеется только один прямой путь: $P_1 = S_{12}$ и нет петель, не соприкасающихся с этим путем, поэтому

$$T(a_2, b_1) = \frac{\Delta_{21}}{\Delta} = \frac{S_{12}}{\Delta} \quad (11.4.10)$$

Вычисляем $T(a_2, b_2)$. Для этой пары узлов имеется два прямых пути: $P_1 = S_{22}$ и $P_2 = S_{12}\Gamma_S S_{21}$. С путем P_1 не соприкасается только петля первого порядка $L(1)^{(1)} = S_{11}\Gamma_S$; петля, не соприкасающихся с путем P_2 , нет. В итоге

$$T(a_2, b_2) = \frac{\Delta_{22}}{\Delta}, \quad (11.4.11)$$

где

$$\Delta_{22} = P_1[1 - L(1)^{(1)}] + P_2 = S_{22}[1 - S_{11}\Gamma_S] + S_{21}S_{12}\Gamma_S. \quad (11.4.12)$$

ДИАГРАММА ВОЛЬПЕРТА-СМИТА

Предисловие

Круговая диаграмма полных сопротивлений (круговая диаграмма, диаграмма Смита, диаграмма Вольперта-Смита, далее – круговая диаграмма, КД) – это графическое средство, или номограмма, помогающее в решении задач, связанных с линиями передачи

и согласующими цепями. Начиная с 1939 г., когда она была предложена независимо Ф. Смитом и Вольпертом, популярность КД неуклонно росла с годами и в 60-е – 80-е годы XX века достигла невиданной высоты: ни одна номограмма в радиоэлектронике, а может быть – в любой области знания, не могла в этом отношении сравниться с КД. Она стала символом радиоэлектроники СВЧ диапазона.

В пике своей популярности КД использовалась не только для облегчения решения задач, но и для наглядной графической иллюстрации поведения радиотехнических параметров с частотой, как альтернатива табличному представлению, а также для представления многих характеристик [12.1, 12.4], таких как окружности сопротивления и проводимости, коэффициент отражения, параметры рассеяния, коэффициент шума, контуры постоянного усиления и области безусловной стабильности генераторов. Наиболее часто КД используется в области круга единичного радиуса, однако построения вне этого круга могут быть также полезны, например, в конструировании генераторов и анализе стабильности.

С появлением и развитием персональных компьютеров и специализированных электродинамических пакетов роль КД несколько угасла, так как во многих случаях расчет и его иллюстрация другими средствами на компьютере стали проще, быстрее и точнее. Однако и сейчас КД широко применяется и является стандартным элементом образования радиоинженера. Не в последнюю очередь причиной этому является привычка, благодаря которой опытному инженеру отображения на КД кажутся наглядными и весьма содержательными. Недаром, например, до сих пор формат КД – один из обязательных форматов дисплейного отображения векторного анализатора цепей.

КД изображается на плоскости комплексного коэффициента отражения в двух измерениях и масштабируется или в нормализованном импедансе (Z -диаграмма), или в нормализованной проводимости (Y -диаграмма). Нормализованное масштабирование позволяет использовать КД для задач, включающих любые характеристические сопротивления, хотя наиболее часто используется характеристическое сопротивление 50 Ом.

При относительно простой графической конструкции, КД устанавливает прямое соответствие между нормализованным сопротивлением (или нормализованной проводимостью) и соответствующим комплексным коэффициентом отражения по напряжению. На КД можно отобразить все многообразие режимов нагруженной однородной линии передачи без потерь. Использование КД и интерпретация результатов, полученных с ее применением, требует хорошего понимания теории цепей переменного тока и теории эквивалентных линий передачи.

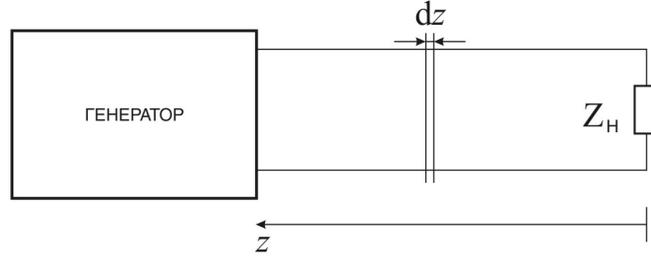
Введение круговой диаграммы

Текущий коэффициент отражения

Выше было введено понятие комплексного коэффициента отражения (по напряжению) Γ ; это понятие относится к сечению $z = 0$ нагрузки линии и является комплексным числом (см. рис. 12.1.1). Теперь нам понадобится обобщить это понятие, введя *текущий коэффициент отражения* $\Gamma(z)$ так, чтобы «старый» коэффициент отражения равнялся $\Gamma = \Gamma(0)$. Комплексная амплитуда напряжения падающей волны в сечении эквивалентной линии передачи (ЛП) с координатой z (рис. 12.1.1): $\dot{U}_{\text{пад}}(z) = \dot{U}_{\text{пад}}(0)e^{j\beta z}$, а отраженной волны: $\dot{U}_{\text{отр}}(z) = \Gamma \dot{U}_{\text{пад}}(0)e^{-j\beta z}$. Текущим коэффициентом отражения в сечении z называется отношение комплексных амплитуд напряжения отраженной и падающей волн в этом сечении:

$$\Gamma(z) = \frac{\dot{U}_{\text{отр}}(z)}{\dot{U}_{\text{пад}}(z)} = \Gamma(0)e^{-i2\beta z}. \quad (12.1.1)$$

Поскольку длина волны в линии равна $\lambda_{\text{л}} = 2\pi/\beta$, (12.1.1) можно переписать в виде $\Gamma(z) = \Gamma(0)\exp(-i \cdot 4\pi z/\lambda_{\text{л}})$, откуда видно, что по мере увеличения z $\Gamma(z)$ описывает окружности на комплексной плоскости с радиусом $\rho = |\Gamma|$ и периодом $\lambda_{\text{л}}/2$. Иногда удобно ввести нормированное расстояние $z' = 2z/\lambda_{\text{л}}$.



Координата в линии передачи

Можно выразить напряжение и ток в линии через текущий коэффициент отражения:

$$\dot{U}(z) = \dot{U}_{\text{пад}}(z)[1 + \Gamma(z)]; \quad \dot{I}(z) = \frac{\dot{U}_{\text{пад}}(z)}{Z_{\text{в}}} [1 - \Gamma(z)], \quad (12.1.2)$$

где $Z_{\text{в}}$ – волновое сопротивление линии. Деля напряжение в сечении z на ток в том же сечении, получаем входное сопротивление как функцию координаты z :

$$Z_{\text{вх}}(z) = Z_{\text{в}} \frac{1 + \Gamma(z)}{1 - \Gamma(z)}. \quad (12.1.3)$$

Удобно пользоваться *нормированным входным сопротивлением*:

$$z_{\text{вх}}(z) = \frac{1 + \Gamma(z)}{1 - \Gamma(z)}, \quad (12.1.4)$$

и наоборот:

$$\Gamma(z) = \frac{z_{\text{вх}} - 1}{z_{\text{вх}} + 1}. \quad (12.1.5)$$

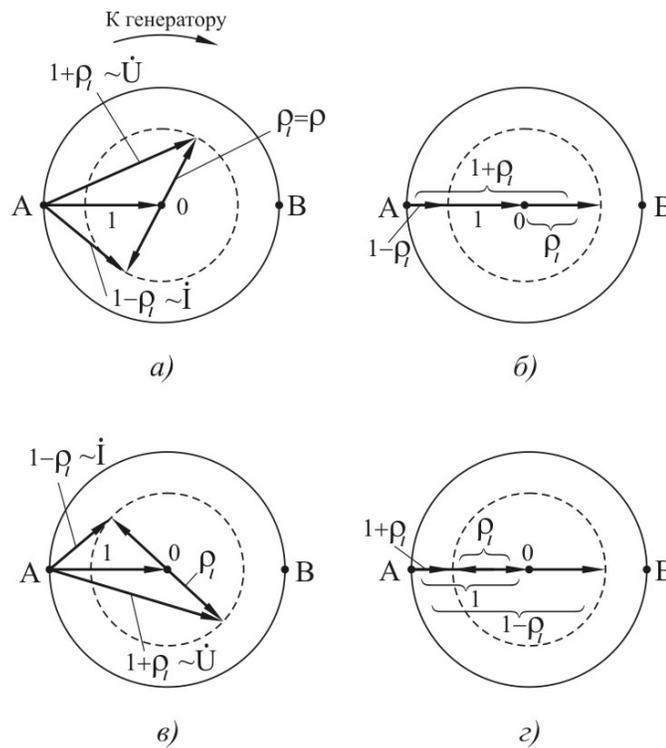
В силу периодичности $\Gamma(z)$ $z_{\text{вх}}$ также периодически с тем же периодом $\lambda_{\text{л}}/2$.

С общематематической точки зрения уравнения (12.1.4), (12.1.5) представляют собой преобразования Мёбиуса, связывающие функции $\Gamma(z)$ и $z_{\text{вх}}(z)$; с позиций теории функций комплексного переменного это частный случай дробно-линейных преобразований. Подобные соотношения часто встречаются в прикладной электродинамике и теории це-

пей. Круговая диаграмма как раз и предназначена для наглядного выполнения и интерпретации этих преобразований. Функции $\Gamma(z)$ и $z_{\text{вх}}(z)$ параметрически зависят от частоты.

Векторная диаграмма падающей и отраженной волн

Идея КД состоит в том, чтобы текущий коэффициент отражения $\Gamma(z)$ выражался своими модулем и углом в полярной диаграмме на комплексной плоскости, а нормированное сопротивление $z_{\text{вх}}$ – в декартовой системе координат на той же плоскости. Так как $\rho = |\Gamma| \leq 1$, то на конкретной частоте и в конкретном сечении z линии $\Gamma(z; \omega)$ будет выражен точкой внутри круга единичного радиуса, а на совокупности частот в некоторой полосе Π $\Gamma(z; \omega \in \Pi)$ отобразится некоторой траекторией внутри этого круга. С другой стороны, как следует из (12.1.1), на данной частоте ω с ростом z (т. е. по мере приближения к генератору) точка $\Gamma(z)$ движется по окружности радиуса ρ в направлении по часовой стрелке с текущим углом $2\beta z$. Прежде чем обратиться к собственно КД, следуя [12.3], рассмотрим векторную диаграмму полных напряжения и тока в линии в зависимости от z (рис.12.2.1).



Векторная диаграмма полных напряжения и тока в различных сечениях линии

Как следует из (12.1.2), числитель $1 + \Gamma(z)$ формулы (12.1.4) в некотором относительном масштабе пропорционален комплексной амплитуде напряжения в точке z , а знаменатель – комплексной амплитуде тока в точке z . На рис. 12.2.1 а) из т. А на комплексной плоскости проведен вектор единичной длины, соответствующий в выбранном масштабе комплексной амплитуде падающей волны при $z=0$ (для определенности фаза вектора при $z=0$ выбрана нулевой). Точка конца вектора O есть центр окружности, по которой движется конец вектора $\Gamma(z)$. Прибавляя к построенному вектору вектор $\Gamma(0)$ и век-

тор $-\Gamma(0)$, получаем в выбранном масштабе вектор $\dot{U}(0)$ и вектор $\dot{I}(0)$. При выбранном для построения диаграммы направлении $\Gamma(0)$ напряжение опережает по фазе ток, значит, тем самым был выбран индуктивный характер нагрузки.

Будем увеличивать z до тех пор, пока вектор $\Gamma(z)$ не примет направления, показанного на рис. 12.2.1 б), т. е. станет вещественным: $\Gamma(z) = \rho$. В этом сечении линии напряжения падающей и отраженной волн сложились синфазно, т. е. имеется пучность напряжения, а токи падающей и отраженной волн – противофазно, т. е. имеется узел тока. Как видим, нормированное входное сопротивление нагруженной линии такой длины (см. (12.1.4)):

$$z_{\text{вх}}(z) = \frac{1 + \Gamma(z)}{1 - \Gamma(z)} = \frac{1 + \rho}{1 - \rho} = k_{\text{св}}, \quad (12.2.1)$$

т. е. вещественно и равно коэффициенту стоячей волны, как и должно быть в пучности напряжения.

Увеличиваем далее z , пока конфигурация векторов не примет вид, похожий на диаграмму рис. 12.2.1 в). Теперь ток опережает по фазе напряжение, т. е. входное сопротивление носит емкостной характер. Увеличиваем далее z до тех пор, пока вектор $\Gamma(z)$ не примет направления, противоположного действительной оси (рис. 12.2.1 г)), т. е. станет вещественным отрицательным: $\Gamma(z) = -\rho$. При этом напряжения падающей и отраженной волн сложились противофазно, образуя узел напряжения, а токи падающей и отраженной волн сложились синфазно, образуя пучность тока. Нормированное входное сопротивление нагруженной линии такой длины (см. (12.1.4)):

$$z_{\text{вх}}(z) = \frac{1 + \Gamma(z)}{1 - \Gamma(z)} = \frac{1 - \rho}{1 + \rho} = 1/k_{\text{св}}. \quad (12.2.2)$$

Увеличиваем далее z до тех пор, пока не повторится конфигурация, показанная на рис. 12.2.1 а), цикл закончился, ситуация повторилась, благодаря периодичности напряжения, тока и входного сопротивления с периодом $\lambda_{\text{д}}/2 = \pi/\beta$ по оси z и полной окружности по вращению $\Gamma(z)$. Таким образом, на рис. 12.2.1 приведены состояния в пяти сечениях линии:

$z_1 = 0$: входное сопротивление равно сопротивлению нагрузки, выбранному как комплексное индуктивного характера;

z_2 : входное сопротивление вещественно, равно КСВ;

z_3 : входное сопротивление комплексное емкостного характера;

z_4 : входное сопротивление вещественно, равно 1/КСВ;

z_5 : повторяется конфигурация при z_1 .

Некоторые элементы векторной диаграммы имеют специальные названия. Если отражение от нагрузки отсутствует: $\Gamma(z) = \Gamma(0) = 0$, т. е. имеется полное согласование, то концы обоих векторов \dot{U} , \dot{I} в выбранном масштабе находятся в точке O , поэтому эту точку называют *точкой согласования*. Если линия нагружена на КЗ, то $\Gamma(0) = -1$ и при $z = 0$ конец вектора \dot{U} находится в точке A , т. е. $\dot{U} = 0$, $Z_{\text{вх}}(0) = Z_{\text{н}} = 0$, поэтому точку A называют *точкой короткого замыкания*. Если линия нагружена на ХХ, то $\Gamma(0) = 1$ и при

$z = 0$ конец вектора \dot{U} находится в точке B , а конец вектора \dot{I} – в точке A , т. е. $\dot{I} = 0$, $Z_{\text{вх}}(0) = Z_{\text{н}} = \infty$, поэтому точку B называют *точкой холостого хода*. Входное сопротивление при некотором z будет активным, если векторы \dot{U} и \dot{I} находятся на одной прямой, а такой прямой может быть только AB . Поэтому отрезок AB называется *линией активных входных сопротивлений*. Если концы векторов \dot{U} , \dot{I} оба находятся на окружности единичного радиуса с центром в точке O , то угол между этими векторами составляет 90° , т. к. вершина угла A находится на окружности и он опирается на ее диаметр. В этом случае входное сопротивление чисто реактивно, поэтому указанная окружность называется *окружностью чисто реактивных входных сопротивлений*. Пусть нагрузка линии неизменна, а z возрастает от нуля в сторону генератора. Тогда, как мы видели на рис. 12.2.1, концы векторов \dot{U} , \dot{I} движутся по окружности радиуса ρ с центром в точке O , а

$k_{\text{св}} = \frac{1+\rho}{1-\rho}$ остается постоянным. Поэтому любая окружность с центром в точке O и радиусом, меньшим единицы, называется *окружностью постоянного КСВ*.

Построение круговой диаграммы

Круговая диаграмма интерпретирует соотношения (12.1.4), (12.1.5), т. е. позволяет геометрически наглядно выразить эти дробно-линейные преобразования. Более конкретно, (12.1.4) преобразует комплексную плоскость текущего коэффициента отражения $\Gamma(z) = \xi + i\eta$ в комплексную плоскость нормированного входного сопротивления $z_{\text{вх}} = r_{\text{вх}} + ix_{\text{вх}}$, а (12.1.5) осуществляет обратное преобразование, причем эти преобразования взаимно-однозначны. Кроме того, благодаря свойствам дробно-линейных преобразований, преобразования (12.1.4), (12.1.5) переводят окружность в окружность и сохраняют углы между пересекающимися кривыми.

Пусть на плоскости $z_{\text{вх}}$ нанесена координатная сетка ортогональной прямолинейной системы координат, состоящая из двух взаимно-ортогональных семейств прямых $r_{\text{вх}} = \text{const}$ и $x_{\text{вх}} = \text{const}$. На плоскости текущего коэффициента отражения $\Gamma(z)$ эти семейства перейдут в два взаимно-ортогональных семейства окружностей, образующих координатную сетку ортогональной криволинейной системы координат. Круговой диаграммой полных нормированных сопротивлений называется векторная диаграмма на плоскости текущего коэффициента отражения $\Gamma(z)$, снабженная указанной координатной сеткой. Найдем уравнения, а тем самым – положения координатных окружностей.

Выразим правую и левую части равенства (12.1.4) через действительные и мнимые части величин $z_{\text{вх}}$, $\Gamma(z)$ соответственно:

$$r_{\text{вх}} + ix_{\text{вх}} = \frac{(1+\xi) + i\eta}{(1-\xi) - i\eta}, \quad (12.3.1)$$

и приравняем по отдельности действительные и мнимые части:

$$r_{\text{вх}} = \frac{1-\xi^2-\eta^2}{(1-\xi)^2+\eta^2}; \quad (12.3.2)$$

$$x_{\text{вх}} = \frac{2\eta}{(1-\xi)^2+\eta^2}. \quad (12.3.3)$$

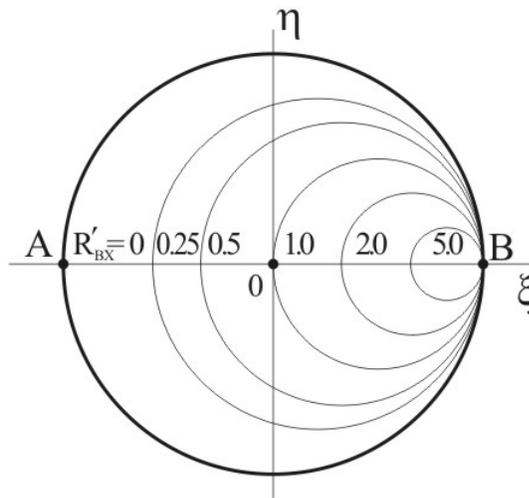
Начнем с семейства окружностей постоянной активной части входного сопротивления: $r_{\text{вх}} = \text{const}$, уравнение каждой такой окружности дается (12.3.2), если считать $r_{\text{вх}}$ постоянной. Как известно, уравнение окружности в декартовой системе координат (ξ, η) имеет вид

$$(\xi - a)^2 + (\eta - b)^2 = R^2, \quad (12.3.4)$$

где (a, b) – центр окружности, R – ее радиус. Представив уравнение (12.3.2) в виде

$$\left(\xi - \frac{r_{\text{вх}}}{1 + r_{\text{вх}}}\right)^2 + \eta^2 = \frac{1}{(1 + r_{\text{вх}})^2}, \quad (12.3.5)$$

и сравнивая его с (12.3.4), находим, что на комплексной плоскости $\xi + i\eta$ центр окружности постоянного $r_{\text{вх}}$ находится в точке $\{\xi = \frac{r_{\text{вх}}}{1 + r_{\text{вх}}}, \eta = 0\}$, а ее радиус равен $1/(1 + |r_{\text{вх}}|)$.



Семейство окружностей постоянной активной части входного сопротивления:

$$r_{\text{вх}} = \text{const}.$$

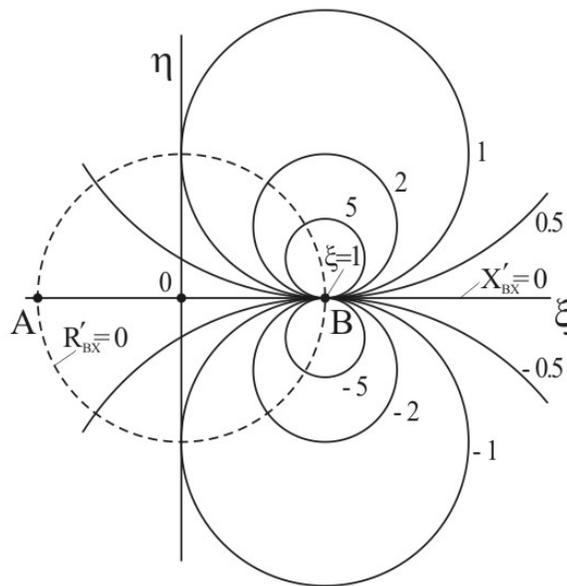
На рис. 12.3.1 показаны некоторые из этих окружностей, соответствующие положительным значениям $r_{\text{вх}}$ (для пассивных цепей активная часть сопротивления неотрицательна). На действительной оси нанесены точки A, O, B из векторной диаграммы рис. 12.2.1, что позволяет совместить круговую и векторную диаграммы и оправдать данное выше определение круговой диаграммы. Как видно из рисунка, окружности постоянных $r_{\text{вх}} \geq 0$ целиком лежат внутри круга единичного радиуса. Как отмечено выше, окружности постоянных $\rho = |\Gamma(z)|$ также лежат внутри единичного круга, но их центры лежат в точке $(0, 0)$. Таким образом, дробно-линейное преобразование (12.1.5) отображает правую полу-

плоскость комплексной переменной $z_{\text{вх}}$ на внутреннюю область единичного круга $|\Gamma(z)| \leq 1$, лежащего в плоскости комплексной переменной $\Gamma(z)$. Все окружности постоянного $r_{\text{вх}}$ касаются друг друга в точке B холостого хода с координатами $(\xi = 1, \eta = 0)$. Точки центров окружностей постоянного $r_{\text{вх}}$ сгущаются к точке B при $r_{\text{вх}} \rightarrow \infty$, что ограничивает точность графических построений при $r_{\text{вх}} \square 1$.

Перейдем к окружностям постоянной реактивной части входного сопротивления: $x_{\text{вх}} = \text{const}$, уравнение каждой такой окружности дается (12.3.3), если считать $x_{\text{вх}}$ постоянной. Представив уравнение (12.3.3) в виде

$$(\xi - 1)^2 + \left(\eta - \frac{1}{x_{\text{вх}}}\right)^2 = \frac{1}{x_{\text{вх}}^2},$$

и сравнивая его с (12.3.4), находим, что на комплексной плоскости $\xi + i\eta$ центр окружности постоянного $x_{\text{вх}}$ находится в точке $\{\xi = 1, \eta = \frac{1}{x_{\text{вх}}}\}$, а ее радиус равен $1/|x_{\text{вх}}|$. При $x_{\text{вх}} > 0$ (индуктивный характер входного сопротивления) центры находятся в верхней полуплоскости, при $x_{\text{вх}} < 0$ (емкостной характер входного сопротивления) – в нижней полуплоскости.



Семейство окружностей постоянной реактивной части входного сопротивления

На рис. 12.3.2 показаны некоторые из окружностей постоянной реактивной части входного сопротивления. Как и на рис. 12.3.1, на действительной оси нанесены точки A, O, B и проведена пунктиром окружность единичного радиуса с центром в точке $(\xi = 0, \eta = 0)$; эта окружность ограничивает область наших интересов на плоскости $\xi + i\eta$. Как видно из рис. 12.3.2, не только центры окружностей, но и сами окружности целиком лежат в одной полуплоскости: при $x_{\text{вх}} > 0$ – в верхней, при $x_{\text{вх}} < 0$ – в нижней. Все окружно-

сти постоянного $x_{вх}$ касаются друг друга в точке B . Точки центров окружностей постоянного $x_{вх}$ сгущаются к точке B при $x_{вх} \rightarrow \infty$, что ограничивает точность графических построений при $x_{вх} \square 1$.

Сравнивая выражения для радиусов окружностей постоянного $r_{вх}$ (радиус $R = 1/(1 + |r_{вх}|)$) и постоянного $x_{вх}$ (радиус $R = 1/|x_{вх}|$), видим, что первый радиус стремится к 1 при $|r_{вх}| \rightarrow 0$, а второй неограниченно и быстро возрастает при $|x_{вх}| \rightarrow 0$, в частности, при $x_{вх} = 0$ радиус равен бесконечности и окружность превратилась в прямую $\eta = 0$. В результате с уменьшением $|x_{вх}|$ все меньшая относительная часть окружностей постоянного $x_{вх}$ попадает в круг единичного радиуса с центром в точке $(\xi = 0, \eta = 0)$.

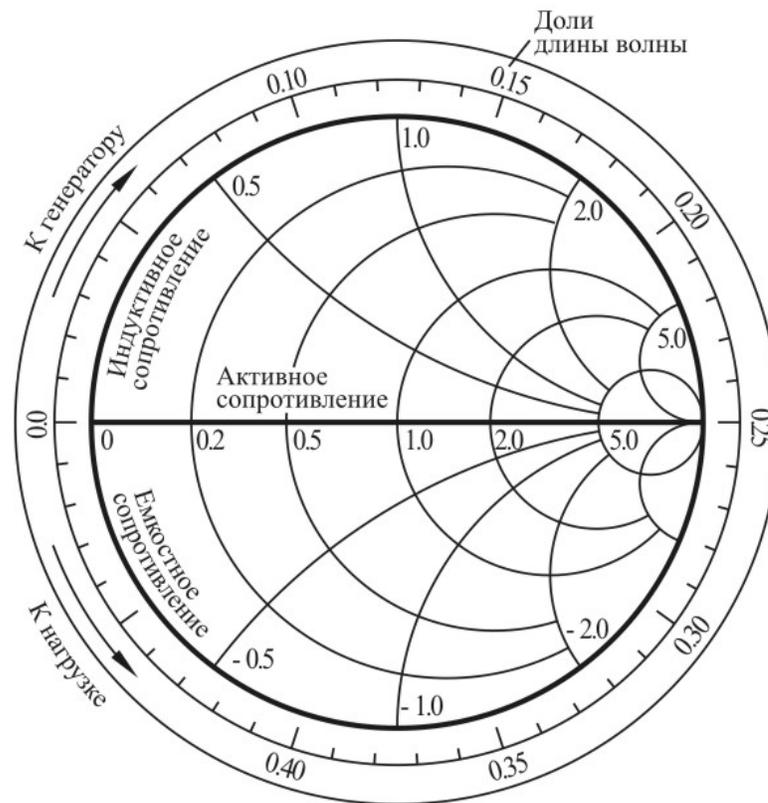


Диаграмма Вольперта-Смита с укрупненными шагами.

Чертеж круговой диаграммы показан на рис. 12.3.3. Для удобства изучения шаги по $r_{вх}, |x_{вх}|, \lambda_{л}$ взяты более крупными, чем обычно берут в рабочих диаграммах. Как видим, диаграмма расположена внутри круга единичного радиуса, оси ξ, η опущены, нанесены (с определенным шагом) окружности постоянной активной части $r_{вх}$ входного сопротивления и попадающие в диаграмму части окружностей постоянной реактивной части $x_{вх}$ входного сопротивления, по границе круга единичного радиуса в долях $\lambda_{л}$ нанесен фазо-

вый угол $4\pi z / \lambda_n = 2\beta z$ текущего коэффициента отражения (индекс «л» в обозначении длины волны опущен), показаны направления движения по окружности «к генератору» и «к нагрузке». Надпись «Активное сопротивление» напоминает, что горизонтальная ось диаграммы есть вырожденная в прямую окружность постоянной $x_{\text{вх}} = 0$, т. е. геометрическое место точек чисто активных входных сопротивлений. Надписи «Индуктивное сопротивление» и «Емкостное сопротивление» напоминают, что окружности постоянных положительных $x_{\text{вх}}$ расположены в верхней полуплоскости, постоянных отрицательных $x_{\text{вх}}$ – в нижней полуплоскости $\xi + i\eta$. Диаграмма, рассчитанная на многократное применение, иногда изготавливается как отдельное графическое устройство, в котором, кроме собственно круговой диаграммы имеется прозрачный вращающийся вокруг центра диаграммы движок с нанесенной на нем шкалой КСВ или модуля ρ коэффициента отражения. Если диаграмма не снабжена движком, то шкалы КСВ и ρ показаны на горизонтальной прямой, расположенной под диаграммой.

Отметим, что нанесенная на диаграмме координатная сетка в виде двух семейств окружностей позволяет отобразить на диаграмме все возможные значения полного нормированного сопротивления $z_{\text{вх}} = r_{\text{вх}} + ix_{\text{вх}}$. Последнее отображается в виде точки пересечения двух окружностей: одна есть окружность постоянного заданного $r_{\text{вх}}$, другая – окружность постоянного заданного $x_{\text{вх}}$. Наоборот, каждая точка диаграммы соответствует некоторому $z_{\text{вх}}$. В то же время, каждая точка диаграммы соответствует некоторому текущему комплексному коэффициенту отражения $\Gamma(z) = \rho \angle \varphi_{\Gamma}$: модуль ρ коэффициента отражения равен радиусу окружности с центром ($\xi = 0$, $\eta = 0$), на которой лежит точка (т. е. расстоянию от точки до центра диаграммы); φ_{Γ} равен углу вектора, проведенного в точку из центра диаграммы. Иногда на внешней окружности диаграммы, кроме шкалы в долях λ_n , наносится азимутальная шкала в радианах (от $-\pi$ до $+\pi$) или в градусах (от -180^0 до $+180^0$), позволяющая определить φ_{Γ} . В этом случае для отсчета φ_{Γ} следует провести прямую из центра диаграммы через данную точку до пересечения с азимутальной шкалой, и по ней отсчитать φ_{Γ} .

Заметим также, что изображающая точка T_y нормированной входной проводимости $y_{\text{вх}} = 1/z_{\text{вх}}$ симметрична изображающей точке T_z нормированного входного сопротивления $z_{\text{вх}}$ относительно центра диаграммы. Действительно, поскольку T_z суть конец вектора $\Gamma(z) = \xi + i\eta$ и

$$z_{\text{вх}} = [1 + \Gamma(z)] / [1 - \Gamma(z)] = (1 + \xi + i\eta) / (1 - \xi - i\eta),$$

то конец вектора $-\xi - i\eta$ суть изображающая точка комплекса

$$(1 - \xi - i\eta) / (1 + \xi + i\eta) = [1 - \Gamma(z)] / [1 + \Gamma(z)] = 1 / z_{\text{вх}} = y_{\text{вх}}.$$

Примеры применения круговой диаграммы

А) Определение модуля коэффициента отражения и КСВ по заданному входному сопротивлению. Пусть задано входное сопротивление $Z_{\text{вх}} = R_{\text{вх}} + iX_{\text{вх}}$ и активное волновое сопротивление $R_{\text{в}}$. Делим $Z_{\text{вх}}$ на $R_{\text{в}}$, получая нормированное входное сопротивление $z_{\text{вх}} = r_{\text{вх}} + ix_{\text{вх}}$, где $r_{\text{вх}} = R_{\text{вх}} / R_{\text{в}}$, $x_{\text{вх}} = X_{\text{вх}} / R_{\text{в}}$. Выделяем окружности постоянного $r_{\text{вх}}$ и постоянного $x_{\text{вх}}$, на их пересечении находится точка T . Модуль ρ коэффициента отражения находится как расстояние точки T от центра диаграммы в выбранном масштабе; его и

$k_{\text{СВ}}$ можно определить по шкале на движке или по шкале, нанесенной на поддиаграммной прямой.

Б) Определение модуля и фазы коэффициента отражения и КСВ по заданному сопротивлению нагрузки. Пусть задано сопротивление нагрузки $Z_{\text{н}} = R_{\text{н}} + iX_{\text{н}}$ и волновое сопротивление $R_{\text{в}}$. Находим нормированное сопротивление нагрузки $z_{\text{н}} = r_{\text{н}} + ix_{\text{н}}$, далее действуем как в п. А). Кроме ρ и $k_{\text{СВ}}$, можно определить $\varphi_{\text{н}} = \varphi_{\Gamma}$, беря отсчет по азимутальной шкале на внешней окружности диаграммы. Если необходимо знать действительную и мнимую части коэффициента отражения $\Gamma(0) = \xi + i\eta$, их отсчет можно взять на осях декартовой системы (ξ, η) .

В) Определение входного сопротивления нагруженной линии заданной длины. Пусть заданы сопротивление нагрузки $Z_{\text{н}} = R_{\text{н}} + iX_{\text{н}}$, волновое сопротивление $R_{\text{в}}$ и длина линии z . Как в п. Б), находим нормированное сопротивление нагрузки $z_{\text{н}} = r_{\text{н}} + ix_{\text{н}}$, изображающую $z_{\text{н}}$ точку T на диаграмме, вектор $\Gamma(0)$ с концом в точке T . Используя азимутальную шкалу на внешней окружности диаграммы, поворачиваем вектор $\Gamma(0)$ по часовой стрелке на $2\beta z = 4\pi z / \lambda_{\text{л}}$ рад и выделяем окружности постоянного $r_{\text{вх}}$ и постоянного $x_{\text{вх}}$, тем самым находя $z_{\text{вх}} = r_{\text{вх}} + ix_{\text{вх}}$.

Г) Определение расстояний от нагрузки до ближайших пучности и узла напряжения. Пусть заданы $Z_{\text{н}}$ и $R_{\text{в}}$. Найдя точку, изображающую $z_{\text{н}}$, поворачиваем вектор $\Gamma(0)$ с концом в этой точке по часовой стрелке до совпадения с отрезком OA (угол поворота φ_A) и до совпадения с отрезком OB (угол поворота φ_B). Расстояние от нагрузки до ближайшего узла напряжения равно $z_{\text{у}} = \varphi_A \lambda_{\text{л}} / 4\pi$, а до ближайшей пучности напряжения – $z_{\text{п}} = \varphi_B \lambda_{\text{л}} / 4\pi$.

Д) Определение сопротивления нагрузки с помощью измерительной линии. Пусть с помощью измерительной линии измерено распределение амплитуды стоячей волны напряжения вдоль линии, в результате чего известны напряжение в пучности U_{max} , напряжение в узле U_{min} , вид ближайшего к нагрузке экстремума (пучность или узел) и расстояние z_0 от него до нагрузки. Вычисляем модуль ρ коэффициента отражения по формуле:

$$\rho = \frac{U_{\text{max}} - U_{\text{min}}}{U_{\text{max}} + U_{\text{min}}},$$

тем самым выделяем окружность радиуса ρ с центром в начале $(0, 0)$ диаграммы. Находим изображающую точку T сдвигом по этой окружности против часовой стрелки на угол $\varphi = 4\pi z_0 / \lambda_{\text{л}}$: от пересечения окружности с отрезком AO , если ближайший к нагрузке экстремум напряжения – узел; от пересечения окружности с отрезком OB , если ближайший к нагрузке экстремум напряжения – пучность. Активную и реактивную части $r_{\text{н}}$, $x_{\text{н}}$ нормированного сопротивления нагрузки $z_{\text{н}}$ находим по пересекающимся в т. T окружностям постоянного $r_{\text{н}}$ и постоянного $x_{\text{н}}$.

Е) Согласование с помощью сосредоточенной реактивности. Пусть линия передачи нагружена на сопротивление $Z_{\text{н}} \neq R_{\text{в}}$, т. е. не согласована. Одним из способов согласования линии является последовательное или параллельное включение в нее сосредоточенной реактивности на некотором расстоянии от нагрузки. С помощью диаграммы необходимо определить как алгебраическую величину $i \cdot X_{\text{согл}}$ включаемой реактивности, так и расстоя-

ние $z_{\text{согл}}$ от места ее включения до нагрузки. Для решения задачи важно заметить, что в результате согласования во всех сечениях слева от включенной реактивности нормированное входное сопротивление должно равняться $z_{\text{вх. испр}} = 1$. Решение таково: отмечаем изображающую точку T нормированного сопротивления нагрузки $z_{\text{н}} = Z_{\text{н}} / R_{\text{в}} = r_{\text{н}} + ix_{\text{н}}$ (см. пп. А), Б)). Двигаясь по часовой стрелке по окружности постоянного ρ , проходящей через T , находим две точки пересечения этой окружности с окружностью постоянного $r_{\text{вх}} = 1$. Одна точка пересечения, точка T_1 , соответствующая повороту $\varphi_1 = 4\pi z_{\text{согл1}} / \lambda_{\text{л}}$, находится в верхней полуплоскости диаграммы; другая, точка T_2 , соответствующая повороту $\varphi_2 = 4\pi z_{\text{согл2}} / \lambda_{\text{л}}$, находится в нижней полуплоскости диаграммы. Выделяем окружности постоянной реактивности входного сопротивления, проходящие через тт. T_1 , T_2 , и тем самым находим соответствующие нормированные реактивности $x_{\text{вх1}} > 0$ (индуктивная) и $x_{\text{вх2}} < 0$ (емкостная), соответствующие расстояниям от нагрузки $z_{\text{согл1}} = \varphi_1 \lambda_{\text{л}} / 4\pi$ и $z_{\text{согл2}} = \varphi_2 \lambda_{\text{л}} / 4\pi$. Для согласования нужно или компенсировать одну из этих реактивностей, или другую, т. е. или включить последовательно к линии в сечении $z_{\text{согл1}}$ реактивность $ix_{\text{согл1}} = -ix_{\text{вх1}}$ (емкость), или в сечении $z_{\text{согл2}}$ реактивность $ix_{\text{согл2}} = -ix_{\text{вх2}}$ (индуктивность). Заметим, что при выполнении одной компенсации другая происходит автоматически. Напомним также, что поскольку величины реактивностей зависят от частоты колебаний, выполненное согласование действительно только на одной частоте (той самой, на которой действительно использованное значение $x_{\text{н}}$).

Аналогичным образом, можно выполнить согласование линии с помощью параллельного включения (шунтирующей) реактивности в некотором сечении линии, для этого нужно использовать не входные сопротивления, а входные проводимости.

ИСТОЧНИКИ ЗОНДИРУЮЩИХ СИГНАЛОВ

Управляемые генераторы – синтезаторы частоты как область радиотехники

Классификация синтезаторов частоты

Современный синтезатор частот (СЧ) это устройство, которое под действием внешнего электрического управления генерирует одну (реже – несколько) частоту из обусловленного множества частот. Обычно СЧ предоставляет возможность достаточно быстрого переключения в произвольный момент времени с текущей частоты на произвольную частоту обусловленного множества. Благодаря своим замечательным свойствам, СЧ используются буквально во всех радиотехнических системах – передачи данных, радиосвязных, радиолокационных, разведывательных, радионавигационных, телевизионных, телеметрических, радиоастрономических, радиоэлектронной борьбы и активного противодействия, радиоконтроля эфира, глобального ориентирования, радиоизмерительных, метрологических частоты и времени, испытательных. Непрерывно ужесточающиеся требования приложений, с одной стороны, прогресс в цифровой технике, в построении сверхстабильных и сверхточных генераторов, сверхминиатюризация и интеграция, с другой стороны, привели к огромному разрастанию числа типов и подтипов СЧ; необходима разветвленная классификация СЧ.

Методы синтеза частот разделяются на прямые (множество частот образуется прямыми линейными и нелинейными преобразованиями опорных частот), косвенные (множество частот образуется с помощью отдельных управляемых и регулируемых генераторов, привязанных к опорным частотам) и непосредственного цифрового синтеза (монохроматические колебания с частотами обусловленного множества создаются цифровой техни-

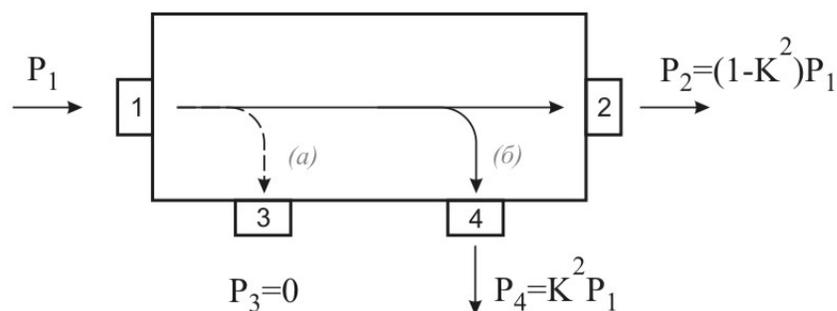
кой с помощью вычислительных алгоритмов). Участвующие в прямом синтезе фильтры могут быть аналоговыми, цифровыми, на поверхностно-акустических волнах (ПАВ) или основанными на каком-либо другом эффекте. В зависимости от числа опорных источников различают когерентный синтез (один опорный источник) и некогерентный синтез (два или несколько источников). Привязка в косвенном синтезе может производиться с помощью петель фазовой автоподстройки (ФАП), частотной автоподстройки (ЧАП), частотно-фазовой автоподстройки (ЧФАП), иных методов. Эти петли отличаются видом дискриминатора. Привязываемый генератор управляется через петлю напряжением (ГУН) с помощью варикапа, током (ГУТ) с помощью ЖИГ-резонатора, или иным способом. При косвенном синтезе петля может быть аналоговой или цифровой. В последнем случае для приведения частот опорного и управляемого генераторов к рабочей частоте дискриминатора в петле ставятся делители частоты. Они могут быть с фиксированным коэффициентом деления (ДФКД) или с переменным коэффициентом деления (ДПКД). ДПКД управляются сигналами извне и могут быть с целым коэффициентом деления или с дробным коэффициентом деления (ДДПКД). Последние предоставляют СЧ важные преимущества, однако являются источником помех дробности (ПД). Многочисленные способы борьбы с ПД разделяются на методы рационального изменения во времени, методы компенсации ПД, комбинированные методы. Методы компенсации могут быть аналоговыми, цифровыми, аналогово-цифровыми. При косвенном синтезе может быть одно преобразование частоты, два преобразования частоты, три или более преобразований частоты. Структура петель может быть различна, различают последовательные и параллельные петли; всего может быть одна петля, две петли, три или более петель. Фильтр в петле привязки может быть первого, второго или более высокого порядка. Для уменьшения ошибки слежения чаще всего применяют фильтр второго порядка с «идеальным интегратором». Аналоговая петля может включать или не включать дополнительную грубую настройку. В узкополосных петлях привязки косвенных СЧ для расширения захвата регулирования применяется или не применяется качание частоты с помощью качания напряжения в цепи управления. ГУНы и ГУТы строятся, в основном, по схеме емкостной трехточки (генератор Колпитца) на биполярном или полевом транзисторе с общей базой (общим затвором). Резонаторы управляемых генераторов строятся на сосредоточенных реактивных элементах, на отрезках линий передачи, на диэлектрике, на ЖИГ-сфере, с включением керамики или лейкосапфира. Иногда для уменьшения фазовых шумов диапазон рабочих частот управляемого генератора разбивается на поддиапазоны путем электронного подключения к резонатору разных конденсаторов с помощью р-і-п-диодов. Возможны варианты: без разбиения, два поддиапазона, три или более поддиапазонов. В цифровых петлях ФАП и ЧФАП применяются дискриминаторы типа «выборка-запоминание», на логических элементах, смешанные. В связи с различием требований к СЧ и вытекающим отсюда различием их структуры, имеет смысл разделять СЧ по диапазонам волн: метрового и более низкочастотных, дециметрового и сантиметрового, миллиметрового.

НАПРАВЛЕННЫЕ ОТВЕТВИТЕЛИ

Свойство направленного ответвления

Поскольку сигналы СВЧ проявляют волновую природу, важную роль в технике СВЧ играют *сепараторы волн*, позволяющие разделить волны, распространяющиеся в ЛП в противоположных направлениях. Например, без сепараторов невозможно измерение коэффициента отражения от нагрузки.

Для выполнения функции сепарации волн необходимо четырехпортовое устройство, обладающее свойством *направленного ответвления*, или просто *направленностью*.



Распределение потоков в НО

В технике СВЧ *направленным ответвителем* (НО) называется восьмиполюсник (четырепортовое устройство) (рис. 15.1.1), основное свойство которого таково: при подаче мощности в порт 1 она (в идеале) не поступает в порт 3 и делится в определенном отношении между портами 2, 4; если же мощность поступает в порт 2, она (в идеале) не поступает в порт 4 и делится в определенном отношении между портами 1, 3. Это свойство, с обеими его частями, можно назвать *свойством направленного ответвления*, что и дает название НО рассматриваемому устройству. Коэффициент деления мощности между выходными плечами зависит от конструкции ответвителя. Свойство направленности имеет место в собственно НО, а также в определенным образом сконструированных сверхвысокочастотных мостах Уитстона (МУ), которые мы рассмотрим отдельно. Далее слово «собственно» опускаем.

Как правило, НО строятся на основе двух близко расположенных и определенным образом связанных отрезков линий передачи, волноводных, коаксиальных или полосковых. Иногда для удобства изложения одну из этих линий называют первой (основной), другую – второй (дополнительной). Тогда свойство направленного ответвления (в идеале) можно представить так, что при передаче СВЧ-мощности по основной линии в одном направлении часть мощности ответвляется в один из портов дополнительной линии, а при передаче по основной линии в другом направлении часть мощности ответвляется в другой порт дополнительной линии.

Названия портов

Примем нумерацию портов как на рис. 15.1.1 (возможны и другие порядки нумерации). Пусть мощность P_1 от генератора подается на порт 1 (этот порт можно назвать *входным*). Мощность P_2 , прошедшая по основной линии в порт 2, меньше P_1 , так как часть P_3 мощности P_1 ответвилась в порт 3 (поэтому порт 2 можно назвать *проходным*, а порт 3 – *ответвительным*) (на самом деле, часть мощности еще ответвляется в порт 4 за счет паразитной протечки (см. далее); вследствие ее малости в энергетическом балансе ее можно не учитывать (но не в других отношениях!)). В идеале на порт 4 ответвится нулевая мощность $P_4 = 0$, как и показано на рисунке (порт 4 можно назвать *запрещенным*). В действительности же некоторая мощность $P_4 \neq 0$ ответвится и в порт 4, что, как мы увидим в дальнейшем, снижает эффективность использования НО как направленного устройства.

Если же мощность от генератора подается на порт 2 (в этом случае порт 2 – входной), то часть мощности (такая же как в предыдущем случае) ответвляется в порт 4 (этот порт является ответвительным в данном случае), часть мощности ответвляется за счет протечки в порт 3 (этот порт в данном случае является запрещенным), остальная мощность поступает в порт 1 (в данном случае – проходной). Как видим, во втором случае порт 1 поменялся местами с портом 2, мощность P_1 – с мощностью P_2 , порт 3 поменялся местами с портом 4, мощность P_3 – с мощностью P_4 , ответвленные части одинаковы, протечки одинаковы. Так проявляется взаимность НО как электродинамической системы.

С топологической точки зрения различают две конструкции НО. Первый вид, называемый *сонаправленным* НО, устроен так, что ответвленная волна во вторичной линии проходит в том же направлении, что и возбуждающая волна в первичной линии. Во втором виде, называемом *противонаправленным* НО, ответвленная волна проходит в направлении, обратном возбуждающей волне.

Консольный порт

Часто один из портов дополнительной линии не используется ни для ввода мощности, ни для вывода ее, а поэтому, в соответствии с общими принципами технической электродинамики СВЧ, оформляется как консоль, нагруженная на согласованную нагрузку (имеется в виду согласование с волновым сопротивлением дополнительной линии, если вследствие каких-то конструктивных соображений оно отличается от волнового сопротивления основной линии), и может быть изготовлен как неразъемный. В зависимости от ориентации НО и направления переноса излучения в основной линии, консольный порт может быть ответвительным или запрещенным. Иногда НО с консольным портом считается *трехпортовым НО*.

Параметры НО

Рассмотрим основные параметры реального НО, характеризующие эффективность его использования как устройства с направленными свойствами и степень его согласованности с подводящей линией. Основные параметры НО без потерь:

- Первая группа. При входном порте 1 коэффициент связи по напряжению $C_1 = \sqrt{P_3 / P_1}$, коэффициент связи по мощности $C_{1P} = C_1^2 = P_3 / P_1$, или то же в логарифмической мере, переходное ослабление $C_{1dB} = -10 \lg(P_3 / P_1)$. При входном порте 2 коэффициент связи по напряжению $C_2 = \sqrt{P_4 / P_2}$, коэффициент связи по мощности $C_{2P} = C_2^2 = P_4 / P_2$, или то же в логарифмической мере, переходное ослабление $C_{2dB} = -10 \lg(P_4 / P_2)$.
- Вторая группа. При входном порте 1 коэффициент направленности по напряжению $D_1 = \sqrt{P_3 / P_4}$, коэффициент направленности по мощности $D_{1P} = D_1^2 = P_3 / P_4$, или то же в логарифмической мере, направленность $D_{1dB} = 10 \lg(P_3 / P_4)$. При входном порте 2 коэффициент направленности по напряжению $D_2 = \sqrt{P_4 / P_3}$, коэффициент направленности по мощности $D_{2P} = D_2^2 = P_4 / P_3$, или то же в логарифмической мере, направленность $D_{2dB} = 10 \lg(P_4 / P_3)$.
- Третья группа. При входном порте 1 коэффициент развязки (просачивания) по напряжению $N_1 = \sqrt{P_1 / P_4}$, коэффициент развязки (просачивания) по мощности $N_{1P} = N_1^2 = P_1 / P_4$, или то же самое в логарифмической мере, развязка (просачивание) $N_{1dB} = 10 \lg(P_1 / P_4)$. При входном порте 2 коэффициент развязки (просачивания) по

напряжению $N_2 = \sqrt{P_2 / P_3}$, коэффициент развязки (просачивания) по мощности $N_{2P} = N_2^2 = P_2 / P_3$, или то же самое в логарифмической мере, развязка (просачивание) $N_{2dB} = 10 \lg(P_2 / P_3)$.

- Четвертая группа. Коэффициент отражения (комплексный) по порту k ($k = 1, 2, 3, 4$) Γ_k , модуль коэффициента отражения по порту k ($k = 1, 2, 3, 4$) ρ_k , или то же в логарифмической мере, возвратные потери $\rho_{k\text{db}} = -20 \lg \rho_k$ ($k = 1, 2, 3, 4$); эквивалентная по информации и по значению (но не по точности и не по смыслу) характеристика – коэффициент стоячей волны по напряжению (КСВН) K_k ($k = 1, 2, 3, 4$); он взаимно-однозначно связан с модулем коэффициента отражения:

$$K_k = \frac{1 + \rho_k}{1 - \rho_k}, \quad \rho_k = \frac{K_k - 1}{K_k + 1} \quad (k = 1, 2, 3, 4). \quad (15.3.1)$$

- Пятая группа. Допустимый диапазон рабочих частот (f_{\min}, f_{\max}), в пределах которого выполняются заданные требования к показателям всех групп.
- Иногда в число параметров НО включают частотную чувствительность ответвления: (шестая группа) Относительное изменение коэффициента связи по напряжению C_1 в заданном частотном диапазоне (f_{\min}, f_{\max})

$$\delta_{C_1} = 2 \frac{|C_1(f_{\max}) - C_1(f_{\min})|}{C_1(f_{\max}) + C_1(f_{\min})},$$

относительное изменение коэффициента связи по мощности C_{1P} , изменение переходного ослабления C_{1dB}

$$\Delta C_1 = \pm |C_{1dB}(f_{\max}) - C_{1dB}(f_{\min})|.$$

Величины C_k ($k = 1, 2$) и однозначно связанные с ними величины C_{kP}, C_{kdB} показывают, какая часть входного потока ответвляется в ответственный порт, оставшаяся часть мощности $(1 - C_k^2)P_k$ распространяется дальше. Минус в выражении для переходного ослабления C_{kdB} поставлен для удобства (теперь величина C_{kdB} положительна и уменьшение ответвляемой доли соответствует увеличению переходного ослабления). Поскольку ответвляемая доля влияет на энергетический баланс потоков в НО, в зависимости от применения НО желательна большая или меньшая ее величина. В определенной мере это зависит от конструкции НО.

Как указывалось выше, в идеальном НО ответвление в запрещенный порт отсутствует, при этом направленность бесконечна: $D_k = D_{kP} = D_{kdB} = \infty$ ($k = 1, 2$); однако такой НО создать невозможно и реальный НО имеет ограниченную, конечную направленность. Этот факт имеет фундаментальное значение для применений НО, в особенности, в измерительных системах. Подробный анализ влияния конечности направленности (но не самой направленности, как часто пишут) на ошибки измерений будет дан в дальнейшем, но качественное объяснение этого влияния можно дать уже сейчас: волна, просочившаяся в запрещенный порт дополнительной линии при одном направлении распространения волны в основной линии, складывается с волной, ответвляемой в этот порт при обратном распространении волны в основной линии, и служит источником ошибки измерения мощности

ответвляемой волны. Чем меньше направленность, тем больше соответствующая ошибка, поэтому в интересах повышения точности измерения мы заинтересованы в увеличении направленности. Достижимая направленность зависит от конструкции НО, частоты излучения, тщательности исполнения конструкции, рабочей полосы частот. Последняя зависимость носит брутальный характер: если в узкополосном применении СВЧ НО достигают направленности порядка 20-40 дБ, то с расширением рабочего диапазона частот излучения минимальная направленность в диапазоне неуклонно падает и при сверхширокой полосе может стать неприемлемо малой. Это особенно фатально при необходимости панорамно-частотных измерений в сверхширокой полосе.

Параметры третьей группы – коэффициенты развязки (просачивания) по напряжению N_k и по мощности N_{kP} , развязка (просачивание) N_{kdB} ($k = 1, 2$) – не являются математически независимыми от показателей первой и второй групп и введены для удобства. Действительно, как нетрудно видеть,

$$N_k = D_k / C_k, \quad N_{kP} = D_{kP} / C_{kP}, \quad N_{kdB} = D_{kdB} + C_{kdB} \quad (k = 1, 2). \quad (15.3.2)$$

Другая причина введения показателей третьей группы – необходимость оценки уровня протечки в запрещенный порт вне зависимости от уровня ответвления в ответственный порт.

Параметры четвертой группы – коэффициент отражения, его модуль, возвратные потери, КСВН – как известно, характеризуют степень согласования НО с подводящими линиями. Идеальный НО – полностью согласованное устройство, в реальности этого достичь невозможно. Если бы свойство симметрии эквивалентной схемы, присущее идеальному НО (см. ниже), сохранялось и в реальном НО, то соответствующие параметры четвертой группы были равны. Однако, и это, вообще говоря, не обязательно выполняется. Индекс показывает номер порта, согласование с которым характеризует параметр.

Не только параметры второй группы, но и параметры третьей и четвертой групп неуклонно «ухудшаются» с расширением рабочего диапазона частот НО (в смысле «наихудшего» значения параметра в полосе); это обстоятельство учитывается параметром пятой группы – допустимым диапазоном рабочих частот.

Идеальный НО

Идеальный НО без потерь избавлен от диссипативных потерь при распространении волн внутри него, полностью согласован с подводящими линиями, является взаимным симметричным относительно двух плоскостей восьмиполосным (четырепортовым) устройством, обладает бесконечной направленностью (нулевой протечкой в запрещенный порт).

Вследствие указанных свойств идеального НО элементы его матрицы рассеяния подчиняются соотношениям:

$$S_{11} = S_{22} = S_{33} = S_{44} = 0, \quad S_{14} = S_{41} = S_{23} = S_{32} = 0, \\ |S_{13}| = |S_{31}| = |S_{24}| = |S_{42}| = C, \quad |S_{12}| = |S_{21}| = |S_{34}| = |S_{43}| = \sqrt{1 - C^2}.$$

Соответственно, с учетом симметрии матрица рассеяния идеального НО имеет вид:

$$S = \begin{bmatrix} 0 & \sqrt{1-C^2} e^{i\varphi_{12}} & Ce^{i\varphi_{13}} & 0 \\ \sqrt{1-C^2} e^{i\varphi_{12}} & 0 & 0 & Ce^{i\varphi_{24}} \\ Ce^{i\varphi_{13}} & 0 & 0 & \sqrt{1-C^2} e^{i\varphi_{34}} \\ 0 & Ce^{i\varphi_{24}} & \sqrt{1-C^2} e^{i\varphi_{34}} & 0 \end{bmatrix} \quad (15.4.1)$$

Аргументы отличных от нуля элементов этой матрицы зависят от конструкции НО и положений плоскостей отсчета в его плечах. Специальным выбором плоскостей отсчета можно сделать $\varphi_{12} = \varphi_{13} = \varphi_{24} = \varphi_{34} = 0$, т. е. все элементы действительными.

Применения НО

Наиболее частые применения НО [15.1] можно разделить на две группы. К первой группе относятся те применения, в которых используется высокая развязка по высокой частоте между каналом главной линии НО и цепью, на которую поступает ответвленная мощность. В этих применениях направленность НО не является определяющим параметром, хотя ее величина связана с количественной мерой указанной развязки. Примерами применений этой группы является мониторинг мощности и (или) частоты генератора СВЧ, *выравнивание мощности* генератора при сканировании частоты или *стабилизация мощности* синтезатора частот при их переключении, построение матриц Батлера для фазированных антенных решеток [15.1], сложение выходных мощностей транзисторных усилителей в широком диапазоне частот [15.1], индикация мощности в фидерных трактах СВЧ радиолокаторов [15.1]. При выравнивании мощности генератора и стабилизации мощности синтезатора частот ответвленная часть мощности генератора или синтезатора подается в цепь отрицательной обратной связи, в которой детектируется, фильтруется и вырабатывает сигнал ошибки, подающийся в цепь управления генератором (синтезатором). Кроме того, НО на связанных ЛП применяются [15.2] в смесителях, переключателях, модуляторах, фазовращателях, балансных усилителях.

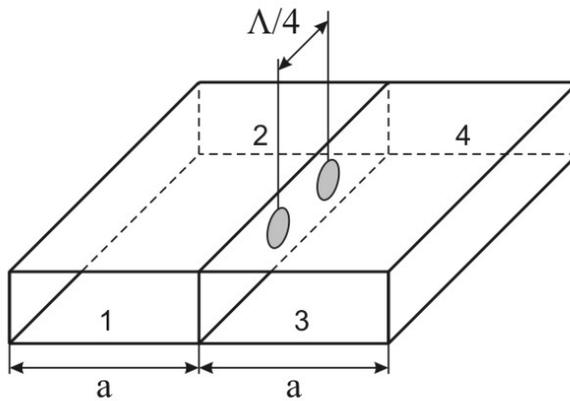
Ко второй группе относятся те применения, в которых эксплуатируется способность НО физически разделять падающий и отраженный потоки, которые затем отдельно обрабатываются. В этих применениях особенно нуждаются измерительные приборы и системы, в частности, анализаторы цепей. Заметим, кстати, что в современных анализаторах цепей НО используются и для стабилизации мощности синтезаторов, и для разделения падающих и отраженных потоков. В применениях второй группы первостепенное значение имеет направленность НО, в особенности в частотно-панорамных измерительных системах, поскольку дефицит направленности является источником существенной ошибки измерения отраженного потока.

Исторические конструкции НО

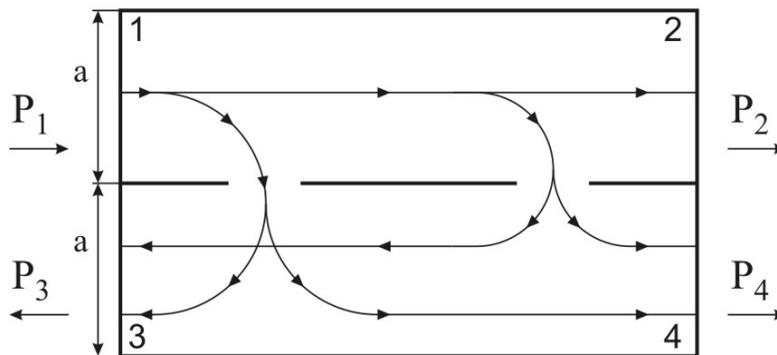
Замечательные свойства идеального НО еще в начале развития техники СВЧ вызвали энтузиазм по поводу возможностей его применения. Однако, создание НО, близких по своим свойствам к идеальным, оказалось трудной задачей, прежде всего в отношении получения высокой направленности и широкополосности. Это вызвало огромный поток работ и изобретений в области принципов создания и конструкций НО, не прекращающийся до сего времени, и, разумеется, им посвящена огромная литература, в частности, [15.1–15.37]. Тем не менее, и сейчас подбор удовлетворительного НО для конкретного применения в измерительной технике СВЧ непросто. Рассмотрим некоторые конструкции НО с целью уяснения возможностей преодоления названных трудностей.

Простейшая и исторически едва ли не первая конструкция НО представляет собой два прямоугольных волновода с общей узкой стенкой, в которой проделаны два неболь-

ших отверстия на расстоянии $\lambda/4$ друг от друга, где λ – расчетная длина распространяющейся волны (рис. 15.7.1) [15.1– 15.4, 15.7– 15.9]. Пусть по основному волноводу из плеча 1 в плечо 2 распространяется волна H_{10} . Это волна через каждое отверстие в узкой стенке создает в дополнительном волновводе по две волны H_{10} , одна из которых направляется в плечо 3, а другая – в плечо 4 (рис. 15.7.2). В плече 3 поступающие волны складываются противофазно, т. к. одна отстает от другой на $\lambda/2$, и в идеале гасят друг друга. В плече 4 поступающие волны складываются синфазно, т. к. проходят одинаковые пути. В результате ответвленная часть мощности на длине волны λ поступает в плечо 4, неответвленная – в плечо 2, в плечо 3 мощность не поступает. Если же на расчетной частоте волна H_{10} поступает из плеча 2 в плечо 1, то ответвленная мощность поступает в плечо 3 и (в идеале) не поступает в плечо 4. Переходное ослабление этого НО зависит от размеров и формы отверстий. Такой НО имеет узкий рабочий диапазон частот, т. к. при отклонении частоты от расчетной, соответствующей расстоянию между центрами отверстий в $\lambda/4$, полной компенсации волн в запрещенном плече не происходит и направленность НО ухудшается.



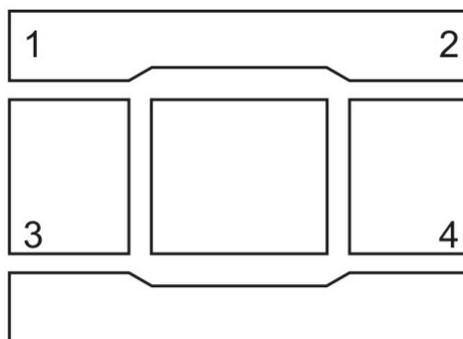
Волноводный НО с двумя отверстиями в узкой общей стенке



Объяснение направленности НО в конструкции рис. 15.7.1

Микрополосковым или симметрично-полосковым аналогом рассмотренного волноводного НО с двухдырочной связью является двухшлейфный НО (рис. 15.7.3). Роль отвер-

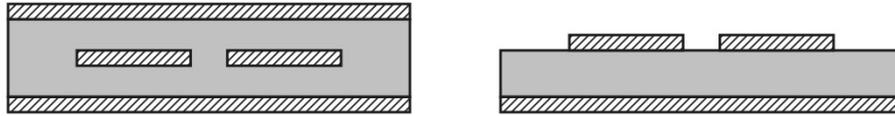
стей, соединяющих волноводы, теперь выполняют шлейфы – отрезки МПЛ или СПЛ длиной $\lambda/4$ (λ – длина волны в линии), отстоящие друг от друга также на $\lambda/4$. Подбирая волновые сопротивления шлейфов и отрезков линий передачи между ними, можно получить требуемое переходное ослабление. Как и его волноводный аналог, такой НО имеет узкий рабочий диапазон частот.



Двухшлейфный НО

Для расширения рабочей полосы частот и увеличения реализуемых значений коэффициента связи применяют многодырочные волноводные НО, для чего в общей узкой стенке двух прямоугольных волноводов проделывают n отверстий связи, отстоящих друг от друга на расстояние l , меньшее или равное четверти расчетной длины волны. Подбирая число, размеры и расстояния между отверстиями связи, можно добиться нужного коэффициента связи, большей максимальной направленности в диапазоне, большей ширины диапазона с заданной минимальной направленностью в нем, ценой сложности изготовления, больших цены, размеров и трудоемкости использования. Например, при умеренных допустимых направленностях можно добиться отношения октавного диапазона частот. Обычно отверстия связи попарно идентичны, коэффициенты связи распределены по биномиальному или чебышевскому закону, а форма и размеры отверстий определяются по полуэмпирическим формулам и графикам [15.9].

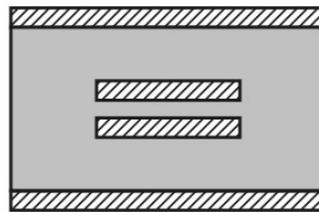
Микрополосковыми или симметрично-полосковыми аналогами волноводного НО с многодырочной связью являются многшлейфные НО [15.2]. В целом многодырочные волноводные и многшлейфные микрополосковые и симметрично-полосковые НО не удовлетворяют требованиям современных векторных анализаторов цепей в отношении направленностей и широкополосности. Значительно большие возможности при меньших сложности изготовления и эксплуатации предоставляют НО на связанных линиях передачи. В этом случае используется распределенная электромагнитная связь между близко расположенными проводниками. Она может реализоваться как боковая связь микрополосковых линий (рис. 15.7.4 а) или симметрично-полосковых линий (рис. 15.7.4 б), или как лицевая связь симметрично-полосковых линий (рис. 15.7.5). Простейшая конструкция НО на связанных МПЛ показана на рис. 15.7.6, для устранения связи между подводящими линиями использован уголкового поворота на 90° в местах соединения подводящих линий с отрезками линий передачи. Аналогичные строятся конструкции НО на связанных полосковых линиях других типов.



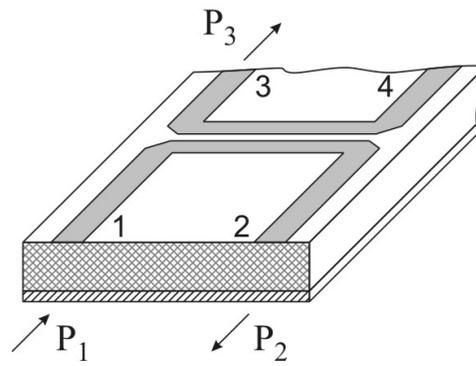
a)

б)

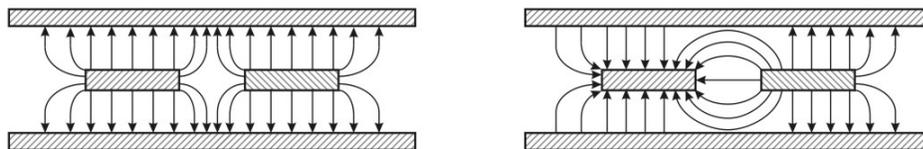
НО с боковой связью МПЛ (а) и СПЛ (б)



НО с лицевой связью СПЛ



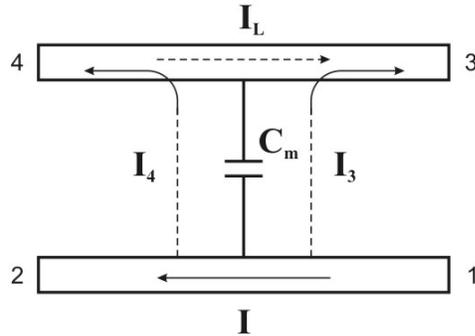
Простейшая конструкция НО на связанных МПЛ



a)

б)

Эпюры силовых линий электрического поля четной (а) и нечетной (б) волн в НО с боковой связью СПЛ



К объяснению однонаправленности в связанной линии

Направленные свойства ответвителя на связанных отрезках ЛП возникают благодаря явлению распространения и сочетания *четных* и *нечетных* нормальных T -волн. Например, в НО с боковой связью СПЛ (рис. 15.7.4) четная и нечетная волны имеют эпюры силовых линий электрического поля, показанные на рис. 15.7.7, и их сочетание обеспечивает направленное свойство ответвителя [15.27]. Хотя строгий электродинамический вывод параметров НО на связанных отрезках ЛП (в идеализированном случае) [15.2, 15.26, 15.27] непросто, саму возможность однонаправленности в связанной линии легко аргументировать на качественном уровне [15.31]. Рассмотрим четырехпортовое устройство на отрезках связанных ЛП, схематично показанное на рис. 15.7.8. Между основной (порт 1 – порт 2) и дополнительной (порт 3 – порт 4) линиями имеется как распределенная индуктивная, так и распределенная емкостная связи. Пусть в основной линии распространяется T -волна или квази- T -волна тока I от порта 1 к порту 2. За счет емкостной связи в дополнительной линии наводятся две волны тока I_3 и I_4 , распространяющиеся в противоположных направлениях, а за счет индуктивной связи – одна волна тока I_L , распространяющаяся, согласно закону электромагнитной индукции Фарадея, в направлении, противоположном волне тока I . На порте 3 волны тока I_3 и I_L синфазны, поэтому некоторая часть входной мощности волны тока I попадает в порт 3. В идеале электрические и геометрические параметры НО должны быть подобраны так, чтобы волны тока I_4 и I_L на порте 4 были противофазны, тогда они взаимно уничтожатся и никакая мощность не попадает в порт 4. В результате в идеале имеем противонаправленный НО с бесконечной направленностью.

Вернувшись к конструкции НО на связанных МПЛ, показанной на рис. 15.7.6, заметим, что она имеет более узкий диапазон рабочих частот, чем НО на СПЛ или НПЛ, и даже после экспериментальной доработки ее направленность не превышает 12...14 дБ [15.26]. Усовершенствованные конструкции НО на МПЛ описаны в [15.2]; среди них наилучшей является конструкция, известная как НО Ланге, использующий несколько связанных проводников и металлические переключки между ними; в этом НО удается добиться [15.32] практически постоянного коэффициента связи и направленности не менее 24 дБ в октавном диапазоне частот.

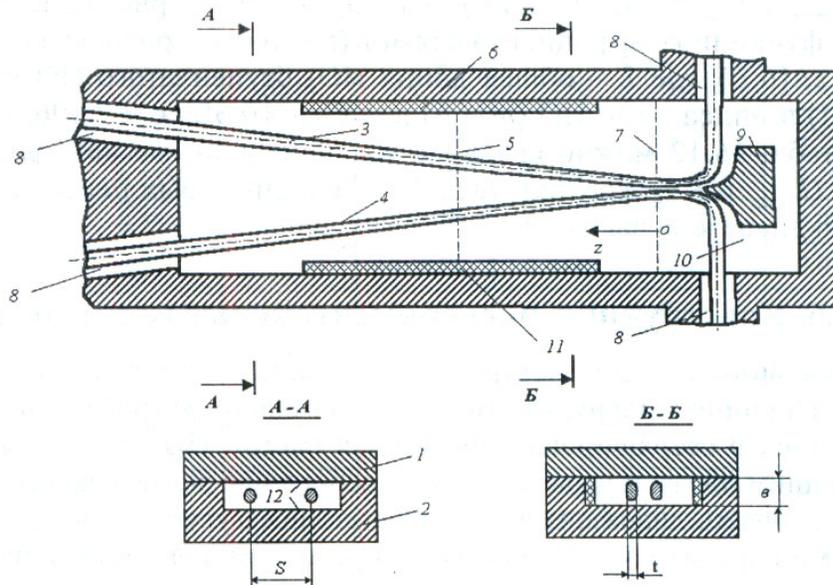
Важнейшие требования приложений к характеристикам НО СВЧ-диапазона: достаточно высокая минимальная по диапазону направленность, ширина рабочего частотного диапазона, заданное значение коэффициента связи и достаточно малая его вариабельность

в рабочем диапазоне. Неудовлетворенность конструкторов СВЧ-техники выполнением этих требований первоначальными конструкциями НО на связанных ЛП стимулировала создание в последнюю четверть века значительного числа все более сложных и критичных к технологическим допускам НО высокой направленности и сверхширокополосности. Значительное разнообразие этих конструкций позволяет составить некоторую классификацию по признакам: регулярности-нерегулярности, однородности-неоднородности связываемых линий, конфигурации линий, ступенчатости-плавности их структуры, симметричности-несимметричности конструкции относительно вертикальной плоскости.

Отечественные конструкции

Приведем несколько конструкций НО отечественной разработки на связанных плавных нерегулярных линиях передачи (НЛП) [15.17]. Сейчас общепринято, что именно использование таких ЛП создает принципиальную возможность реализации сверхширокополосных НО, наилучшим образом сочетающих выполнение вышеприведенных требований. Необходимым условием этого является стабильность диэлектрических свойств заполнения связанных линий, чему в наибольшей степени удовлетворяют высокочастотные линии с воздушным заполнением.

На рис. 15.8.1 показана конструкция унифицированного несимметричного НО на нестандартных полосковых линиях. Расстояние «в» между частями 1 и 2 корпуса в зоне проводников равно 5,54 мм, при этом волновое сопротивление одиночной полосковой линии с круглым внутренним проводником диаметром 3,04 мм равно 50 Ом. Проводники 3 и 4 основного и дополнительного каналов имеют слева от плоскости 5 постоянное круглое сечение диаметром 3,04 мм (см. сечение А-А), а затем их ширина монотонно убывает (см. сечение Б-Б) вплоть до начала изгиба проводников 7 в зоне перехода от связанных ЛП к подводящим коаксиальным линиям 8.

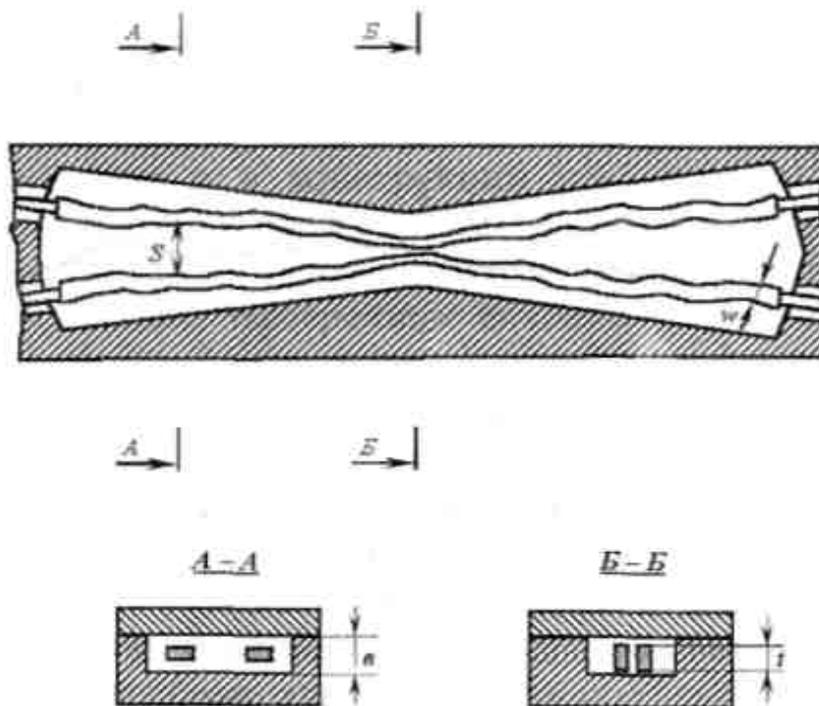


Конструкция несимметричного НО на нестандартных полосковых линиях

Оси проводников в области связи расположены под углом друг к другу; этот угол выбирается из компромисса между максимальной направленностью в рабочей полосе и шириной этой полосы; при угле менее 3 градусов достигается максимальная направленность не менее 40 дБ. Форма и положение подстроечного элемента 9 определяются из компромисса

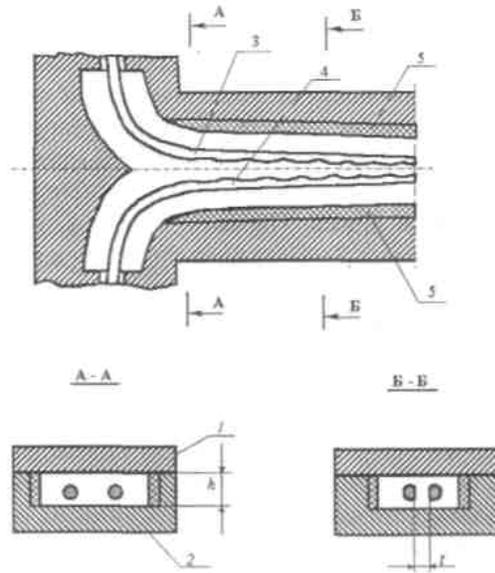
между максимальной направленностью в рабочей полосе и степенью согласования плеч НО. Для повышения направленности приняты дополнительные меры: введение в конструкцию поглощающих диэлектрических пластин 6, 11 вдоль стенок камеры связи, уменьшение сечения проводников и размеров камеры связи.

Конструкция симметричного НО [15.17] на связанных полосковых линиях с прямоугольными внутренними проводниками показана на рис. 15.8.2. В частотном диапазоне с перекрытием $f_2 / f_1 = 13$ он имеет переходное ослабление C_{1dB} равное 10 дБ или 20 дБ. Для простоты изготовления и лучшего согласования параметры линий t и b были выбраны равными, соответственно, диаметрам внутреннего и внешнего проводников подводящих коаксиальных линий. Параметры w и S при расчете линий варьировались: расстояние S между проводниками и ширина проводников w плавно увеличиваются от центра области связи ($z = l/2$) к ее краям. В качестве подводящих линий были выбраны коаксиальные линии стандартного поперечного сечения 7/3,04 мм. Отклонение переходного затухания от номинального значения в полосе 0,5...12,5 ГГц не превышает 0,3 дБ, направленность в той же полосе выше 20 дБ. Введение подстроечных элементов, положение которых выбирают экспериментально, позволяет увеличить минимальную направленность до 25 дБ в рабочем диапазоне частот 1..13 ГГц. В качестве подводящих линий были выбраны коаксиальные линии стандартного поперечного сечения 7/3,04 мм.



Конструкция симметричного НО на ППЛ

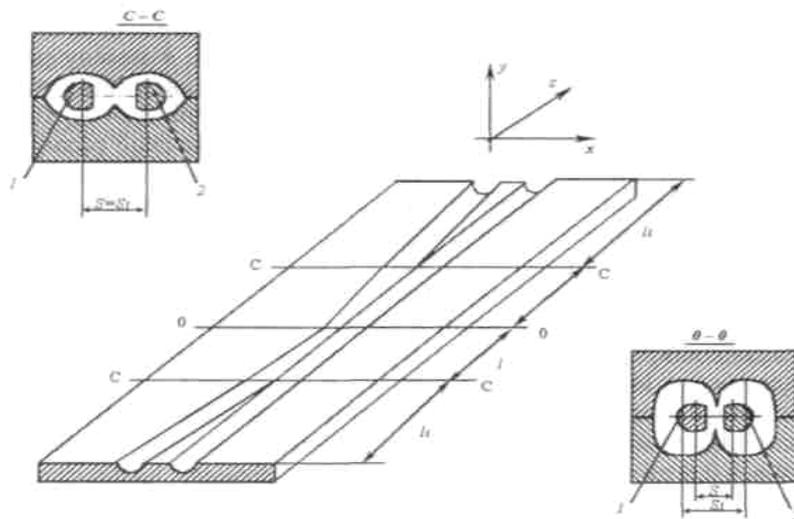
Другая конструкция симметричного НО [15.17] использует ЛП с усеченными круглыми внутренними проводниками (рис. 15.8.3).



Симметричный НО на ПЛ с усеченными внутренними проводниками

Такие линии обладают достоинствами связанных полосковых ЛП с внутренними проводниками как прямоугольной, так и круглой формы поперечного сечения. Это позволяет технологически просто и эффективно изменять значение коэффициента связи, исключая необходимость изгибания проводников, а также обеспечивает возможность хорошего согласования области связи с подводщими коаксиальными линиями с круглыми внутренними проводниками. Корпус НО состоит из двух частей 1 и 2, содержит проводники 3 и 4 линий основного и вспомогательного каналов, а также поглощающие пластины 5. Выходы НО заканчиваются присоединительными разъемами. Расстояние h между рабочими поверхностями корпуса в зоне одиночных подводщих коаксиальных линий, как и в конструкции несимметричного НО (рис. 15.8.1), постоянно и равно 5,54мм. Геометрические размеры t и h связанных полосковых линий плавно меняются вдоль области связи так, что в каждом поперечном сечении выполняются условия идеальной направленности и согласования. При изготовлении НО значения t и h контролируются в 180 точках по длине области связи. Поглощающие пластины 5 обеспечивают подавление волн высших типов, возбуждаемых на частотах выше 8 ГГц. Их введение в область связи способствует также повышению направленности НО. Измерения экспериментальных образцов в диапазоне 0,5 – 12 ГГц дали следующие результаты: переходное ослабление $C_{1dB} = 10,6$ дБ; его неравномерность $\Delta C_{1dB} = \pm 0,6$ дБ; направленность $D_{1dB} \geq 28$ дБ; КСВ $K_1 \leq 1,29$.

В [15.17] выражено мнение, что для достижения сверхширокополосным НО минимальной в рабочем диапазоне частот направленности более 30 дБ необходимо использовать связанные коаксиальные ЛП с деформированными внешними проводниками. Конструктивно такие ЛП состоят из двух половин (рис. 15.8.4). Каждую из них предложено обрабатывать фрезой с шарообразной режущей частью, погружаемой в заготовку на глубину внешнего про-



Нижняя половина корпуса НО на связанных коаксиальных ЛП с деформированными внешними проводниками

водника коаксиальной ЛП. Этот подход реализован [15.17] в серии унифицированных НО с переходными ослаблениями 10, 20, 30 дБ. Оптимальная конфигурация отрезка плавной НЛП характеризуется тремя варьируемыми размерами, которые обеспечивают технологичность и плавность конфигурации линии: деформированная коаксиальная ЛП в средней части отрезка НЛП длиной l на его краях плавно переходит в коаксиальную ЛП с круглыми проводниками. Одновременно связанные ЛП плавно переходят на границах области связи (сечение СС) в подводящие однородные ЛП. Расстояние S между осями внутренних проводников изменяется от центра области связи НО (сечение ОО) до ее краев (сечение СС) по линейному закону $S = S_{min} - 2z / l(S_{min} - D)$, где S_{min} - расстояние между осями проводников в центре области связи, z - продольная координата, отсчитываемая от центра области связи. Линейный закон изменения S задан для исключения технологически сложной операции изгибания внутренних проводников. Компенсировать изменение диаметра внутреннего проводника на центральном участке области связи можно путем варьирования расстояния S_1 между осями внешних проводников. Расстояние S между осями внутренних проводников 1 и 2 дополнительно не меняется. На центральном участке области связи $S_1 > S$; вне данного участка S_1 и S совпадают.

Геометрические размеры были оптимизированы в 180 поперечных сечениях области связи по ее длине l и в 140 поперечных сечениях подводящих (одиночных) ЛП по их длине l_1 . При этом полагалось $S_{min} = 4,00$ мм (для $S_0 = 20$ дБ); $d = 3,04$ мм, $D = 7,00$ мм. Наиболее сложные детали – внутренние проводники и корпус симметричного НО на плавной НЛП – были изготовлены на фрезерном станке с ЧПУ. При изготовлении корпуса требовалось обеспечить заданное перемещение фрезы по всем трем пространственным координатам (x, y, z) одновременно и реализовать заданные размеры в указанных точках. Верхняя половина корпуса была выполнена составной из трех частей. В результате оказался возможным в про-

цессе настройки и эксплуатации доступ к области связи без какого-либо нарушения жесткости крепления внутренних проводников относительно внешних (корпуса). В результате, поставленная создателями данного НО цель была достигнута: например, для НО с переходным ослаблением 20 дБ в диапазоне 1...15 ГГц оказалось $\Delta C_{1дБ} = \pm 0,5$ дБ, $D_{1дБ} \geq 30$ дБ, $K_1 \leq 1,3$.

НАПРАВЛЕННЫЕ ОТРАЖАТЕЛЬНЫЕ МОСТЫ

Предисловие

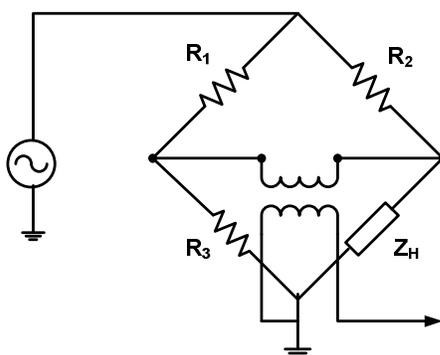
Как указано выше, в качестве сепаратора прямой и отраженной волн можно использовать или направленный ответвитель (НО), или отражательный мост (ОМ). В данной главе рассматривается второй вариант. Выбор одного из этих видов сепараторов в каждом конкретном применении диктуется рядом факторов и требований, таких как энергетическая эффективность, стоимость, влияние на ошибку измерения отраженной волны, широкополосность и т. д. Свидетельством нешаблонности этого выбора могут служить случаи, когда в одном и том же измерительном приборе в различных его частях сделан разный выбор, например, в тестовой сети скалярного анализатора цепей применены ОМ, а в схеме выравнивания мощности сканирующего генератора зондирующих сигналов (синтезатора частот) – НО. В обоих видах сепараторов важнейшим фактором, влияющим на ошибку оценки мощности отраженной волны, является размер дефицита направленности.

Начиная со знаменитого моста Уитстона, мостовые методы детектирования и измерения в различных частотных диапазонах, от постоянного тока до СВЧ, чрезвычайно развились и составляют основу ряда конструкций различного функционального назначения. В этом разнообразии ОМ представляют лишь частное, специфическое приложение.

В последние десятилетия происходило своеобразное соревнование между НО и ОМ за право реализации функции сепарации волн СВЧ. Развивались и совершенствовались оба вида; на каждом этапе этого соревнования ответ на вопрос: какой вид сепаратора предпочтителен в данной ситуации с точки зрения комплексного критерия, включающего направленность, широкополосность, согласованность и т.д., был разным. Например, долгое время почти догмой было мнение, что для обеспечения функции сепарации волн в тестовой сети скалярного анализатора цепей (САЦ) предпочтительней скалярный ОМ (датчик КСВ, отражательный мост с вшитым диодным детектором), нежели НО, т.к. первый дешевле и лучше согласован; а для тестовой сети векторного анализатора (ВАЦ) предпочтительней НО, нежели векторный ОМ, т.к. последний включает СВЧ-трансформатор, ограничивающий рабочую полосу сепаратора, и относительно дорог. Сейчас, благодаря прогрессу в технологии СВЧ-трансформаторов, по крайней мере вторая часть этого мнения часто не верна.

Датчик комплексного коэффициента отражения

В интересующем нас диапазоне частот от нескольких МГц до десятков ГГц существует две возможности построения мостового сепаратора волн.

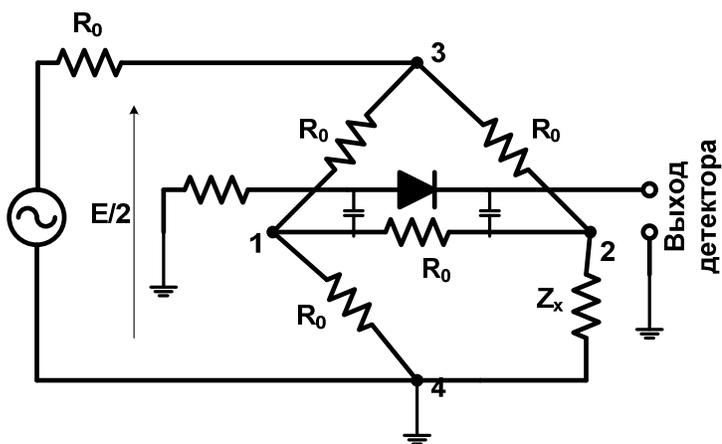


Мостовой датчик комплексного коэффициента отражения нагрузки, включенной в плечо моста

Первую возможность реализует мостовой датчик комплексного коэффициента отражения нагрузки, включенной в плечо моста (рис. 16.1.1). Физическая информация о комплексном коэффициенте отражения заложена в комплексной амплитуде колебания основной (фундаментальной) частоты f_0 , снимаемого со вторичной диагонали моста (на первичную диагональ подается волна от генератора). Обработка выходной волны совместно с ответвленным опорным колебанием позволяет определить как модуль, так и фазу искомого коэффициента отражения; поэтому такой датчик годится для использования в векторном анализаторе цепей.

Принципиальной трудностью реализации такого мостового датчика является необходимость трансформации напряжения вторичной диагонали моста в выходное напряжение, т.к. несимметричное включение этой диагонали (заземление одного из ее узлов) нарушило бы баланс моста. Требуемый трансформатор (рис. 16.1.1) как блок перехода «симметричное-несимметричное» существенно ограничивает широкополосность мостовой схемы в целом и затрудняет решение проблемы приемлемого согласования. И хотя к настоящему времени разработаны и производятся коммерческие модели таких мостовых датчиков с приемлемыми характеристиками, они все еще достаточно дороги. К счастью, для скалярного анализатора цепей такой датчик не нужен.

Датчик модуля коэффициента отражения



Мостовой датчик модуля коэффициента отражения нагрузки, включенной в плечо моста

Вторую возможность реализует мостовой датчик модуля коэффициента отражения нагрузки, включенной в плечо моста (рис. 16.2.1). Поскольку сохранение информации о фазе отраженной волны на частоте f_0 не требуется, напряжение на вторичной диагонали можно непосредственно продетектировать. В схеме этого датчика балансный детектор помещен между узлами 1 и 2, выход схемы может быть несимметричным, линия передачи соединяет порт между узлами 2 и 4 (тестовый порт) с нагрузкой, прямая волна входит в порт между 3 и 4. В отечественной литературе такой мостовой датчик иногда называют «датчиком КСВ», имея в виду, что он позволяет определить скалярную величину – модуль коэффициента отражения $|\Gamma_L| = \rho$, которую легко пересчитать в КСВ k_{CB} по формуле:

$$k_{CB} = \frac{1 + |\Gamma_L|}{1 - |\Gamma_L|} = \frac{1 + \rho}{1 - \rho}.$$

Будем считать, что внутреннее сопротивление источника напряжения E равно характеристическому сопротивлению линии R_0 . Если бы линия, подключенная к тестовому порту, имела бесконечную длину, ее входное сопротивление для входящей прямой волны было бы также равно R_0 , мостовой импеданс между точками 3 и 4 был бы согласован с внутренним сопротивлением источника R_0 , напряжение между этими точками равнялось бы $E/2$, а напряжение прямой волны на тестовом порте составляло половину этого напряжения и равнялось $E/4$ (-6 дБ относительно напряжения $E/2$). На самом деле, в общем случае линия имеет конечную длину и нагружена на несогласованное комплексное сопротивление Z_X . Комплексный коэффициент отражения от этой нагрузки равен:

$$\Gamma_L = \frac{Z_X - R_0}{Z_X + R_0},$$

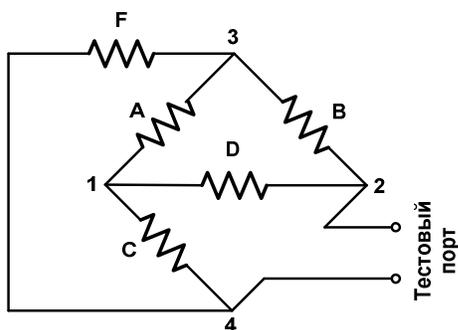
а комплексная амплитуда напряжения отраженной волны на тестовом порте равна (с точностью до фазового сдвига на длине линии):

$$V_R = \frac{E}{4} \cdot \Gamma_L.$$

Чтобы найти комплексную амплитуду E_R напряжения на входе детектора (в первом случае – на входе трансформатора), рассмотрим схему рис. 16.2.1, удалив из нее генератор и элементы детектора [16.1] (рис. 16.2.2) и введя очевидные обозначения (сопротивление в плече D – шунтирующее сопротивление детектора). Из схемы рис. 16.2.3, полученной из схемы рис. 16.2.2 с помощью теоремы Тевенина, видно, что если сопротивления в плечах C и D равны (это выполняется на рис. 16.2.1), то E_R равна половине V_R :

$$E_R = \frac{V_R}{2} = \frac{E}{8} \Gamma_L.$$

Как видим, между входной прямой волной $E/2$ и детектируемой отраженной волной образуются дополнительные потери 12 дБ.



Эквивалентная схема моста при удалении генератора и элементов детектора

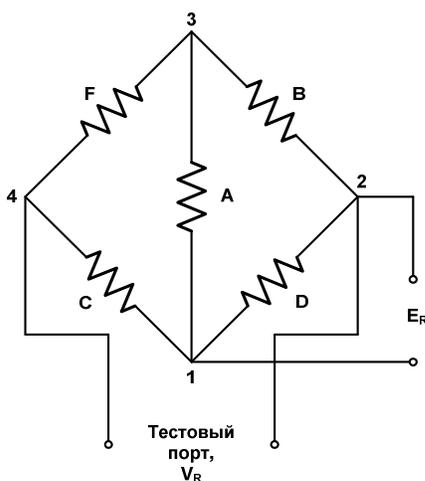
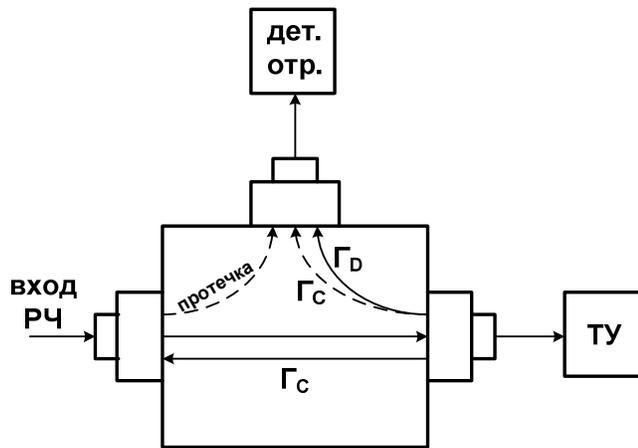


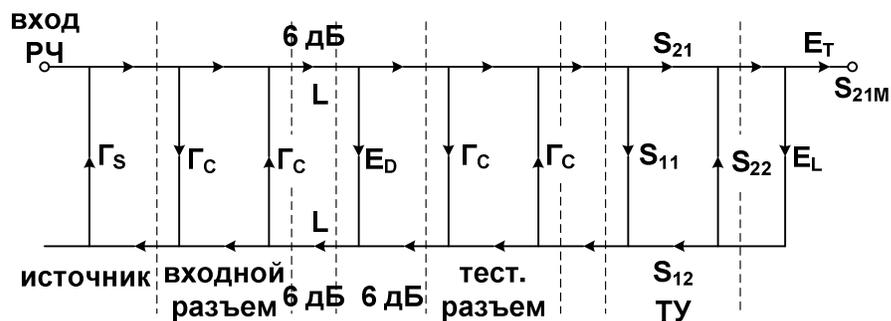
Схема после преобразования по теореме Тевенина

Более подробный анализ схем рис. 16.2.1 и рис. 16.2.2 показывает, что плечо D связано от прямой входной волны (между тт. 3-4), а плечо A (между тт. 1-3) – от отраженной волны (на детекторном порте и на входе детектора). Калиброванное сопротивление в плече C выполняет функцию внутренней опоры, определяющей характеристический импеданс моста. Конструктивно опорная и детекторная цепи интегрированы в мост, который, таким образом, становится сигнальным сепаратором со встроенным ослаблением в 6 дБ в прямом направлении и 6 дБ – в обратном направлении.

Вернемся к вопросу выбора сепаратора волн: НО или ОМ? НО не испытывает 12-децибельных двухпутевых потерь, и в этом смысле энергетически более выгоден. С другой стороны, многократные отражения между источником и тестируемым устройством, между которыми находится сепаратор в виде НО, фактически увеличивают негативный эффект дефицита направленности НО, что при очень «скромных» возможностях калибровки скалярного анализатора цепей может сделать соответствующую составляющую ошибки измерения модуля коэффициента отражения недопустимо большой. В то же время именно 12-децибельные потери в ОМ, игравшие в энергетическом аспекте негативную роль, в аспекте усиления ошибки направленности играют позитивную роль, фактически разрывая цепь многократных отражений.

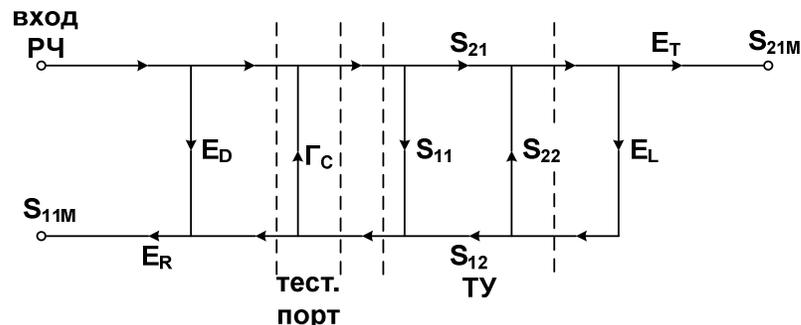


Пути прямого и отраженного сигналов в отражательном мосте



Упрощенный потоковый граф нагруженного отражательного моста

Сигнальные пути в отражательном мосте показаны на рис. 16.2.4 (Γ_c, Γ_D – коэффициенты отражения разъемов). Здесь есть два отражательных пути: один, формирующий многократную петлю с источником; другой с разъемным рассогласованием детектора отражения. Первый пренебрежим, поскольку путьевые потери по кругу составляют 12 дБ; второй также очень мал, поскольку детектор отражения неотделим от собственно моста. Упрощенный потоковый граф на рис. 16.2.5 не включает многократные отражательные эффекты между тестовым и детекторным портами, но учитывает отражения от источника во входном канале. Здесь L представляет 6-децибельные мостовые потери, E_D – сигнал, обусловленный дефицитом направленности между прямым и детекторным портами, E_T, E_L относятся к пропускательным потерям и отражению нагрузки в детекторе пропускания соответственно. ТУ представлено своими S -параметрами.



Потоковый граф после следующего упрощения

Сигнал дефицита направленности складывается на тестовом порте с отраженной волной его разъема, в результате эффективная направленность сепаратора волн ограничена качеством разъема тестового порта. Стоит заметить, что гибридная природа моста способствует некоторому погашению разъемного рассогласования путем старательного выбора импеданса в плече А, чтобы оптимально перераспределить переотражения на детекторный и тестовый порты. На рис. 16.2.6 показан граф, упрощенный за счет суммирования ветви E_D с одной из ветвей Γ_C (сумма снова обозначена E_D) и игнорирования Γ_S (в связи с чем отпадают две разнонаправленные ветви Γ_C).

АТТЕНЮАТОРЫ

Предисловие

Как ни удивительно, столь простое по замыслу устройство, как аттенюатор, находит многочисленные применения, отвечает широкой классификации, характеризуется многочисленными параметрами, выполняет многочисленные требования. Осуществляя банальную функцию ослабления, аттенюаторы находят значительное число применений: предотвращение перегрузки или выгорания, согласование импедансов, измерение потерь или усиления, увеличение изоляции, расширение динамического диапазона.

Они могут быть пассивными или активными, поглощающими или отражающими, фиксированными или переменными, шаговыми или непрерывными, регулируемые вручную или электронно, резистивными или реактивными, симметричными по импедансу или несимметричными, калибруемыми или некалибруемыми, принадлежать к различным классам по степени точности установки ослабления. Для характеристики аттенюаторов применяются такие параметры, как девиация ослабления, частотная, температурная и мощностная чувствительности, внесенные потери, входной и выходной КСВ. Данная глава ограничена пассивными аттенюаторами.

Характеристики и параметры

Аттенюатор это линейная, пассивная или активная цепь или прибор, который ослабляет электрические или микроволновые сигналы, представленные током или напряжением. Он может быть встроен в коаксиальную, полосковую или волноводную линию передачи. *Ослабление (аттенюация)* обычно выражается отношением входной мощности P_{in} к выходной мощности P_{out} в линейных величинах:

$$Atten_{lin}(A) = P_{in} / P_{out} ,$$

в децибелах:

$$Atten_{dB}(A) = 10 \lg \frac{P_{in}}{P_{out}} = 20 \lg \frac{E_{in}}{E_{out}} = 20 \lg \frac{E_1}{E_2} ,$$

или в неперах:

$$Atten_{Np}(A) = \alpha \Delta x = \ln \frac{E_1}{E_2} ,$$

где E_1, E_2 – амплитуды напряжений, α – коэффициент ослабления (Нп/м), Δx – длина ослабляющего участка линии передачи. Соотношение между неперами и децибелами:

$$1 \text{ Нп} = 8.686 \text{ дБ}.$$

Девиация ослабления – разность между действительным и номинальным ослаблениями при комнатной температуре и входной мощности 10 мВт на специальной частоте.

Частотная чувствительность – разность между максимальным и минимальным ослаблениями в заданном частотном диапазоне.

Номинальная частота – специальная частота, на которой задается точность ослабления аттенюатора.

Внесенные потери – величина потерь мощности из-за внесения аттенюатора в передающую систему. Выражается отношением доставленной мощности к той части системы, которая следует за аттенюатором, перед и после его внесения.

Характеристические внесенные потери – внесенные в линию передачи потери при отсутствии отражений в обоих направлениях от вставленного в линию аттенюатора.

Возможности управления мощностью – максимальная мощность, которая может быть подана на аттенюатор в специальных условиях и при заданной длительности без появления долговременных, выходящих за специальные границы, изменений рабочих характеристик.

Мощностная чувствительность – временное изменение ослабления (дБ/Вт) в стабильных условиях, когда входная мощность меняется от 10 мВт до максимального значения.

Стабильность ослабления – способность аттенюатора сохранять свои параметры под воздействием изменений окружающих условий.

Температурный диапазон действия – Температурный диапазон, в котором аттенюатор может действовать при максимальной входной мощности.

Температурная чувствительность – температурные изменения ослабления [дБ/(дБ×⁰С)] в температурном диапазоне действия.

Входной КСВн – коэффициент стоячей волны по напряжению на входе аттенюатора, нагруженного на сопротивление, равное характеристическому сопротивлению источника.

Выходной КСВн – коэффициент стоячей волны по напряжению на выходе аттенюатора, вход которого нагружен на сопротивление, равное характеристическому сопротивлению источника.

Применения

Существует много примеров применений, когда необходимо вести ослабление в поле, в мощность, в напряжение, в ток. Приведем некоторые из этих примеров:

- Уменьшение уровня сигнала для предотвращения перегрузки или выгорания.
- Согласование импедансов источника и нагрузки для уменьшения их взаимодействия.
- Измерение потерь или усиления двухпортового устройства.
- Увеличение изоляции между элементами цепи, или цепями путем уменьшения взаимодействия между ними.
- Расширение динамического диапазона оборудования.

Классификация

В зависимости от природы используемого элемента цепи, типа конфигурации и вида регулирования различают [17.1]:

- Пассивные и активные аттенюаторы.
- Поглощающие и отражающие аттенюаторы.
- Фиксированные и переменные аттенюаторы.

Фиксированные аттенюаторы применяют для получения постоянного ослабления. Переменные аттенюаторы создают переменное ослабление под воздействием различных регулирующих факторов. Изменение ослабления может быть шаговым или непрерывным, регулировка производится вручную или программно (электронные переменные аттенюаторы). Обычно аттенюаторы реверсируемы, за исключением некоторых типов высокоомощных аттенюаторов. Различают аттенюаторы резистивные и реактивные, симметричные (по импедансу) и несимметричные, калибруемые и некалибруемые. Неоднократно предпринимались попытки классификации аттенюаторов по степени точности ослабления в увязке с приложениями. Например, стандарт США IEEE Std 474 [17.5] различает аттенюаторы:

Класс I – Стандартные

Класс II – Прецизионные

Класс III – Общецелевые

Класс IV – Полезные

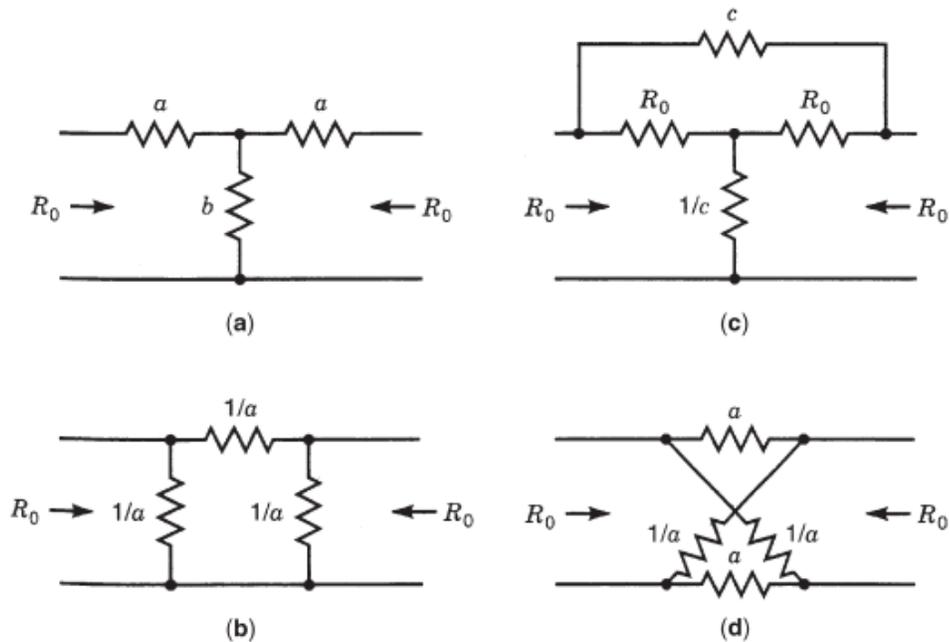
Требования

Количественные требования к аттенюаторам зависят от целей их использования. При выборе аттенюатора необходимо учитывать частотный диапазон его работы, поскольку точность установки ослабления зависит от частоты. Ослабление предполагает помещение резистивного материала в поле для его поглощения; это означает, что будет проявляться некоторое отражение. Поэтому одной из целей разработки аттенюаторов является минимизация отражения. Другой фактор – внешние потери, оцениваемые отношением уровней мощности до и после внесения поглощающего материала. Для работы переменного шагового аттенюатора важен размер шага. Вообще в список требований входят:

- Номинал ослабления.
- КСВн входа и выхода.
- Точность реализуемого ослабления.
- Номинальный уровень входной мощности.
- Размер шага (для переменного шагового аттенюатора).
- Рабочая полоса частот.
- Степень стабильности (мера изменений ослабления из-за вариаций температуры, влажности, частоты и уровня мощности).
- Характеристический импеданс на входе и на выходе.
- Повторяемость.
- Продолжительность работы и параметры надежности.
- Разрешающая способность установки ослабления (при непрерывном регулировании переменных аттенюаторов).

Фиксированные аттенюаторы

Обычно в аттенюаторах используются Т-образные, П-образные или L-образные резистивные цепи. На рис. 17.5.1 приведены примеры симметричных (входные и выходные резисторы одинаковы) конфигураций. Технология изготовления резисторов на этих схемах соответствует частотному диапазону, например, в диапазоне выше 26.5 ГГц в коаксиальном тракте используются тонкопленочные резисторы. Формулы для сопротивлений резисторов и внесенных потерь в этих схемах можно найти, например, в[17.1].



Фиксированные аттенюаторы на симметричных резистивных цепях

Фиксированные аттенюаторы, в англоязычной литературе – pad (прокладка), уменьшают мощность входного сигнала в фиксированное число раз, например, на 3, 10 или 50 дБ. Обычно точность ослабления (относительно номинального значения) и входной КСВн зависят от частоты в пределах рабочего диапазона частот. Этот вид аттенюаторов используется в многочисленных приложениях, в том числе в скалярных и векторных анализаторах цепей, с целью уменьшения уровня сигнала и (или) согласования источника с нагрузкой. Уменьшение уровня сигнала может понадобиться для приведения диодного детектора в квадратичный режим, при тестировании устройств (например, усилителей) в малосигнальной области, для удовлетворения требований мощностных сенсоров и т. п. Улучшение импедансного согласования с помощью аттенюаторов может понадобиться для уменьшения частотнозависимых вариаций сигнала, обусловленных рассогласованиями, для уменьшения частотного затягивания твердотельных источников за счет изменения высокоотражающей нагрузки и т. п.



Аттенюатор на параллельном (а) и последовательном (б) проволочном или тонкопленочном поглотителе на изолирующей подложке

В зависимости от конструкции, фиксированные аттенюаторы могут быть совместимы с коаксиальным, волноводным или полосковым трактом. Существуют конструкции с волноводной лопаткой, с вращающейся лопаткой, на основе направленного ответвителя, с Т-цепью или с П-цепью, на основе линии передачи с потерями или распределенно-резистивной пленки. В частности, в коаксиальном тракте наибольшее распространение получили аттенюаторы с Т-цепью или П-цепью: на основе проволочных резисторов (низкие частоты) или тонкопленочных резисторов (высокие частоты) (рис. 17.5.2). На СВЧ вместо угольных резисторов используются тонкопленочные резисторы. Они используют легированную никромом пленку на высококачественной керамической подложке, чтобы гарантировать надежно закрепленную пленку с низкотемпературными коэффициентами. Такая конструкция делает резисторы чрезвычайно стабильными на высоких частотах. В микроволновых применениях широко используется исключительно малый скин-эффект таких резисторов. Т-образные и П-образные конфигурации, получаемые путем подключения резисторов последовательно и параллельно к центральному проводнику коаксиала, контактируют и с внешним проводником. Поэтому Т-конфигурация может быть изготовлена с помощью одного шунта, фланкированного двумя последовательными резисторами, а П-конфигурация – с помощью одного последовательного резистора, фланкированного двумя шунтирующими резисторами. Последовательные резисторы не способны поглощать мощность более 1 Вт, но для интересующих нас приложений (векторные и скалярные анализаторы цепей) это не является ограничением.

Переменные аттенюаторы

Каждый переменный аттенюатор имеет диапазон регулирования ослабления, например, 0-20 дБ; это регулирование, как указано выше, может быть непрерывным или шаговым, ручным или программным.

Шаговые аттенюаторы

Чтобы создать шаговый аттенюатор, набор фиксированных аттенюаторов механически упорядочивают в барабане или на пластине для получения набора ослаблений путем переключения контактов. Этот способ обеспечивает высокую надежность. Шаги могут быть, например, 0,1, 1, или 10 дБ. На концах устройства устанавливаются стационарные коаксиальные разъемы. Такие конструкции применяются в диапазоне от 0 до 18-20 ГГц для приложений, требующих лишь гладкой неравномерности широкополосной частотной характеристики с низким КСВн и удовлетворительной восстанавливаемости в диапазоне от 0 до 120 дБ.

Переменные аттенюаторы с непрерывной регулировкой ослабления

Существуют следующие разновидности этих аттенюаторов:

- «Стенка потерь»
- «Подвижная лопасть (заслонка)»
- «Роторная лопасть»
- Переменный ответвитель
- Абсорбционный
- Коаксиальная резистивная пленка
- «Переменное Т»
- Запредельный волновод (плунжер)

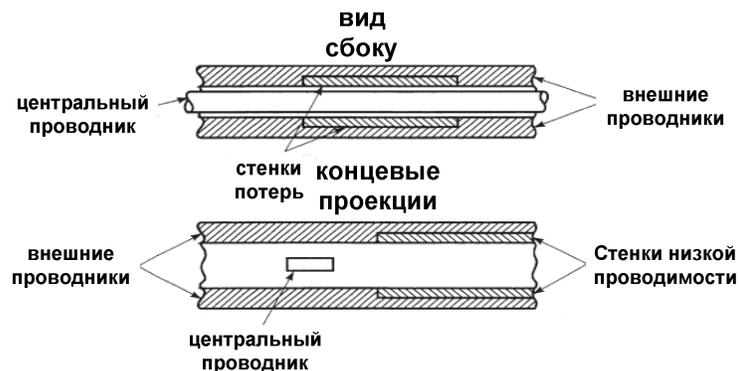
Быстро переключаемые (перестраиваемые) аттенюаторы

Среди существующих переменных электронных (то есть управляемых сигналом) аттенюаторов (шаговых или управляемых напряжением) особое значение для интересующих нас приложений имеют такие, которые допускают очень быстрое переключение ослабления (с временем переключения порядка 30 нс), имеют высокую точность и повторяемость. Управляющие сигналы могут иметь вид тока смещения или цифровой. Ток смещения может быть импульсным, меандровым

или синусоидальным. В коаксиальном шаговом варианте типичные параметры таких аттенюаторов таковы [17.1]:

- Рабочие частоты: 0 – 40 ГГц
- Ослабление: 0 – 11 или 20 дБ
- Шаг: 1 дБ
- Максимальный КСВн: 1.3 на 12.4 ГГц
1.7 на 34 ГГц
1.8 на 40 ГГц
- Повторяемость: в пределах 0 – 0.03 дБ
- Средняя мощность на входе: до 1 Вт

Аттенюатор с поглощающей стенкой

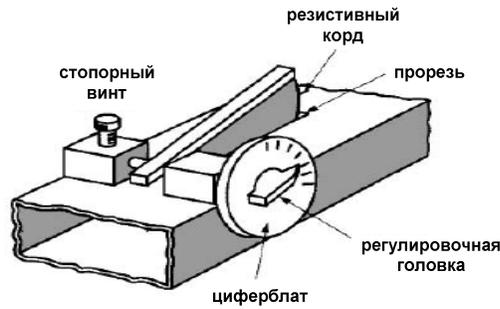


Конструкция волноводного переменного аттенюатора с «поглощающей стенкой»

На рис. 17.6.1 показана конструкция волноводного переменного аттенюатора с «поглощающей стенкой». Он состоит из стеклянной лопасти, покрытой поглощающим материалом, подобным аквадагу или углероду. Лопасть помещается в центре широкого размера направляющего устройства, где интенсивность электрического поля максимальна. Приводной механизм с циферблатом сдвигает лопасть от центра в меру изменения ослабления. Аттенюатор требует калибровки. Чтобы согласовать аттенюатор с волноводом, лопасть может быть клиновидной на каждом конце. Обычно сужение в половину волноводной длины волны обеспечивает приемлемое согласование. Поэтому последнее частотно зависимо и стеклянный диэлектрик лопасти вносит существенный фазовый сдвиг.

Ослабления можно также достигнуть вставлением резистивного элемента сквозь заслонку. Плоскость элемента лежит в распределении электрического поля поперек широкого размера волновода, в результате степень ослабления увеличивается с глубиной вставки. К сожалению, в этой конструкции получается разрыв, обуславливающий плохое согласование.

Аттенюатор с подвижной лопастью (заслонкой)



Конструкция переменного волноводного диссипативного аттенюатора

На рис. 17.6.2 схематично показана конструкция переменного волноводного диссипативного аттенюатора. Пластина вводится в волновод через щель в широкой стенке, перехватывая и поглощая часть волны TE_{10} . Проникновение пластины и, следовательно, ослабление контролируется с помощью шарнирного оборудования. Подобные конструкции с диапазоном ослабления порядка 30 дБ широко используются в микроволновых устройствах. Однако, ослабление оказывается частотно-чувствительным и фаза выходного сигнала является функцией проникновения пластины и, следовательно, ослабления. Это может отразиться на формировании нулей, когда аттенюатор является частью мостовой схемы. Поскольку в данном случае нет простого метода расчета потерь в дБ, такой аттенюатор должен калиброваться с помощью стандарта превосходящего качества.

Аттенюатор с роторной лопастью

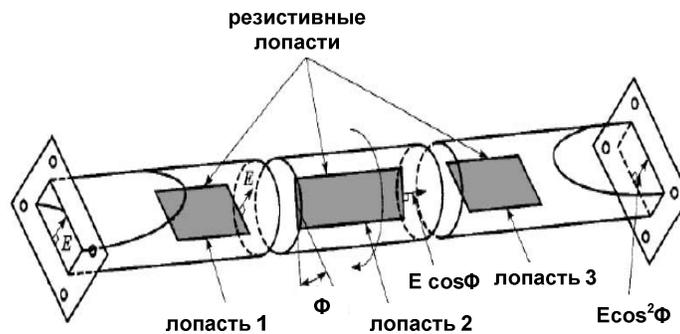


Иллюстрация принципа действия аттенюатора с роторной лопастью

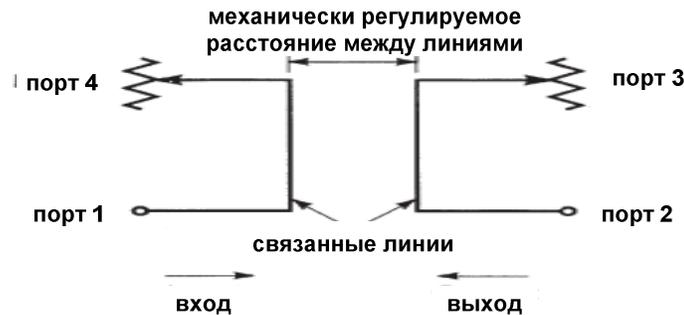
Этот вид аттенюаторов в значительной мере свободен от недостатков предыдущего вида. Он представляет собой прецизионный аттенюатор с прямым отсчетом, подчиняющийся простому математическому закону $A = -20 \lg \cos^2 \theta = -40 \lg \cos \theta$, где A – ослабление в дБ. По существу, он частотно-независим, что очень важно для аттенюаторов. Рис. 17.6.3 иллюстрирует принцип действия аттенюатора. Он состоит из трех последовательно расположенных секций волновода. Прямоугольно-круглый волноводный переход, содержащий горизонтальную аттенюаторную полоску, соединен с вращающимся круглым волноводом, содержащим аттенюаторную полоску. Он, наоборот, соединен с кругло-прямоугольным переходом, содержащим горизонтальную аттенюаторную полоску.

Входящая мода TE_{10} трансформируется в моду TE_{11} в круглом волноводе с помощью прямоугольно-круглого волноводного перехода с пренебрежимыми отражениями. Поляризация моды TE_{11} такова, что электрическое поле перпендикулярно тонкой резистивной пластинке в переходной секции. Как таковая, эта резистивная пластинка имеет пренебрежимое влияние на моду TE_{11} . Поскольку резистивная пластинка в центре может вращаться, ее ориентация относительно элек-

трического поля входящей моды TE_{11} может варьировать так, что регулируется величина ослабления этой моды.

Когда все полосы выровнены, электрическое поле приложенной волны нормально к полоскам, и следовательно, нет текущих потоков в аттенуаторных полосках, а поэтому нет ослабления. В положении, когда центральная аттенуаторная полоска повернута на угол θ , электрическое поле приложенной волны может быть разложено на две ортогонально поляризованные моды: одна перпендикулярная, а другая – параллельная резистивной пластине. Та часть, которая параллельна резистивной пластине, будет поглощаться, в то время как другая часть будет передаваться дальше.

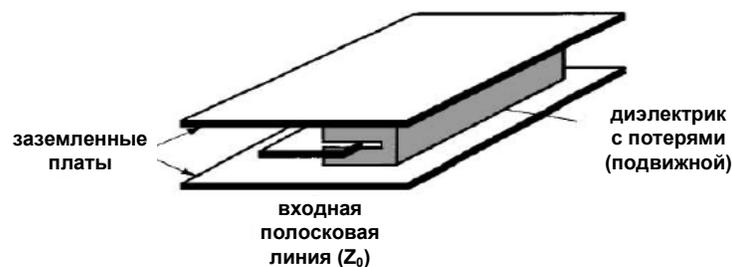
Аттенуатор на базе переменного направленного ответвителя



Принцип действия аттенуатора на базе переменного направленного ответвителя

Созданные и возможные аттенуаторы на базе направленных ответвителей составляют целый класс, в зависимости от вида линии передачи (волноводные, коаксиальные, полосковые), вида и способа регулирования (механического или электрического) связи между первичной и вторичной линиями и т. д. Любой направленный ответвитель уже является аттенуатором, так как мощность, направляемая в ответвительный канал, меньше мощности в первичной линии. Ясно также, что регулируя коэффициент ответвления, мы тем самым регулируем ослабление, вносимое таким аттенуатором. Эта идея схематически поясняется на рис. 17.6.4, где для конкретности взято регулирование связи путем вариации расстояния между линиями. Такие аттенуаторы, как правило, обладают высокой развязкой входа и выхода, широким диапазоном регулирования ослабления, высокой допустимой мощностью на входе, свойством сохранения калибровки в широком диапазоне изменений окружающих условий.

Абсорбционный аттенуатор

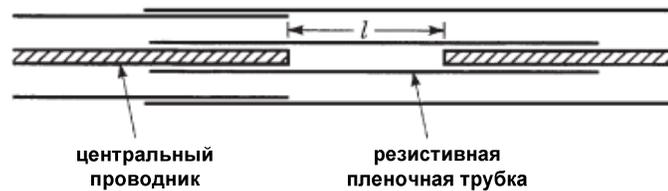


Принцип действия переменного волноводного диссипативного аттенуатора

Рис. 17.6.5 схематично поясняет конструкция переменного волноводного диссипативного аттенуатора. Ослабление обеспечивается потерями в диэлектрике заполнителя. Электрическое поле моды TEM концентрируется в области центральной полоски полосковой линии передачи. По мере вставления поглощающего материала в область высокой плотности электрического поля, большая или меньшая часть TEM -волн пересекается и поглощается диэлектриком, тем самым регули-

руется поглощение аттенюатора. При этом КСВн также стремится изменяться в силу зависимости характеристического сопротивления полосковой линии от объема вставленного диэлектрика. Чтобы минимизировать этот эффект, концы поглощающего материала клиновидно сужаются, что сглаживает импедансную трансформацию на входе и на выходе поглотительной секции. Тем не менее, не удастся достичь хорошего согласования, особенно на нижних частотах, где КСВн существенно возрастает. Кроме того, ослабление аттенюатора увеличивается с ростом частоты при заданной установке регулятора. Это влечет громоздкую процедуру калибровки аттенюатора. Для избавления от этого недостатка иногда применяют компенсационные методы.

Коаксиальный резистивно-пленочный переменный аттенюатор

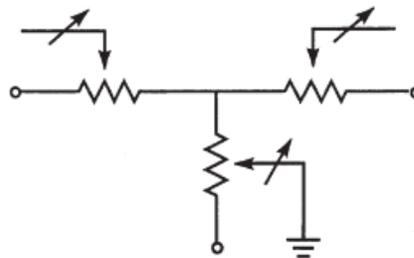


Принцип действия и способ регулировки ослабления коаксиального резистивно-пленочного переменного аттенюатора

Принцип действия и способ регулировки ослабления этого типа аттенюатора показан на рис. 17.6.6. Подбирая размер l свободной от внутренней «жилы» части резистивной пленочной трубки, мы меняем сопротивление $R = rl$ (r – сопротивление трубки на единицу длины), включенное последовательно с внутренним проводником коаксиальной линии передачи; падение напряжения на этом сопротивлении равно $V = Irl$ (I – ток проводимости внутреннего проводника) и вычитается из входного напряжения E_i линии передачи, так что ослабление равно

$$A = 20 \lg \frac{E_i}{E_i - Irl}, \text{ дБ.}$$

Аттенюатор с переменной Т-образной схемой



Принцип действия аттенюатора с переменной Т-образной схемой

Схема этого типа переменного аттенюатора повторяет постоянный аттенюатор с Т-образной резистивной схемой, но с регулируемыми резисторами (рис. 17.6.7). Все три резистора регулируются одновременно. Большое достоинство этого аттенюатора – возможность получения хороших значений входного и выходного КСВн.

Переменный аттенюатор на основе закритического ослабления волновода

Плунжерный аттенюатор использует хорошо известный эффект ослабления волны в волноводе на частоте, меньшей критической для данного волновода. Специальная кругло-волноводная конструкция с закритическим ослаблением на частоте f и регулируемой плунжером рабочей длиной вставляется в линию передачи (волноводную или коаксиальную), которая на частоте f образует моду TEM . Ослабление

$$A = 54.6 \frac{l}{\lambda_c} \sqrt{1 - \left(\frac{f}{f_c}\right)^2}, \text{ дБ}$$

(f_c , λ_c – критические частота и длина волны конструкции) регулируется длиной l рабочей части конструкции. Хотя такой аттенюатор обладает рядом положительных черт (если не считать сложности конструкции и высокой цены), он не обладает сверхширокополосностью и не годится для интересующих нас приложений.

Активные аттенюаторы

Существует широкий (и все расширяющийся) класс активных аттенюаторов (на пин-диодах, на арсенид-галлиевых полевых униполярных транзисторах, на канальных полевых униполярных МОП-транзисторах и т. д.), которых мы не будем касаться в данном обзоре, т. к. из-за их сложности и высокой цены вероятность их применения в анализаторах цепей мала.

ПРИНЦИПЫ ИЗМЕРЕНИЯ МОЩНОСТИ НА СВЧ

Предисловие

Совершенно очевидно, зачем нужно измерять мощность генерируемых, излучаемых, проходящих некоторый тракт, принимаемых сигналов, поскольку мощность – один из основных параметров сигналов, а в диапазоне СВЧ – главный и единственный энергетический параметр, т. к. вследствие волновой природы сигналов понятия тока и напряжения в этом диапазоне вообще не «работают».

В скалярном анализаторе цепей (САЦ) и скалярном спектроанализаторе (СА) мощность – единственный измеряемый параметр сигналов (зондирующих, опорных, отраженных, пропущенных, калибровочных), в векторном анализаторе цепей (ВАЦ) и векторном анализаторе спектра (ВАС) кроме него измеряются еще фазы многочисленных сигналов. Ошибки оценки мощностей в первом случае и мощностей и фаз – во втором напрямую определяют ошибки итоговых результатов: характеристик цепей и сигналов.

Хотя измерение мощности в диапазоне СВЧ – очень «старая» задача, ее решение во все новых разработках всегда не банально, т. к. должно удовлетворять все более высоким и взаимно противоречивым требованиям точности, мощностного и временного разрешений, быстродействия, помехоустойчивости, повторяемости, надежности и экономичности. Это становится возможным в основном благодаря технологическому прогрессу сенсоров мощности и развитию техники оцифровки и компьютерной обработки сигналов. Вторжение элементов цифровой и вычислительной техники произвело революцию в технике измерения мощности, как и вообще в измерительной технике.

Все измерительные приборы диапазона СВЧ должны измерять мощность тех или иных колебаний – в узкой сканирующей полосе или в широкой полосе, на коротком интервале времени («мгновенная»), или на длинном интервале («средняя»), непрерывных или импульсных сигналов.

В этом ряду приборов ВАЦ и САЦ имеют только одну «поблажку»: поскольку эти приборы основаны на принципе зондирования, в них измеряются мощности сигналов с высоким уровнем априорной информации. В частности, как правило, известны временная, частотная и модуляционная структуры сигналов. Зато ряд других требований выступают в ужесточенной форме. Это особенно относится к требованию широкого динамического диапазона (в особенности в ВАЦе) в сочетании с малым временем измерения мощности, обусловленным быстрым и широкодиапазонным сканированием или переключением частоты.

Другое жесткое требование – предельно малая случайная ошибка в условиях ограниченности времени измерения и невысокого отношения сигнал-шум. Это иногда не вполне осознанное требование возникает в связи с предпринимаемыми большими усилиями по согласованию трактов и

калибровке и коррекции систематической ошибки (особенно в ВАЦе): случайная ошибка не должна превалировать над малой скорректированной систематической ошибкой.

Перечисленные выше факторы эффективности измерения мощности взаимно противоречивы, поэтому требуется знать, как методически и аппаратурно оптимально разрешить эти противоречия. Например, общая ошибка измерения мощности складывается из ряда парциальных ошибок, обусловленных различными взаимопротиворечивыми факторами, и понижение общей ошибки есть результат методических и конструктивных компромиссов.

С другой стороны, точность и чувствительность, точность и динамический диапазон мощностей и т. д. предъявляют противоположные требования к параметрам аппаратуры, и требуются многочисленные оптимальные конструктивные решения. Ясно, что лучший способ разрешения всех этих проблем – обобщить исторический опыт создания измерителей мощности, накопленный по мере прогресса технологий и последовательного наступления тенденций микроминиатюризации, цифровых технологий, компьютеризации, – с одной стороны, и по мере развития техники скоростной связи, радиолокации, радионавигации и других, и, следовательно, умеренного усложнения структуры сигналов, – с другой стороны.

Вопрос измерения мощностей опорных и тестовых сигналов ВАЦ, САЦ, СА, ВАС тесно связан с энергетическим балансом прибора, то есть с уровнем мощности генератора, чувствительностью и полосой ПЧ приемников, динамическим диапазоном при тестировании различных устройств, как линейных, так и нелинейных. Учет и увязка этих факторов с эффективностью прибора и экономическими реалиями – столь сложная задача требует выработки определенной концепции оптимизации измерительных приборов СВЧ, которая, в свою очередь, нуждается в усвоении мирового опыта измерения мощности СВЧ-сигналов.

Основные понятия

Единицы

Ватт. Международная система единиц СИ устанавливает в качестве единицы мощности *ватт* (Вт) [21.1]: *один Вт есть один джоуль в секунду.*

Децибелы. Во многих случаях удобно пользоваться *относительной мощностью*, то есть отношением R мощности P к некоторой опорной мощности P_{ref} . Относительная мощность безразмерна. В свою очередь, относительную мощность удобно выражать в *децибелах* (дБ):

$$R_{dB} = 10 \lg\left(\frac{P}{P_{ref}}\right) \quad (21.1.1)$$

Использование децибелов дает два преимущества. Во-первых, сокращается диапазон используемых чисел; например, диапазон от +63 дБ до -153 дБ более компактен, чем диапазон от $2 \cdot 10^6$ до $0.5 \cdot 10^{-15}$. Во-вторых, вместо умножения чисел в натуральных величинах суммируются числа в децибелах, что удобно, например, при каскадном соединении усилителей или последовательном соединении аттенуаторов и т. д.

дБм. Это популярная и удобная логарифмическая единица абсолютной мощности. Формула для мощности P_{dBm} в этих единицах получается из (21.1.1) при $P_{ref} = 1$ мВт:

$$P_{dBm} = 10 \lg\left(\frac{P}{1 \text{ мВт}}\right). \quad (21.1.2)$$

Смысл исчисления абсолютной мощности в этих единицах: «столько-то дБ выше одного мВт» (отсутствие знака означает «плюс», при отрицательном значении P_{dBm} смысл: «столько-то дБ ниже 1 мВт»). Преимущества исчисления абсолютной мощности в дБм те же, что исчисления относительной мощности в дБ.

Категории мощности

Только для сигнала в виде постоянного тока (напряжения) термин мощность имеет прозрачный смысл, не требующий пояснений: мощность равна произведению тока и напряжения. В остальных случаях в зависимости от временной структуры сигнала и опорного масштаба времени различают «среднюю мощность», «импульсную мощность» и «пиковую мощность огибающей».

Средняя мощность.

Для идеализированной модели сигнала в виде периодического переменного напряжения (тока) бесконечной длительности термин «мощность» относится к постоянным составляющим тока и напряжения; здесь опорный масштаб времени бесконечен. В более реалистической модели сигнала в виде синусоидального тока и напряжения на сопротивлении (в общем случае – комплексном) длительностью в целое число n периодов T_0 используется *средняя мощность*, понимаемая как среднее по времени от произведения переменных тока и напряжения на длительности сигнала:

$$P = \frac{1}{nT_0} \int_0^{nT_0} E_0 \sin(\omega_0 t) I_0 \sin(\omega_0 t + \phi) dt. \quad (21.1.3)$$

Здесь $\omega_0 = 2\pi/T_0$ – круговая частота, E_0, I_0 – амплитуды напряжения и тока, ϕ – фазовый угол между током и напряжением. Вычисление интеграла в (21.1.3) дает:

$$P = \frac{I_0 E_0}{2} \cos \phi \quad (21.1.4)$$

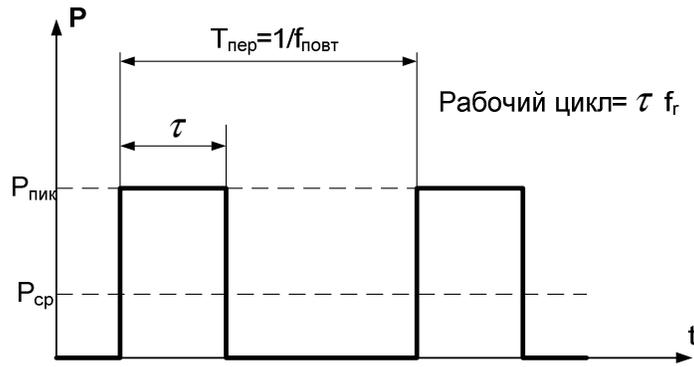
Если длительность сигнала T_s в этой модели составляет очень много периодов T_0 , то можно усреднение проводить на любом интервале $T_0 \lll T_{aver} \leq T_s$, причем неважно, включает ли он целое или нецелое число периодов T_0 , так как разница между этими случаями исчезающе мала. В той же синусоидальной модели среднюю мощность P можно выразить и через среднеквадратичные напряжение E_{rms} и ток I_{rms} :

$$P = E_{rms} I_{rms} \cos \phi, \quad (21.1.5)$$

поскольку $E_0 = \sqrt{2} E_{rms}$, $I_0 = \sqrt{2} I_{rms}$.

В еще более реалистической модели с током и напряжением, имеющими полосовой спектр, измерение средней мощности должно включать усреднение по интервалу T_{aver} , составляющему очень много периодов нижней частоты спектра: $T_{aver} = nT_{max}$, $T_{max} = 2\pi / \omega_{min}$. Если спектр узкополосный, T_{max} можно заменить на период центральной частоты $T_0 = 2\pi / \omega_0$. Если сигнал имеет амплитудную модуляцию, усреднение должно проводиться по многим периодам нижней модуляционной частоты.

Импульсная мощность.



Огибающая импульсного радиосигнала

Рассмотрим модель сигнала в виде последовательности радиоимпульсов, в какой-то степени близких к прямоугольным (рис. 21.1.1). Примем разумное определение эффективной ширины импульса τ_{eff} , например, по уровню -3 дБ или каким-либо другим образом. Пусть f_R – частота повторения, $T_R = 1/f_R$ – период повторения импульсов, P_P – импульсная мощность, P_{aver} – средняя мощность импульсной последовательности. Импульсная мощность по определению:

$$P_P = \frac{1}{2\tau_{eff}} \int P_{envel}(t) dt, \quad (21.1.6)$$

где $P_{envel}(t)$ – мгновенная мощность огибающей радиоимпульса, интегрирование ведется по одному импульсу. Импульсная и средняя мощности последовательности связаны как

$$P_P = \frac{P_{aver}}{Q} = \frac{P_{aver}}{\tau_{eff} f_R}, \quad (21.1.7)$$

где Q – *рабочий цикл последовательности*. В модели сигнала в виде последовательности видеоимпульсов равенство (21.1.6) переходит в формулу:

$$P_P = \frac{1}{\tau_{eff}} \int P_{inst}(t) dt,$$

где $P_{inst}(t)$ – мгновенная мощность импульса, равенство (21.1.7) остается в силе. Таким образом, импульсная мощность получается путем усреднения текущей мощности импульса по его эффективной длительности.

Пиковая мощность

Для идеализированных моделей последовательностей прямоугольных радиоимпульсов или видеоимпульсов понятия импульсной и пиковой мощностей совпадают, поэтому в последнем понятии нет необходимости. Однако, для реальных моделей непрямоугольных импульсов концепция импульсной мощности не вполне удовлетворительна. Трудности растут, когда импульс нарочито не прямоугольный, или когда искажения формы импульса не позволяют произвести точное аппаратное измерение эффективной ширины импульса. Кроме того, само определение понятия эффективной ширины импульса не единственно; например, определения

$$\tau_{eff} = \frac{1}{P_{envel.max}} \int P_{envel}(t) dt$$

и по уровню -3 дБ дадут разный результат. В этих условиях измерение импульсной мощности по (21.1.6) и по (21.1.7) (это бывает аппаратно более удобным) также могут дать различные резуль-

таты. По этим причинам, кроме импульсной мощности используют как независимый энергетический параметр *пиковую мощность* P_{peak} , понимая ее как максимальное значение текущей мощности огибающей импульса: $P_{peak} = P_{envel.max}$. Эта аргументация станет особенно хорошо понятной, если обратиться к стандартной модели видеоимпульса со всеми сложностями его формы («проскоками», «импульсным звоном» и т. п.) (см. IEEE-стандарт ANSI/IEEE Std 194-1977 [21.7]).

Мощностные сенсоры

Важнейшей и принципиальной частью измерителя мощности СВЧ-волны является *сенсор*: прибор, воспринимающий мощность волны и конвертирующий высокочастотную мощность в измеримый сигнал постоянного тока или низкочастотный сигнал. Из всех мыслимых чувствительных к мощности переменного тока приборов в историческом развитии «выжили» три вида мощностных сенсоров, актуальных для измерительных приборов диапазонов ВЧ и СВЧ – *термистор*, *термопара* и *диодный детектор*. Каждый из них имеет определенные преимущества и недостатки по отношению к другим сенсорам. Хотя по сумме современных требований к анализаторам цепей будет выбран один из этих типов, полезно обсудить все три типа.

В целом каждый блок измерения мощности анализатора цепей кроме сенсора включает соответствующий соединитель сенсора с линией передачи, мощность волны в которой должна быть измерена, соединитель выхода сенсора с собственно измерителем, который, в свою очередь, может включать оцифровщик, элемент временного окна, регистратор и т. п. Наличие цифровой обработки в блоке измерения мощности при условии подходящего математического обеспечения позволяет расширить список вычисляемых мощностных параметров; например, кроме средней, импульсной и пиковой мощности, блок может отображать такие мощностные характеристики, как интегральная функция распределения мощности, позволяющая определять процент времени превышения отношением «пик-среднее» специфического сигнального уровня [21.8].

Термисторные сенсоры

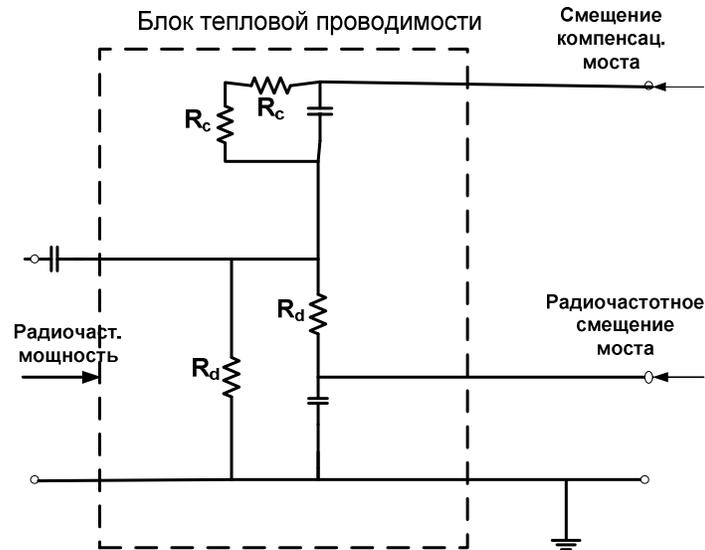
Термисторные сенсоры – разновидность *болометрических* сенсоров. Другая разновидность – *сенсоры на барретерах*. Барретер это тонкая проволочка с положительным температурным коэффициентом сопротивления и сравнительно небольшим сроком службы. Болометрические сенсоры, особенно термисторные, сыграли важную историческую роль в радиочастотных и СВЧ измерениях мощности, но в последние годы термопарные и диодные технологии отвоёвали существенную часть этих приложений, вследствие их увеличенной чувствительности, более широкого динамического диапазона и более высоких мощностных характеристик. Тем не менее, термисторы все еще выбирают для переноса мощностных стандартов, вследствие их подстановочной возможности к мощности по постоянному току. Поэтому производитель измерительных приборов СВЧ даже в случае, если в их блоках измерения мощности не будет термисторов, обязательно встретится с последними в порядке мощностной трассируемости, в частности, при калибровке примененных мощностных сенсоров. И в настоящее время важное значение имеет возможность применения термисторных сенсоров в переносимых стандартах, в процессах калибровки других типов сенсоров, в технике реализации трассируемости и циклического обмена между пользователями по всему миру. Термисторные сенсоры позволяют производить измерения калибровочного фактора с высокой точностью и повторяемостью, они высоко портативны и стабильны.

Вообще болометры это мощностные сенсоры, которые действуют путем изменения сопротивления при изменении температуры, которые, в свою очередь, образуются в результате превращения радиочастотной или микроволновой энергии в тепло внутри болометрического элемента. В отличие от барретера, термистор это полупроводниковый элемент с отрицательным температурным коэффициентом сопротивления.

Термисторные сенсоры, используемые для радиочастотных измерений, это маленькие бусинки металлического окисла, обычно 0.4 мм диаметром с проволочными усиками диаметром 0.03 мм. В балансно-мостовой технике термисторный элемент всегда устанавливается на постоянном сопротивлении R с помощью постоянного или низкочастотного смещения. Когда радиочастотная мощность рассеивается на термисторе, уменьшая сопротивление, мощность смещения стремится

сбалансировать мост и поддерживать R тем же. Это уменьшение мощности смещения затем показывается на измерителе и индицирует радиочастотную мощность.

Термисторные элементы устанавливаются в коаксиальных структурах, поэтому они совместимы с общими системами на линиях передачи, работающими как на радиочастотах, так и на микроволновых частотах. Современные термисторные сенсоры имеют второй комплект компенсирующих термисторов, предназначенных для поправки на вариации температуры окружающей среды.



Термисторная установка с тепловой компенсацией

Для большей конкретности рассмотрим, например, термисторную установку Agilent 478A (рис. 21.2.1). Она содержит 4 согласованных термистора: два *измерительных* R_d и два *компенсационных* R_c , и может измерять мощность в диапазоне частот от 10 МГц до 10 ГГц, или дл 18 ГГц. Измерительные термисторы по 100 Ом каждый включены последовательно (200 Ом) в балансном мосте по постоянному току, и параллельно по радиочастоте, образуя нагрузку 50 Ом для измеряемого сигнала. Принципиальное достоинство такой схемы включения – образование моста с радиочастотной «землей».

Компенсационные термисторы, которые отслеживают изменения температуры окружающей среды, но не изменений радиочастотной мощности, также соединены последовательно, образуя второй мост, называемый компенсационным. Они полностью заключены в полость для электрической изоляции от радиочастотного сигнала, но установлены в том же теплопроводящем блоке, что и измерительные термисторы. Тепловая масса блока достаточно велика, чтобы предотвратить резкие температурные градиенты между термисторами. Это улучшает изоляцию системы от тепловых влияний, подобных эффекту человеческих рук.

Описанное коаксиальное включение термисторов, при котором два термистора параллельны на радиочастоте, но последовательны по постоянному току, порождает специфическую ошибку компенсации, а следовательно, измерения: если два термистора не одинаковы по сопротивлению, то больший радиочастотный ток течет (а следовательно, большая радиочастотная мощность выделяется) на меньшем сопротивлении, а большее постоянное напряжение падает (а следовательно, большая мощность постоянного тока выделяется) на большем сопротивлении. Ошибка, возникающая из-за неэквивалентности распределения диссипируемых мощностей на радиочастоте и на постоянном токе, называется *дуально-элементной*, пропорциональна измеряемой мощности и для современных термисторов составляет меньше 0.1 % на правом конце диапазона измеряемых мощностей.

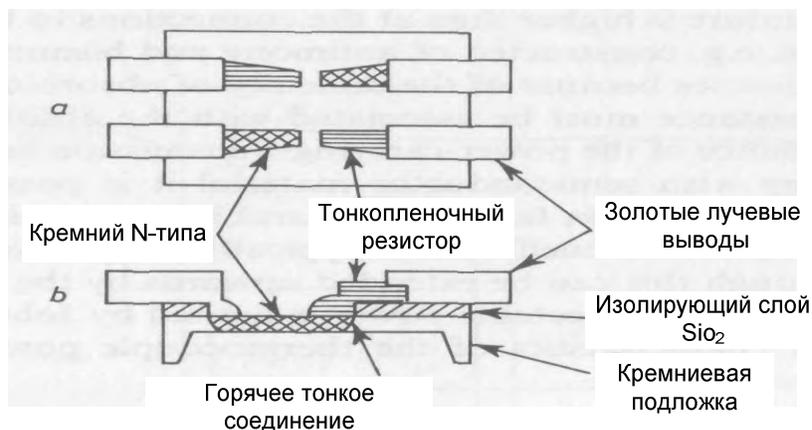
Термопарные сенсоры

В результате технической эволюции комбинирования тонкопленочной и полупроводниковой технологий термопарные сенсоры стали в целом предпочтительными для восприятия радиочастотной и СВЧ мощности в сравнении с термисторными сенсорами примерно в 1974 году. Главные причины этого: преимущество в чувствительности и внутренне свойственная квадратичная характеристика. Другие преимущества: бо́льшая механическая жесткость, бо́льшая чувствительность и меньшая измерительная неопределенность.

В термопаре, состоящей из двух разнородных металлических проводов, соединенных на одном конце и подсоединенных к измерителю на другом, генерируется термоэлектрическое напряжение, если контактная температура выше, чем на соединении с измерителем. Простые проволочные термопары, сконструированные из сурьмы и висмута, редко используются на СВЧ из-за трудностей поглощения мощности в горячем стыке. Комбинируя тонкопленочную технологию с полупроводниковым материалом, возможно произвести термопары с калибровочным коэффициентом, сравнимым с тем, что имеют термисторные установки, но со много меньшими коэффициентами отражения [21.8], обычно 0.1. Мощностной измерительный диапазон: от 0.1 мкВт до 1 Вт; он может быть расширен вверх с помощью использования калибровочных аттенуаторов, но не вниз, т. к. ограничен шумом Джонсона в элементе.

Главные характеристики термопарного сенсора таковы:

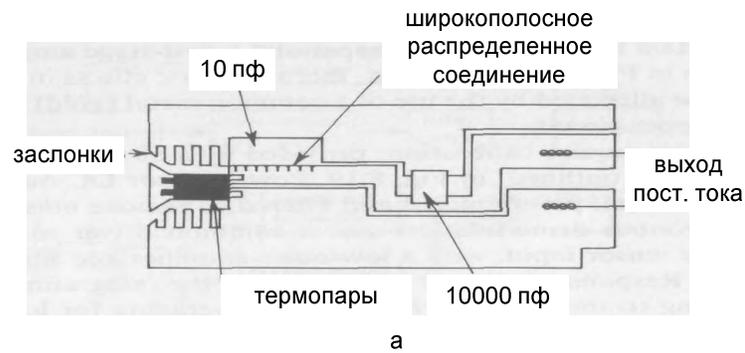
- Очень широкий динамический диапазон порядка 70 дБ.
- Высокая температурная стабильность.
- Частотный диапазон от постоянного тока до 40 ГГц.
- Высокая точность \square 1 %.
- Быстрый отклик.
- 300 % кратковременная перегрузка.



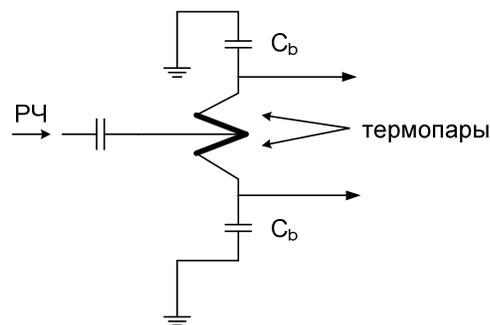
Упрощенный эскиз конструкции термопарного сенсора HP 8480 фирмы Хьюлетт-Паккард

В качестве примера на рис. 21.2.2 показана конструкция термопарного сенсора HP 8480 фирмы Хьюлетт-Паккард. Две термопары заложены в один кремниевый чип, чтобы обеспечить последовательно-параллельную цепь для разделения цепей постоянного тока и радиочастотной. Тонкопленочный резистор из нитрида тантала следует за горячим стыком, согласно характеристическое сопротивление полосковой линии на соединении с лучевым выводом. Он отделен от кремния слоем диоксида кремния, исключая зону вблизи центра чипа, где имеется маленькая дырка в диоксидном слое. Центр чипа очень тонкий, а вне кромок толстый, так что поглощение тепла в рези-

стороне выращивает центральную температуру, формируя горячий стык. Внимательная формовка вложения резистора концентрирует токовую плотность и, следовательно, тепло в центре чипа.



Соединение лучевых выводов с копланарными полосковыми линиями в сенсоре HP 8480 фирмы Хьюлетт-Паккард



Последовательно-параллельная ПТ-PC конфигурация подсоединения термопар

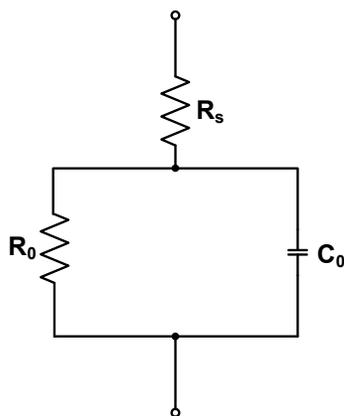
Соединение лучевых выводов с копланарными полосковыми линиями показано на рис. 21.2.3. Полный размер приблизительно 7x4 мм, с соединением термопары с радиочастотой через блокирующий конденсатор (не показан). Сдвоенные емкости для последовательно-параллельной ПТ-PC конфигурации показаны на рис. 21.2.4 (на рис. 21.2.2 соединение ПТ сделано прямым к одной термопаре, в то время как другие заземлены на PC через 10 пФ и 10000 пФ емкости). Широкополосное согласование достигается распределенным штыревым контактом на 10 пФ емкость и согласующими дросселями на PC входе.

Конструкция мощной головки зависит от частотного и мощностного диапазона. Например, низкие частоты требуют большой блокирующей емкости, которая может компрометировать высокочастотную эффективность. Масса и форма резистивной нагрузки имеет важное значение для динамического диапазона и полосы согласования. Дрейф из-за вариаций внешней температуры уменьшается до очень низких уровней, благодаря горячим и холодным стыкам, генерирующим эдс противоположной полярности, поэтому не дающим итогового выхода, когда внешняя температура меняется. Температурная зависимость конвективного охлаждения комплексной сборки может иметь значительное влияние на PC и ПТ чувствительность, но внимательное конструирование может уменьшить это до величин меньших 0.1 % на градус Цельсия. Выход мощностного сенсора по ПТ подается на мощемер, который состоит из очень маломощного УПТ высокого усиления и измерителя. Термопарный сенсор требует калибровки [21.9, 21.10], техники которой мы здесь не касаемся.

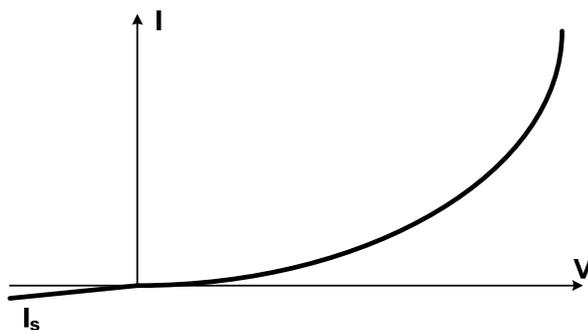
Диодные сенсоры (диодные детекторы)

Диодные мощностные сенсоры используются в скалярных анализаторах цепей и приложениях мониторинга мощности [21.10], например, в устройствах автоматической регулировки мощности (АРМ) генератора. Эти сенсоры обычно действуют в диапазоне измеряемых мощностей от -50 дБм до +20 дБм. Диапазон квадратичности диодной характеристики простирается от -70 дБм до -20 дБм, при большей мощности характеристика становится линейной. В сочетании с термопарным сенсором полный динамический диапазон достигает 90 дБ. Характерная измерительная неопределенность диодного сенсора ± 0.5 дБ может быть уменьшена микропроцессорной коррекцией. Как и термопарный сенсор, диодный требует калибровки.

Диоды, используемые в данных сенсорах в качестве детекторов с нулевым смещением, имеют сопротивление, задаваемое наклоном вольт-амперной характеристики в начале координат (*начальное сопротивление*), *рассеянным* или *массовым* сопротивлением материала и контактной емкостью. Упрощенная модель диода показана на рис.21.2.5, в которой мы проведем дальнейшие упрощения, предположив, что рассеянное сопротивление R_s пренебрежимо в сравнении с начальным сопротивлением R_0 , а частотный диапазон позволяет игнорировать контактную емкость C_0 .



Упрощенная эквивалентная схема диода



Схематический вид статической диодной характеристики (масштабы при положительных и отрицательных V разные)

Статическая диодная характеристика (рис.21.2.6) математически может быть записана в виде:

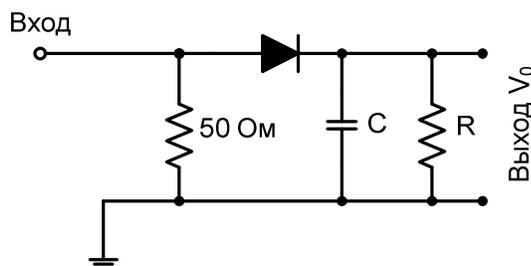
$$i = I_s [\exp(v/D) - 1], \quad (21.2.1)$$

где $-\infty < i < \infty$ – ток, $-\infty < v < \infty$ – напряжение на диоде, I_s – ток обратного насыщения, $D = nkT/e \approx 25\text{-}50$ мВ при температуре 290°K , k – постоянная Больцмана, T – диодная температура, e – заряд электрона, $n=1.2$ для диода Шоттки, $n=2$ – для точечно-контактного диода.

Разлагая (21.2.1) по степеням отношения v/D , при $v \ll 50$ мВ можно ограничиться квадратичным членом:

$$i \approx I_s \left[\frac{v}{D} + \frac{v^2}{2D^2} + \dots \right]. \quad (21.2.2)$$

На первый взгляд, при очень малом напряжении можно было бы ограничиться линейным членом, но при переменном напряжении на входе детектора выпрямленный выходной ток в резистивной нагрузке сильно зависит от квадратичного члена в разложении (21.2.2). Действительно, рассмотрим рис. 21.2.7, на котором показана детекторная цепь, состоящая из линии, нагруженной на 50 Ом, диода и RC -фильтра с детекторным выходом V_0 на сопротивлении R .



Детекторная цепь

Напряжение на диоде:

$$v = V_{RF} - V_0 = V_s \cos \omega t - V_0; \quad (21.2.3)$$

подставляя это выражение в (21.2.2) и усредняя по периоду, находим средний ток \bar{i} в R :

$$\bar{i} = \frac{I_s}{2\pi} \int_0^{2\pi} \left[\frac{V_s \cos \omega t - V_0}{D} + \frac{(V_s \cos \omega t - V_0)^2}{2D^2} \right] d\omega t. \quad (21.2.4)$$

Учитывая, что интеграл от $\cos \omega t$ по периоду равен нулю, из (21.2.4) получаем

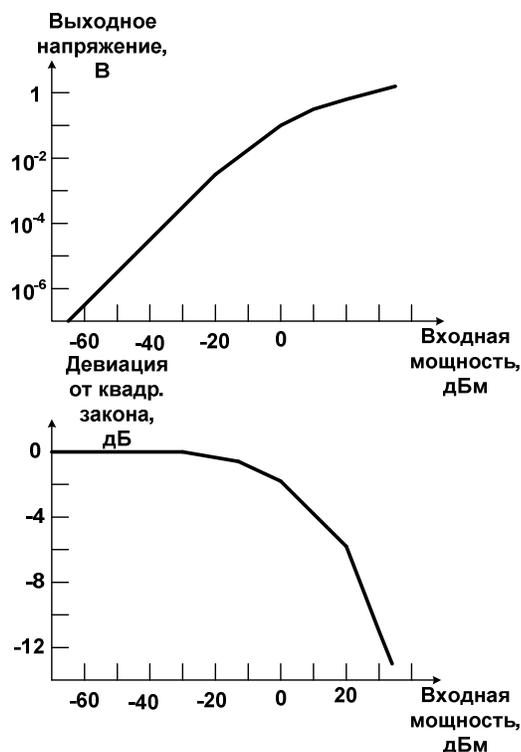
$$\bar{i} = I_s \left[-\frac{V_0}{D} + \frac{V_s^2}{4D^2} + \frac{V_0^2}{2D^2} \right]. \quad (21.2.5)$$

Поскольку $V_0^2/2D^2 < V_0/D$, из (21.2.5) находим выходное напряжение детектора:

$$V_0 = \bar{i} R = \frac{I_s R}{4[D^2 - (I_s R)/2]} V_s^2. \quad (21.2.6)$$

Уравнение (21.2.6) показывает, что выходное напряжение детектора пропорционально уровню входной радиочастотной мощности [21.10, 21.11]. Для больших входных сигналов детекторная цепь ведет себя как пиковый детектор с емкостью C , заряженной до пикового напряжения на каждой прямой половине периода. После сглаживания получается $V_0 \approx V_s$, то есть линейный пико-

вый детектор. Таким образом, диапазон квадратичного закона детектирования ограничен шумом на своем нижнем пределе и наступлением линейного закона детектирования на своем верхнем пределе.



Верхний рисунок: зависимость выходного напряжения от входной мощности; нижний рисунок: зависимость девиации от квадратичного закона от входной мощности

Рис. 21.2.8 показывает типичную кривую диодного детектирования [21.12], начинающуюся близко к шумовому уровню порядка -70 дБм и продолжающуюся до уровня $+20$ дБм. В нижней области «квадратичного закона» выходное напряжение диодного детектирования прямо пропорционально входной мощности (V_{out} пропорционально V_{in}^2) и поэтому измеряет мощность прямо. Выше -20 дБм диод совершает характерный переход к линейной детекторной функции (переходная зона) (V_{out} пропорционально V_{in}) и соотношение квадратичного закона уже не верно.

Традиционно диодные мощностные сенсоры специфицировали для измерения мощности в диапазоне от -70 до -20 дБм, делая их преимущественным сенсорным типом для применений, которые требуют высокой чувствительности измерений, подобных верификационным входным уровням в тестах чувствительности приемника. В применениях, которые требуют высокой скорости измерений, диодные сенсоры предпочитают термопарным типам за их более быструю реакцию на изменения входной мощности. Примеры этой высокочувствительной диодной технологии – серия 8480 D -адических сенсоров фирмы Аджилент.

Когда необходимо тестирование от -70 дБм до $+20$ дБм, что становится все более частым случаем, традиционный подход – использовать диодный сенсор для покрытия нижнего конца диапазона, а термопарный сенсор – верхнего конца диапазона. Однако, эта дуальная измерительная конфигурация не отвечает требованию малости тестового времени, особенно если должна достигаться оптимальная точность измерения. Чтобы выполнить эти требования, часто идут по другому пути: расширяют динамический диапазон диодных мощностных сенсоров путем использования корректирующих факторов. Последние выводятся из экспериментов с генератором непрерывных волн разных частот, заносятся в таблицу с двумя входами – частота и мощность, и сохраняются в памяти компьютера. С помощью корректирующих факторов можно компенсировать отклонения детекторной характеристики от квадратичного закона в переходной области (примерно от -29 дБм

до 0 дБм) и области линейного детектирования (выше 0 дБм). В применении к единичному сенсору это приводит к возможности точного измерения мощности непрерывной волны и сигналов постоянной амплитуды от -70 дБм до +20 дБм. Этот рецепт расширения динамического диапазона годится для скалярного анализатора цепей, пока и поскольку в нем используются сигналы без модуляции и с периодической модуляцией, но не обязательно годится, если используется непериодическая модуляция и, тем более, сложные виды модуляции, типичные для систем беспроводной связи. В последнем случае корректирующие факторы, установленные в экспериментах с непрерывной волной, могут привести к дополнительным ошибкам, сравнимым с самыми большими ошибками, обусловленными измерительными неопределенностями опорного источника и инструментальными неопределенностями.

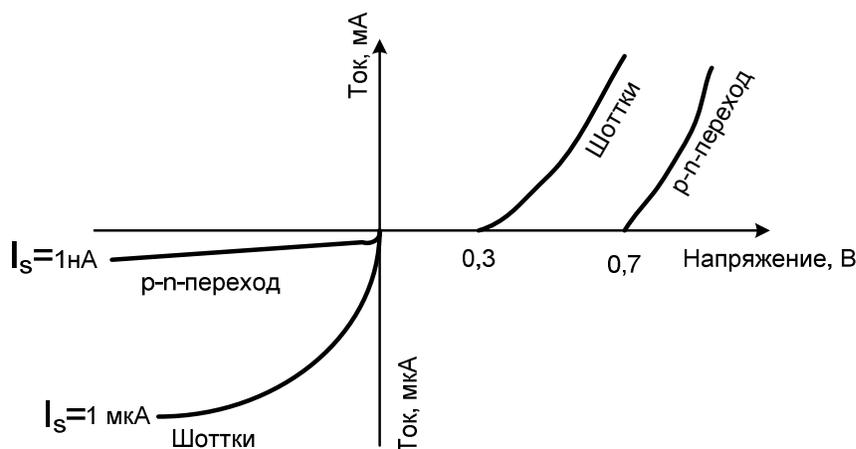
Дифференцируя (21.2.1) по i (для этого сначала находим di/dv , а затем «переворачиваем» дробь) и полагая $v = 0$, получаем начальное сопротивление:

$$R_0 = \left. \frac{dv}{di} \right|_{v=0} = \frac{nkT}{I_s e}. \quad (21.2.7)$$

Если ко входу детектора подключается коаксиальная линия передачи с характеристическим сопротивлением 50 Ом, то было бы идеально, если R_0 было равно 50 Ом; тогда детектор был бы согласован с подводимой линией и осуществлял бы максимальное преобразование мощности и имел бы максимальную чувствительность. Однако это невозможно. Действительно, подставляя в (21.2.7) значения n, k, e и $T = 290^0 K$, находим

$$R_0 \approx \frac{0.025}{I_s},$$

и для того, чтобы было $R_0 = 50$ Ом, необходимо, чтобы обратный ток насыщения I_s достигал 0.5 мА. На самом деле, для точечно-контактных диодов, диодов на $P-N$ переходе и на металл-полупроводниковом переходе значения I_s достигают лишь величин 10 мкА, 1 нА, 1 мкА соответственно, а минимальные значения R_0 – 2.5 кОм, 25 Мом, 25 кОм. Продолжая рассмотрение вопроса, заметим, что точечно-контактные диоды хрупки, нестабильны и не могут использоваться для нормального точного измерения мощности; остается сравнить диоды на $P-N$ переходе и на металл-полупроводниковом переходе (диоды Шоттки), которые, наоборот, высоко стабильны.



При включении нормального диода Шоттки вместо точечно-контактного диода на р-п-переходе наблюдается большое увеличение обратного тока и уменьшение прямого порога

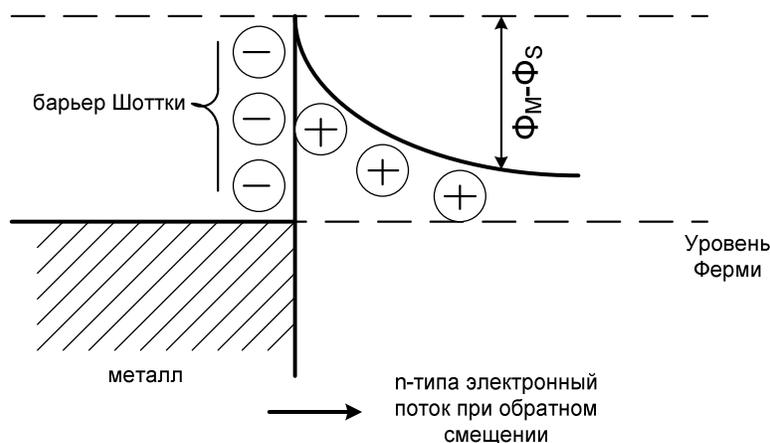
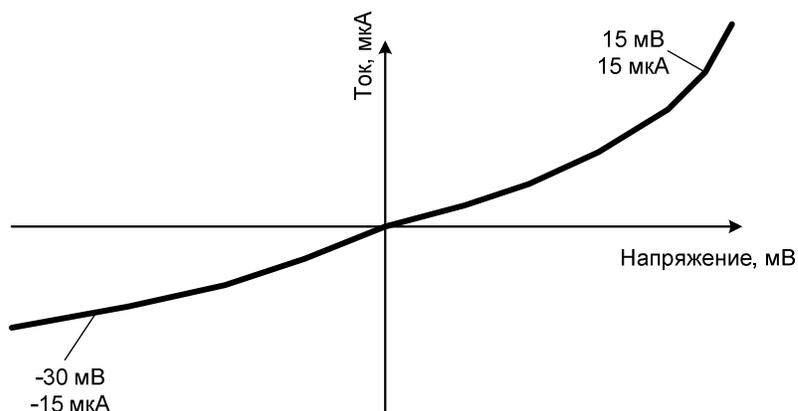
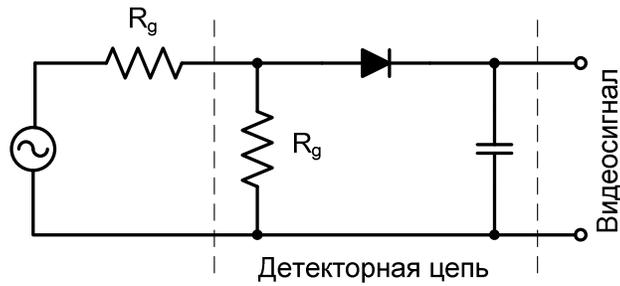


Схема энергетических уровней для диода Шоттки



Результирующая вольтамперная характеристика после уменьшения высоты барьера сглаживанием поверхности перехода в диоде Шоттки

Это сравнение иллюстрируется рис.21.2.9, показывающим большое увеличение обратного тока и уменьшение прямого порога для нормального диода Шоттки. На рис. 21.2.10 приведена схема энергетических уровней, показывающая, что работа выхода металла ϕ_M больше работы выхода ϕ_S полупроводника n -типа. При обратном смещении электроны способны преодолеть крутой разрывный *барьер Шоттки*. Возможно уменьшить высоту барьера, изменяя состояние поверхности перехода, чтобы достигнуть обратного тока 10 мкА, обычного для точечно-контактного диода, но при грубой повторяемости, характерной для металл-полупроводникового диода Шоттки [21.12]. Результирующая характеристика показана на рис.21.2.11.

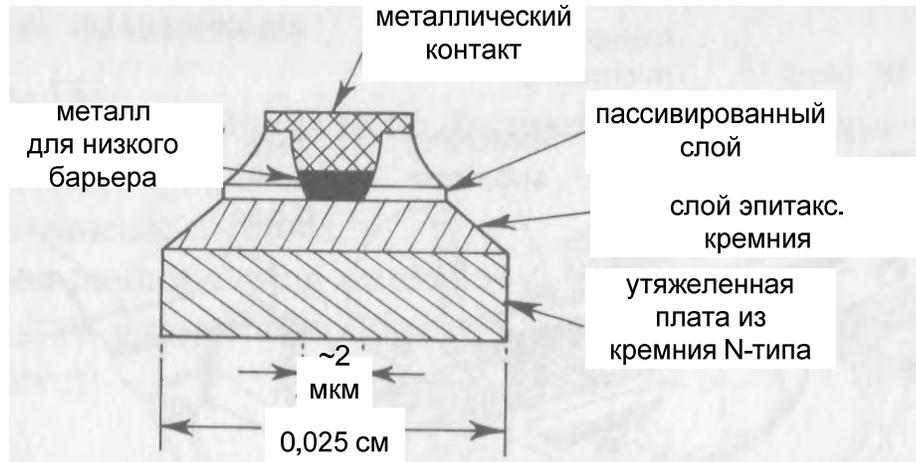


Модифицированная детекторная цепь с согласующим шунтирующим резистором

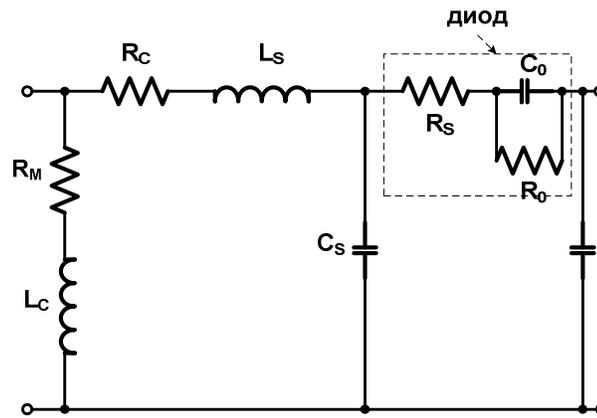
Поскольку даже при таком сглаживании барьера в диоде Шоттки R_0 не может быть ниже 2.5 кОм, в детекторной цепи (рис.21.2.12) применяют специальный согласующий шунтирующий резистор, чтобы согласовать эту цепь с внутренним сопротивлением R_g генератора. Это подразумевает, что $R_0 \gg R_g$; но дифференцируя (21.2.7) по температуре T :

$$\frac{dR_0}{dT} = \frac{nk}{I_s e} = \frac{R_0}{T},$$

закключаем, что для уменьшения температурной зависимости R_0 оно должно быть по возможности как можно меньше. На практике, разумный компромисс достигается при $R_0 = 2.5$ кОм.



Планарная эпитаксиальная конструкция мезоструктурной формы диода Шоттки с малой областью перехода



Эквивалентная схема детекторной цепи на диоде Шоттки с малой областью перехода

Для того чтобы детектор имел хорошие высокочастотные характеристики, емкость перехода C_0 должна быть как можно меньшей. Планарная эпитаксиальная конструкция (рис.21.2.13) мезоструктурной формы с малой областью перехода сводит C_0 до величины порядка 0.1 пФ при действии до 18 ГГц. Эпитаксиальный слой слегка легирован в сравнении с сильно легированной кремниевой подложкой. Оксидированный пассивационный слой изолирует эпитаксиальный слой от верхних слоев, исключая центральный, где вытравлено малое отверстие для металлополупроводникового перехода. Выбор металла перехода критичен для производства низкого барьера [21.8]. Широкополосная характеристика получается погружением диода в компенсирующую согласующую цепь на сапфировой подложке с помощью тонкопленочных методов. Это показано в форме эквивалентной цепи на рис.21.2.14, где R_M ($= 50$ Ом), L_C, R_C – компенсирующие компоненты, L_S, C_S отражают паразитные эффекты, а диодный кристалл находится внутри штриховой линии [21.13]. Наконец, корпус сенсора должен быть тщательно сконструирован для минимизации температурных градиентов поперек диода путем размазывания тепловой однородности до обоих концов блока.

В применениях к диодным детекторам предъявляются различные наборы требований. Так, для использования диодных детекторов в скалярных анализаторах цепей (САЦ) важны следующие параметры и характеристики детекторов:

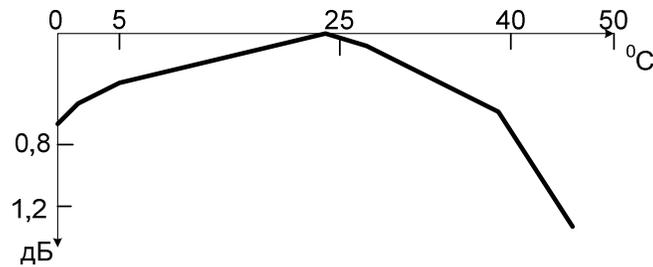
- Динамический диапазон
- Мощностная точность
- Частотный отклик
- Входной КСВн
- Тангенциальная чувствительность
- Температурная стабильность
- Диапазон квадратичного закона

Наиболее часто в САЦ используются металлически-полупроводниковые барьерные диоды Шоттки. Они имеют широкий динамический диапазон (от -55 до +15 дБм при полосе 1 кГц), низкий шум: на 15 дБ ниже диода с точечным контактом, мощностную точность ± 0.1 дБ, широкий частотный диапазон от 0.01 до 20 ГГц, и входной КСВн ≤ 1.5 . Точечно-контактные диоды менее стабильны и более ненадежны, чем диоды Шоттки, но их много меньшая емкость перехода делает их единственно подходящими детекторами на более высоких частотах вплоть до 200 ГГц.

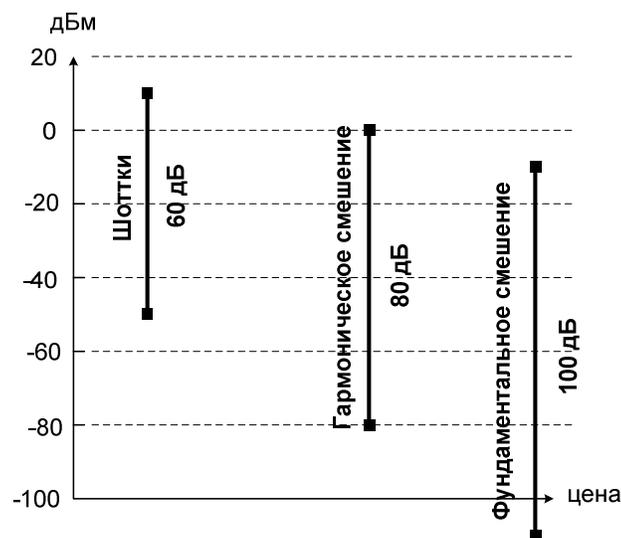
Диодная тангенциальная чувствительность [21.14] обычно измеряется отысканием входного уровня импульса, который при детектировании показывает метку с низким шумовым контуром,

как раз касательную к верхнему шумовому контуру статического уровня детектирования. Для барьерных диодов Шоттки она варьирует от -50 до -55 дБм в полосе 10 МГц

Типичный температурный отклик детектора показан на рис.21.2.15, где над диапазоном шириной 50⁰С выход меняется на 1.2 дБ при радиочастотном входе постоянной амплитуды. В температурно-компенсационной цепи сохраняется кусочно-линейная аппроксимация отклика в виде трех линейных областей, которая используется для хардверной или софтверной процедуры коррекции.



Типичный температурный отклик детектора



Сравнение различных видов детекторов по цене и динамическому диапазону

Границы действия квадратичного закона обсуждались выше.

На рис. 21.2.16 детекторы на диодах Шоттки ориентировочно сравниваются по цене и динамическому диапазону с применяемыми в векторных анализаторах цепей детекторами гармонического смещения и детекторами фундаментального смещения [21.15]. Как видим, динамический диапазон детекторов на диодах Шоттки примерно совпадает с динамическим диапазоном сигналов тестовой сети, ограниченным сверху влиянием нелинейностей, а снизу – уровнем шумов.

Заключение

Описанные три вида сенсоров (термисторные, термопарные, диодные) покрывают широкий динамический диапазон от -70 дБм до +20 дБм, или даже больше, если прибегнуть к помощи аттенуаторов. Быстродействие и простота, особенно при измерении пиковой мощности, наилучшим образом представлены диодными детекторами, также как возможность измерений мощности на очень низком уровне. Термисторы действуют в цепях с замкнутой петлей, поэтому они предпоч-

тительны там, где изменения окружающей температуры становятся большими, вместо того, чтобы принимать специальные меры для гарантирования температурного выравнивания в термопарных и диодных головках. Термопары имеют наименьший КСВ. Калибровка мощностных сенсоров зависит от стандартных головок с калибровочным фактором и эффективностью, трассируемыми до национальных лабораторий стандартов. Эти вторичные стандарты непосредственно сравниваются с лабораторными или производимыми головками в порядке установочного точного тестового эксперимента.

Ошибки измерения мощности

Источники ошибок

В данном разделе мы коснемся только трех принципиальных составляющих систематической ошибки измерения мощности. С более подробными методиками анализа ошибки (в том числе со стандартизованными международными методиками) можно познакомиться в литературе, например, [21.3–21.6]. В качестве сенсора в необходимых случаях имеется в виду диодный сенсор с квадратичной характеристикой.

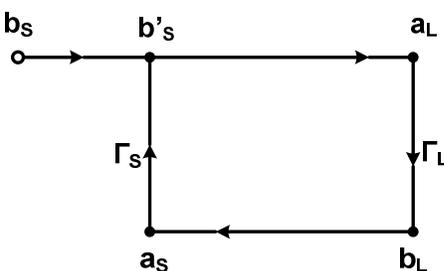
Главные источники ошибки измерения мощности на СВЧ таковы:

- Рассогласование мощностной головки с подводящим трактом СВЧ сигнала.
- Неопределенность из-за многократных отражений между мощностной головкой и источником измеряемого сигнала.
- Неточность калибровки мощностной головки и ее к.п.д.

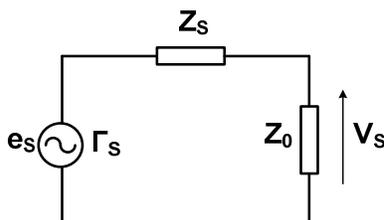
Кроме того, существуют систематические и случайные составляющие ошибки, возникающие в цепях обработки последетекторных сигналов, оцифровки, регистрации и т. д.



Схематичное представление процесса измерения мощности



Потоковый граф измерителя мощности радиочастотного сигнала



Эквивалентная схема источника с линией передачи

Схематично и условно процесс измерения мощности СВЧ сигнала можно представить в виде рис.21.3.1, где Γ_L, Γ_S – коэффициенты отражения головки и «источника»; последний символизирует цепи, подводящие сигнал к «линии передачи», «мощностная головка» включает сенсор и подает на «измеритель» протектированный сигнал, «измеритель» может включать оцифровку, регистратор и т. д.

Рассогласование мощностной головки

Линия передачи как двухпортовое устройство, к которому подводится внешняя волна \mathbf{b}_s измеряемого сигнала и подсоединены «источник» и головка, описывается потоковым графом [21.8], показанным на рис.21.3.2, где $\mathbf{a}_L, \mathbf{a}_S$ – входные, а $\mathbf{b}_L, \mathbf{b}'_S$ – выходные узлы. В условном узле \mathbf{b}'_S суммируются волна \mathbf{b}_s и отраженная от «источника» волна, то есть

$$\mathbf{b}'_S = \mathbf{b}_s + \Gamma_S \mathbf{a}_S. \quad (21.3.1)$$

Нам понадобится также эквивалентная схема с источником эдс (рис.21.3.3), где Z_0 – характеристическое сопротивление линии передачи, а Z_S учитывает рассогласование источника с этой линией.

Прямая волна в согласованную нагрузку Z_0 :

$$\mathbf{b}_s = \frac{V_s}{\sqrt{Z_0}}. \quad (21.3.2)$$

Но источник не согласован с линией передачи, поскольку $Z_S \neq Z_0$, или $\Gamma_S \neq 0$. Поэтому

$$V_s = e_s \frac{Z_0}{Z_S + Z_0}, \quad (21.2.3)$$

то есть

$$\mathbf{b}_s = e_s \frac{\sqrt{Z_0}}{Z_S + Z_0}. \quad (21.3.4)$$

Согласно теории эквивалентных линий передачи коэффициент отражения от источника:

$$\Gamma_S = \frac{Z_S - Z_0}{Z_S + Z_0}. \quad (21.3.5)$$

Мощность, поглощаемая нагрузкой, можно найти как разность между падающей и отраженной мощностями. Из графа рис. 21.3.2 находим падающую волну как:

$$\mathbf{a}_L = \frac{\mathbf{b}_s}{1 - \Gamma_S \Gamma_L}, \quad (21.3.6)$$

и отраженную волну как:

$$\mathbf{b}_L = \frac{\mathbf{b}_s \Gamma_L}{1 - \Gamma_S \Gamma_L}. \quad (21.3.7)$$

Разность мощностей этих волн есть поглощаемая в нагрузке мощность:

$$P_L = \frac{1}{2}|\mathbf{a}_L|^2 - \frac{1}{2}|\mathbf{b}_L|^2 = \frac{1}{2}|\mathbf{b}_s|^2 \frac{1-|\Gamma_L|^2}{|1-\Gamma_s\Gamma_L|^2}. \quad (21.3.8)$$

Как видим, она зависит от двух коэффициентов отражения: Γ_s и Γ_L . Из этого выражения легко получить известный результат: максимум поглощаемой мощности P_{av} по переменной Γ_L достигается при условии сопряженного согласования:

$$\Gamma_L^* = \Gamma_s. \quad (21.3.9)$$

Подставляя это условие в (21.3.9), получаем:

$$P_{av} = \frac{|\mathbf{b}_s|^2}{2(1-|\Gamma_s|^2)}. \quad (21.3.10)$$

Потери сопряженного рассогласования называется [21.16] отношение

$$\frac{P_L}{P_{av}} = \frac{(1-|\Gamma_s|^2)(1-|\Gamma_L|^2)}{|1-\Gamma_s\Gamma_L|^2}, \quad (21.3.11)$$

которое всегда не больше единицы. С другой стороны, если бы нагрузка была согласована с линией передачи, то есть $Z_L = Z_0$ и $\Gamma_L = 0$, то петля на графе рис. 21.3.2 была разомкнута и, как следует из (21.3.8), было бы

$$P_L = P_{z_0} = \frac{1}{2}|\mathbf{b}_s|^2. \quad (21.3.12)$$

Потери Z_0 -рассогласования называется отношение

$$\frac{P_L}{P_{z_0}} = \frac{1-|\Gamma_L|^2}{|1-\Gamma_s\Gamma_L|^2}, \quad (21.3.13)$$

которое может быть больше единицы; например, при $\Gamma_L^* = \Gamma_s$:

$$\frac{P_L}{P_{z_0}} \Big|_{\Gamma_L^* = \Gamma_s} = \frac{1-|\Gamma_s|^2}{(1-|\Gamma_s|^2)^2} = \frac{1}{1-|\Gamma_s|^2} \geq 1. \quad (21.3.14)$$

Неопределенность из-за многократных отражений между генератором и головкой

Предположим, что измеритель мощности идеально зарегистрировал мощность P_L , поглощенную чувствительной поверхностью мощностного сенсора. На самом деле, задача заключается в измерении мощности $P_{z_0} = \frac{1}{2}|\mathbf{b}_s|^2$, отдаваемой «источником» в согласованную с линией передачи нагрузку, или мощности $P_{av} = \frac{|\mathbf{b}_s|^2}{2(1-|\Gamma_s|^2)}$, отдаваемой в сопряженно-согласованную с «источником» нагрузку (чаще – первое). Казалось бы, что для пересчета измеренной мощности P_L в одну

или другую названные мощности нужно использовать соотношения (21.3.11) или (21.3.13). Как следует из этих соотношений, для этого нужно знать оба коэффициента отражения: $\Gamma_S = |\Gamma_S| e^{i\varphi_S}$ и $\Gamma_L = |\Gamma_L| e^{i\varphi_L}$, то есть 4 скалярные величины: модули $|\Gamma_S|, |\Gamma_L|$ и фазы φ_S, φ_L .

Незнание или неполное знание этих величин или некоторых из них является одним из источников ошибки измерения мощности. В отношении информации об этих величин могут быть различные варианты.

1) Неизвестны ни модули, ни фазы коэффициентов отражения $|\Gamma_S|, |\Gamma_L|$. В этом случае пришлось бы полагать знаменатель в правых частях (21.3.11), (21.3.13) равным единице, что привело бы к относительной ошибке оценки мощности, могущей принимать как угодно большое значение. Поэтому данный случай никогда не допускается.

2) Известны только модули $|\Gamma_S|, |\Gamma_L|$, но не фазы коэффициентов отражения. В этом случае диапазон неопределенности оценки знаменателя в правых частях (21.3.11), (21.3.13) дается максимумом и минимумом для худшего случая: $(1 \pm |\Gamma_L| |\Gamma_S|)^2$ [21.16]; это означает, что соответствующая составляющая относительной ошибки оценки мощности принимает значения в примерном диапазоне $\pm 2|\Gamma_S| \cdot |\Gamma_L|$.

3) С некоторыми точностями известны и модули $|\Gamma_S|, |\Gamma_L|$, и фазы φ_S, φ_L (достаточно знать сумму фаз $\varphi_S + \varphi_L$), например, в результате исследования цепи детектирования на векторном анализаторе цепей. Тогда соответствующая составляющая ошибки оценки мощности определяется значениями амплитуд и фаз и ошибками их задания.

4) Амплитуды и фазы коэффициентов отражения являются результатом применения скользяще-винтового тюнера [21.16], встроенного в конструкцию головки детектора. Этот вариант редко используется в современных головках, поскольку остаточное отражение от винтов не является незначительным в сравнении с отражением от головки.

5) Модули и фазы коэффициентов отражения неизвестны, но задан так называемый калибровочный коэффициент головки. Этот случай рассмотрен ниже.

Калибровочный коэффициент мощностной головки и КПД

Мощностная головка почти идеально собирает всю мощность, рассеянную на эффективной нагрузке, представленной детекторным элементом и его схемой. Но выше мы уже видели, что несовершенное согласование означает, что только часть $(1 - |\Gamma_L|^2)$ падающей мощности поглощается нагрузкой. Однако даже не вся эта часть окончательно достигает детекторного элемента, поскольку имеются некоторые потери в стенках и некоторая часть излучается в пространство, или стекает в проходные элементы конструкции. *К.п.д. мощностной головки η_e* , определяемый как отношение собранной мощности на детекторном элементе к полной рассеянной на нагрузке мощности, таким образом, уменьшается (кроме того, как мы видели, отраженная от нагрузки волна, дойдя до «источника», отражается от него, и вливается с некоторой фазой на чувствительную поверхность). Из падающей мощности P_i часть $(1 - |\Gamma_L|^2)P_i$ не отражается, но только часть поглощается детекторным элементом. Результирующее отношение падающей мощности к поглощенной элементом называется *калибровочным коэффициентом K_e* , который, таким образом, равен:

$$K_e = \eta_e (1 - |\Gamma_L|^2). \quad (21.3.15)$$

Производители мощностных головок обычно снабжают их таблицей калибровочных коэффициентов на специальных частотах (эти коэффициенты должны трассироваться до национальных стандартов), и современные мощмеры могут быть установлены корректно как по к.п.д, так и по потере рассогласования через единственный коэффициент K_e .

Пусть целью измерения является *согласованная мощность* P_{Z_0} , падающую на головку мощность обозначим P_i , результат измерения до коррекции (*кажущаяся мощность*) – P_i . Коррекция с помощью калибровочного коэффициента дает

$$P_i = \frac{P_M}{K_e}. \quad (21.3.16)$$

С другой стороны, с помощью (21.3.6) получаем:

$$P_i = \frac{1}{2} |\mathbf{a}_L|^2 = \frac{|\mathbf{b}_s|^2}{2|1 - \Gamma_s \Gamma_L|^2}, \quad (21.3.17)$$

и подставляя сюда правое равенство (21.3.12), имеем:

$$P_i = \frac{P_{Z_0}}{|1 - \Gamma_s \Gamma_L|^2}. \quad (21.3.18)$$

Сравнивая (21.3.18) с (21.3.16), видим, что даже после применения коррекции к кажущейся мощности P_M с помощью калибровочного коэффициента, лучшей оценкой мощности P_{Z_0} является величина P_i , полученная по (21.3.16); эта оценка имеет неопределенность в диапазоне худших случаев: $(1 \pm |\Gamma_L| |\Gamma_s|)^2$, то есть относительная ошибка принимает значения в примерном диапазоне $\pm 2|\Gamma_s| \cdot |\Gamma_L|$.

Поскольку мы получили те же неопределенность и диапазон относительной ошибки оценки, что и в **21.3.2 2)**, возникает вопрос: зачем нужно знать и применять калибровочный коэффициент K_e ? Ответ состоит в том, что знание K_e , определяемого заводом-изготовителем мощностной головки, избавляет пользователя от необходимости проведения непростого калибровочного определения модулей $|\Gamma_s|, |\Gamma_L|$ коэффициентов отражения.

АНАЛИЗАТОРЫ ЦЕПЕЙ (по книге т. 1 и по Балло)



Предисловие

Исторические замечания

Проблема аппаратного анализа цепей, предназначенных для работы в СВЧ диапазоне волн, возникла сразу же, как появились генераторы СВЧ-сигналов, то есть в первой половине XX века. Техника анализа быстро развивалась под воздействием актуальных стимулов, от ручных измерений отдельных параметров единичных элементов тракта – до автоматизированной индикации комплексного коэффициента отражения в классе устройств СВЧ – в 50-е годы XX столетия (см., например, [1.13, 1.14]). Исторической вехой стала разработка и совершенствование классического рефлектометра, для чего понадобились устройства сепарации волн, в частности, направленные ответвители (НО); с этого началась самостоятельная история развития устройств сепарации волн – НО и отражательных мостов СВЧ (векторных и скалярных), которая продолжается на патентном уровне до сих пор. Затем последовали этапы развития, связанные с миниатюризацией и сверхминиатюризацией компонент, появлением интегральных схем, развитием компьютерной техники, импульсной техники и т.д.

Современный ВАЦ как представитель четвертого поколения радиоизмерительных приборов

К настоящему времени анализаторы цепей стали важнейшим измерительным средством характеристики высокочастотных и сверхвысокочастотных компонент и устройств. Современный векторный анализатор цепей СВЧ относится к четвертому поколению радиоизмерительных приборов; позволяет вести автоматическую регистрацию амплитудных и фазовых параметров цепей в частотной области (панорамно по частоте, иногда от нуля до 110 ГГц) и во временной области; управляется внутренним или внешним компьютером; тестирует в режиме малых или больших сигналов линейные и нелинейные цепи различной степени интегрированности, в том числе – цепи с преобразованием частоты, с задержкой, нестационарные, параметрически управляемые и т.д.; вычисляет параметры рассеяния и другие характеристики цепей; применяет модулированный или немодулированный зондирующий сигнал; тестирует пассивные, активные и смешанные цепи, оснащенные разъемами, или непосредственно на подложке, а также полуоснащенные разъемами, например, антенны;

автоматически представляет и сохраняет характеристики цепей в различных форматах, масштабах и подробностях – в линейном и децибельном форматах, в декартовом и полярном базисах, с дискретом до долей Гц – в частотной области, и до 10 пс – во временной области; оснащен изоэкранным устройством механических и электронных измерительных калибровок; может быть включен в систему метрологического трассирования вплоть до национальных стандартов.

Направления развития

Тем не менее, возможности увеличения эффективности векторных и скалярных анализаторов цепей далеко не исчерпаны. Косвенным свидетельством этому является ежегодное появление на рынке измерительных приборов СВЧ все новых, более совершенных моделей анализаторов цепей ведущих мировых фирм. Можно предположить, что в связке с технологическими прорывами радиотехнического значения и рыночной конъюнктурой основными направлениями развития будут:

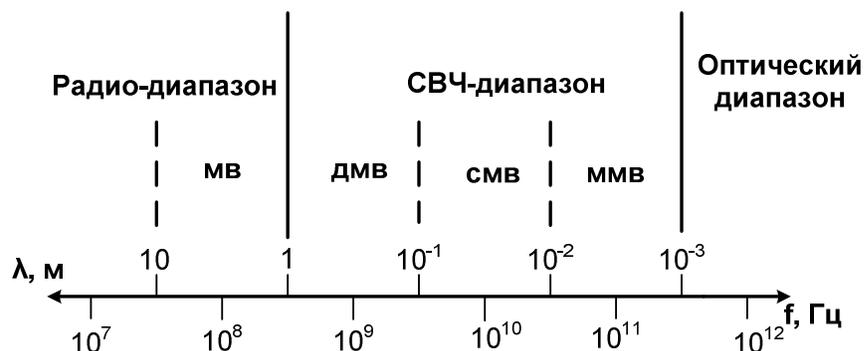
- 1) дальнейшая виртуализация системы, то есть передача выполняемых функций от радиотехнических устройств к компьютерам;
- 2) совершенствование измерительных функций во временной области;
- 3) совершенствование измерительных калибровок во временной области;
- 4) расширение возможностей измерения и интерпретации характеристик нелинейного взаимодействия зондирующего сигнала с устройством СВЧ;
- 5) дальнейшая автоматизация цикла измерений, вплоть до почти полного устранения ручных операций;
- 6) дистанционное управление работой прибора;
- 7) расширение возможностей включения прибора в измерительный комплекс без доработки интерфейсов и т.п.

Портовое представление цепей

Тракт СВЧ

Аппаратный анализ цепей является довольно специфической частью радиотехнических измерений. С технической точки зрения его удобно разделить на две части по критерию используемых для анализа частот: *низкочастотный анализ*, *анализ на СВЧ*. Далее, за редким исключением, имеется в виду аппаратный анализ цепей на СВЧ. Ключевые слова для этой темы: измерения на СВЧ, электродинамика СВЧ, теория и техника СВЧ, тракт СВЧ.

К *диапазону сверхвысоких частот* (СВЧ) обычно относят волны частотой от 300 МГц до 300 ГГц, то есть длиной волны от 1 мм до 1 м. Этот диапазон условно делят на поддиапазоны *дециметровых волн* (частоты от 300 МГц до 3 ГГц, длины волн от 0.1 м до 1 м), *сантиметровых волн* (частоты от 3 ГГц до 30 ГГц, длины волн от 0.01 м до 0.1 м), *миллиметровых волн* (частоты от 30 ГГц до 300 ГГц, длины волн от 0.001 м до 0.01 м). Место диапазона СВЧ в общем спектре электромагнитных волн показано на рис. 1.1.1.



Диапазоны электромагнитных волн

В дальнейшем основное внимание будет уделено измерениям в дециметровом и сантиметровом поддиапазонах, т.к. техника миллиметровых волн отличается большой специфичностью и требует отдельного рассмотрения. С другой стороны, по практическим причинам, рабочий диапазон *анализаторов цепей* (АЦ), которым посвящена эта глава, обычно шире указанного объединения поддиапазонов за счет спектрального участка *метровых волн* (частоты от 30 МГц до 300 МГц, длины волн от 1 м до 10 м), который формально не относится к диапазону СВЧ, а является частью диапазона УКВ. Иногда нижняя частота рабочего диапазона АЦ еще ниже и доходит до 10 МГц и менее.

Свойства электромагнитного поля существенно зависят от характера его изменения во времени; для модели монохроматических волн – от частоты волны. Особенности техники измерений на СВЧ можно условно разделить на особенности излучения, распространения, отражения и поглощения СВЧ-волн и особенности СВЧ-техники применительно к измерениям. Освоение СВЧ-диапазона началось позже, чем соседних диапазонов – только в 30-х годах двадцатого века, что объясняется особенностями СВЧ-волн, к важнейшим из которых относятся:

- Соизмеримость длины волны λ и характерного размера приборов, неоднородностей тракта и других, что делает малоэффективным применение обычных линий передачи и колебательных контуров. Традиционные элементы аппаратуры – соединительные проводники, резонансные контуры, конденсаторы, катушки индуктивности и т.д. становятся распределенными системами. Вместо соединительных проводников используют линии передачи конечной длины, колебательные системы представляют собой объемные резонаторы; сосредоточенные компоненты заменяются различными неоднородностями в линии передачи, расчет которых проводят электродинамическими методами.
- Соизмеримость периода колебаний T и времени пролета носителей заряда в активной области прибора τ . Проявляющаяся в этих условиях инерция носителей заряда нарушает нормальную работу активных приборов, эффективно функционирующих в радиодиапазоне.
- Малая энергия кванта электромагнитного поля $E_k = hf$ по сравнению с энергией теплового движения частиц, заданной в единицах $E_T = kT$, при обычной температуре, исключает проявление и возможное использование квантовых эффектов, успешно используемых в оптическом диапазоне. Например, на частоте $f = 20$ ГГц $E_k = 1.33 \cdot 10^{-23}$ Дж, а при температуре $T = 290$ К $E_T = 4 \cdot 10^{-21}$ Дж, то есть их отношение $E_T / E_k = 300$. Интересно отметить, что этот вывод меняется при очень низких температурах, например, на той же частоте, но температуре жидкого гелия $T = 4.2$ К отношение $E_T / E_k = 4.34$, а при частоте 50 ГГц даже $E_T / E_k = 1.74$.

Для реализации генерации, усиления и преобразования электромагнитных колебаний СВЧ-диапазона необходимо было разработать специальные методы, что потребовало значительных усилий многочисленных коллективов инженеров и ученых. В частности, при переходе на измерения в диапазоне СВЧ потребовалось, в ряде случаев, разработать специальные измерительные методы и средства. Измерения на СВЧ, по сравнению с традиционными измерениями на более низких частотах, имеют свои особенности:

- Ряд проблем вызывает подключение измерительной аппаратуры к исследуемому устройству и элементам калибровки. Пришлось разработать элементы подключения измерительных датчиков и преобразователей к соединительным линиям, вообще, в большой мере заново создавать измерительный тракт.
- Излучение электромагнитных волн из отверстий и щелей в измерительной аппаратуре приводит к потерям и искажениям сигналов. Требуется тщательное экранирование измерительных элементов и датчиков, что, однако осложняет введение измерительных элементов в исследуемые объекты.

- Отражения на соединениях, неоднородностях, сгибах, концах линий и т.д. (*несогласованность*) вызывают искажение сигналов, ухудшают точность измерений, требуют изошренной калибровки измерительного средства. Вообще, *проблема согласования* «нагрузок» в диапазоне СВЧ настолько всеобъемлюща, что иногда высказывается мнение, что «проблема оптимального конструирования тракта СВЧ сводится к проблеме согласования» [1.1].

- Неоднозначность отсчета напряжений и токов на СВЧ заставляет переходить к мощности как единице уровня сигналов.

- В технике СВЧ необходимо измерять новые параметры – КСВ, коэффициент отражения и т.д. Вообще, в рамках матричной характеристики цепей приходится переходить от матриц, связанных с импедансами, токами и напряжениями, к матрицам рассеяния; появляются также новые виды графов – *потокосые графы*, и т.д.

- Измерениям на СВЧ свойственны ограничения технического порядка, связанные с несовершенством аппаратуры на этих частотах. В частности, трудно или невозможно добиться достаточно высоких добротностей резонаторов, трудно создавать измерительные генераторы достаточной мощности со сверхширокополосной перестройкой, трудно обеспечить высокую частотную стабильность генераторов, трудно создавать измерительные усилители со стабильными метрологическими характеристиками и т.д.

- В диапазоне СВЧ ограничен выбор невзаимных устройств, позволяющих разделять волны разных направлений и улучшать согласование трактов. На малых частотах построение невзаимных устройств облегчается возможностями различных соединений в цепях, в диапазоне СВЧ, с его полевой структурой волн, приходится применять специальные волновые среды (например, ферриты) или комбинации линий передачи с полевой связью.

В силу перечисленных особенностей измерений на СВЧ, для решения каждой конкретной измерительной задачи в этом диапазоне часто приходится индивидуально подбирать или заново разрабатывать свою собственную методику, а в ряде случаев и уникальную измерительную аппаратуру. Всегда требуется предварительный теоретический и расчетный анализ измерительной схемы, особенно в отношении калибровки и оценки погрешностей измерения.

Вышеперечисленные особенности СВЧ-волн, особенно первые две, могут создать впечатление, что сам диапазон «невыгоден» для радиоэлектронных приложений и в измерениях на СВЧ, в частности, в анализе СВЧ-цепей нет большой необходимости. Это впечатление в корне неверно: практически диапазон СВЧ очень важен, а в некоторых случаях – незаменим, для целого ряда приложений, вследствие важных преимуществ СВЧ-волн. Эти преимущества таковы:

- Как известно, информационная емкость канала связи пропорциональна произведению полосы пропускания канала Δf на отношение сигнал-шум по мощности P_S / P_N . Полосу пропускания можно оценить как 1-10% от несущей частоты f_0 , следовательно, чем выше несущая частота, тем выше информационная емкость канала связи. Поэтому в СВЧ-диапазоне информационная емкость канала, в частности, скорость передачи информации, при том же отношении сигнал-шум, в сотни и тысячи раз больше, чем в радиодиапазоне. Но этот выигрыш еще больше вследствие минимума шумов атмосферного, космического и искусственного происхождения в диапазоне 1-10 ГГц, что повышает отношение сигнал-шум.

- С увеличением частоты возрастают возможности направленной передачи электромагнитной энергии, так как они зависят от отношений диаметра антенны к длине волны. В диапазоне СВЧ эти возможности, важные для многих приложений, значительно выше, чем в радиодиапазоне.

- В целом земная атмосфера весьма прозрачна в диапазоне СВЧ; это является ценным энергетическим фактором для многочисленных радиоэлектронных приложений.

- Ряд веществ проявляет селективно-частотное (избирательное) поглощение СВЧ-излучения, особенно при низких температурах, вследствие полосового, линейчатого характера спектра поглощения на этих частотах, а также, в ряде случаев, высоких градиентов поглощения по частоте. Это создает ряд возможностей для создания оригинальных устройств техники СВЧ.

Отмеченные преимущества диапазона обусловили широкое применение волн СВЧ в различных областях науки и техники. Перечислим некоторые из них:

- Радиолокация – первая исторически и до сих пор наиболее важная область применения колебаний СВЧ. Эта область в последние десятилетия испытала большой прогресс, как в военном, так и научном и народно-хозяйственном аспектах. Многообразие объектов локации (сосредоточенные: наземные, надводные, воздушные, космические; распределенные: элементы поверхности Земли, облака, полярные льды, северные сияния, метеорные ионизированные следы, планеты Солнечной системы и т.д.), многообразие способов формирования и обработки радиолокационных сигналов (когерентная радиолокация, радиолокация бокового обзора, сжатие сигналов и т.д.), многообразие носителей радиолокационной техники (наземные, надводные, воздушные, космические) и степени их подвижности, многообразие антенных систем управления лучами (одиночные антенны, синтезированные антенны, пассивные и активные антенные решетки) и видов их сканирования – все это и многое другое предопределило огромное поле для создания разнообразных устройств и систем техники СВЧ, нуждающихся в их тестировании и измерении.

- Радиорелейная связь, обеспечивающая передачу пакетов телефонных, телевизионных каналов и профессиональной информации различного назначения на различные, в том числе – большие расстояния с высокой скоростью передачи.

- Кабельная связь, в том числе распределенные системы вещательного телевидения, промышленного и охранного телевидения.

- Космическая связь, то есть передача больших объемов информации через спутники связи (непосредственное телевизионное вещание, связь с космическими кораблями и орбитальными станциями, с автоматическими космическими аппаратами).

- Радионавигация.
- Радиоастрономия (прием и анализ электромагнитного излучения космических объектов).
- Радиоразведка и радиоконтроль.
- Экспериментальная физика (нагрев и диагностика плазмы, молекулярная спектроскопия газов, жидкостей и твердых тел).
- Мазеры и их применение.
- Метрология (атомные стандарты времени и частоты).
- Технология (использование СВЧ-излучения для нагрева различных материалов с целью их сушки и полимеризации, приготовления пищи и т.п.).
- Медицина и биология (воздействие СВЧ-колебаний на биологические объекты с целью изучения и коррекции процессов в них, анализ собственного излучения объектов).

Продолжается быстрое расширение областей применения электромагнитных колебаний СВЧ-диапазона в науке и технике. Такое широкое разнообразие применений СВЧ-колебаний привело к появлению огромного числа видов и типов СВЧ-устройств, систем, цепей. Их число постоянно возрастает по мере развития технологии синтеза веществ (в частности, нанотехнологий, гетероструктурных технологий, полупроводниковых технологий), методов микроминиатюризации, создания интегральных схем и т.д. Все это стимулирует не только развитие методов расчета и проектирования с помощью ЭВМ устройств СВЧ, но и методов и приборной базы *аппаратного анализа цепей СВЧ*, чему и посвящена данная работа.

Метод декомпозиции

В диапазоне СВЧ квантовые эффекты еще пренебрежимы, поэтому закономерности протекания электромагнитных явлений в различных средах составляют предмет изучения *классической электродинамики*. Ее технические аспекты составляют область науки и техники, иногда называемую *технической электродинамикой*. В этой области в основном изучается и разрабатывается *техника СВЧ*, которая объединяет разделы радиотехники и радиоэлектроники, касающиеся вопросов анализа, синтеза и практической реализации *устройств СВЧ*.

Устройства СВЧ, соединенные отрезками *линий передачи*, составляют *тракт СВЧ*. Тракт СВЧ любой радиотехнической системы, в том числе анализатора цепей или анализируемого устройства, состоит из большого числа различных устройств СВЧ. К их числу относятся отрезки линий передачи, разъемы, изгибы и скрутки, согласующие устройства, фазовращатели, фильтры, делители мощности, направленные ответвители, переключатели и многие другие. Общим для этих и им подобных устройств является то, что они относятся к *устройствам с распределенными параметрами*. Геометрические размеры этих устройств сравнимы с длиной волны электромагнитных колебаний. Это определяет всю специфику расчета и проектирования устройств СВЧ, т.к. происходящие в них процессы имеют *волновой характер*.

Теория устройств СВЧ тесно связана с электродинамикой и включает в себя два больших раздела: *анализ устройств СВЧ* и *синтез устройств СВЧ*. Задача анализа состоит в изучении внешних характеристик устройств СВЧ, а также в определении этих внешних характеристик из решения соответствующей внутренней задачи методами прикладной электродинамики или из эксперимента. Задача синтеза устройств СВЧ состоит в определении структуры и геометрических размеров элементов устройства СВЧ по заданным его характеристикам.

Изучение внешних характеристик устройств СВЧ может производиться без конкретизации их внутренней структуры, что позволяет рассматривать устройство СВЧ как некий «черный ящик», имеющий определенное число выходящих из него линий передачи СВЧ. Каждая из этих линий передачи также является устройством с распределенными параметрами, для которого неизменным является волновой характер электромагнитных процессов. Это приводит к необходимости фиксировать продольные координаты поперечных сечений линий передачи или, как говорят, фиксировать *опорные плоскости*. Относительно этих опорных плоскостей проводится отсчет фаз, а в некоторых случаях и амплитуд падающих и отраженных волн. Смещение опорных плоскостей вдоль входных линий передачи приводит к изменению внешних характеристик устройств СВЧ. В большинстве случаев во входных линиях передачи таких устройств единственной распространяющейся волной является волна основного типа. Остальные типы волн находятся в закритическом режиме, то есть быстро затухают при отдалении от неоднородности, где они могут возникнуть. Опорные плоскости устройства СВЧ стремятся расположить таким образом, чтобы амплитудами закритических волн в них можно было пренебречь.

Основой теоретического и расчетного анализа и проектирования СВЧ-устройств является *теория цепей СВЧ*. В силу упомянутого выше представления цепи в виде «черного ящика», эта теория основана на *матричном представлении* описания внешних параметров устройств, использует *метод декомпозиции* анализа сложных объектов и компьютерные методы их моделирования. Для расчета параметров базовых блоков *декомпозиционных схем* привлекаются методы *прикладной электродинамики* и *численные модели* систем автоматического проектирования (САПР) СВЧ-устройств. При практической реализации ключевую роль играют *экспериментальные методы* исследования, настройки и контроля характеристик и параметров спроектированных устройств. Измерения используют также в случаях, когда для моделирования отдельных узлов и блоков нет достоверных расчетных данных.

Порты

Радиосистемы СВЧ-диапазона (а также более высокочастотные) обычно можно представить в виде некоторых устройств, соединенных линиями передачи. Часть такой системы,

расположенную между начальным и конечным устройствами (например, между антенной и радиопередающим или радиоприемным устройством, называют *трактом СВЧ* (иногда *цепью СВЧ*). Подобный тракт осуществляет передачу электромагнитной энергии от передатчика к антенне, или от антенны к приемнику, или от генератора сигналов специальной формы ко входу некоторого устройства и т.п. Тракт СВЧ может, например, обеспечивать требуемый режим работы выходных и входных цепей передатчика или приемника, выполнять частотное и поляризационное разделение и объединение передаваемых сигналов и ряд других функций. Отдельные части тракта также могут называться цепями СВЧ. Наиболее распространенными элементами СВЧ-цепей являются отрезки линий передачи, переходные и стыковочные узлы между линиями различных типов, согласующие и настроечные элементы, сумматоры, делители и ответвители мощности, поляризационные устройства, фильтры, фазовращатели, коммутаторы и переключатели, невзаимные устройства с намагниченными ферритами и др.

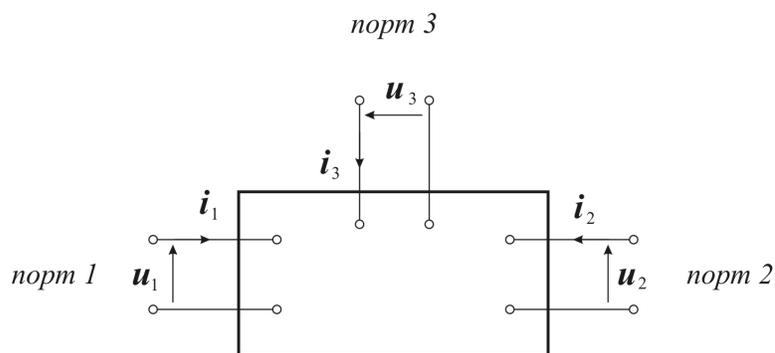
Процессы передачи электромагнитных сигналов в цепях СВЧ и в образующих их элементах весьма сложны. Можно было бы попытаться проанализировать эти процессы на основе соответствующих *краевых задач электродинамики*. Однако строгая постановка и решение таких задач возможны только для очень упрощенных форм и конструкций, почти всегда выходящих за рамки реальных. А для применяемых на практике цепей СВЧ из-за их конфигурационной сложности решение краевых задач в строгой постановке в настоящее время практически невозможно.

Некоторую помощь в решении проблемы передачи электромагнитных сигналов в сколь-нибудь сложных цепях СВЧ оказывает всегда применяемый на практике *метод декомпозиции* (разбиения). Он состоит в том, что тракт СВЧ или цепь СВЧ разбивается на ряд *элементов*, которые анализируются независимо, но теперь также нужно адекватно учесть взаимодействия между элементами через распространяющиеся между ними волны. При этом каждый такой элемент рассматривается как независимая электродинамическая система.

Постановка и решение краевых электродинамических задач, соответствующих отдельным элементам, существенно проще, чем для всего устройства в целом. Используя или решение электродинамической задачи или результаты экспериментального исследования, если подобное решение получить не удастся, для каждого выделенного элемента строят такое описание, которое позволяет находить влияние этого элемента на передаваемые электромагнитные сигналы. Обычно описание элементов цепи представляют либо формально-математически с помощью матриц специального вида (матрицы рассеяния, или матрицы передачи и т.п.), либо в виде *эквивалентной схемы*, состоящей из отрезков эквивалентной линии передачи, в которую тем или иным способом включены сосредоточенные элементы L, C, R и трансформаторы. Имея подобные универсальные описания всех элементов тракта или цепи СВЧ, можно определить все требуемые характеристики тракта или цепи.

Обычно при построении *математической модели* тракта или цепи СВЧ для упрощения анализа отрезки линий передачи, соединяющие входящие в этот тракт устройства заменяют отрезками *эквивалентной линии*, а устройства рассматриваются как некоторые *многополюсники*.

Сделаем одно терминологическое замечание. Понятия полюсов, клемм и т.п., так хорошо подходившие к концепции «черного ящика» на достаточно низких частотах, где цепи строятся на сосредоточенных элементах, а соединения производятся с помощью проводов, становятся неудобными в диапазоне СВЧ, где по крайней мере линии передачи суть распределенные элементы, а соединения отрезков этих линий производятся с помощью коаксиальных разъемов, фланцевых соединений волноводов и т.п. В этом диапазоне больше подходит понятие *порта*, которым и пользуются во всей научно-технической радиоэлектронной литературе, кроме русскоязычной. С точки зрения формально-математического описания порт полностью эквивалентен паре полюсов или клемм, но физическая интерпретация этого понятия более прозрачна. В этой работе мы будем пользоваться именно этим понятием.



Токи и напряжения в трехпортовом устройстве



Комплексные амплитуды токов и напряжений в двухпортовом устройстве

Математическая модель тракта, портовое представление

Теперь, повторяя вышесказанное, можно заметить, что математическая модель тракта СВЧ включает некоторые устройства (элементы) в виде N -портовых цепей ($N=1, 2, 3, \dots$), соединенные отрезками эквивалентных линий передачи, которые, при необходимости или для удобства, также можно представлять как двухпортовые ($N=2$) цепи. Такое представление тракта или многокомпонентной цепи назовем *портовым представлением*. Следует еще раз подчеркнуть, что портовое представление цепи есть некоторое упрощение реальных электродинамических процессов. Это видно хотя бы из того, что процессы в эквивалентной линии описываются скалярными функциями (напряжением \dot{U}_m и током \dot{I}_m), зависящими лишь от продольной координаты z , в то время как на самом деле эти процессы должны быть представлены векторными функциями \vec{E}_m, \vec{H}_m , являющимися решением соответствующей электродинамической задачи. Более того, такую упрощенную модель отрезка линии передачи можно использовать лишь в диапазоне частот, где соблюдается одноволновый режим работы линии. Кроме того, эта модель непригодна для определения ряда характеристик цепи СВЧ, например, таких как максимальная мощность, передаваемая по тракту СВЧ, или величина взаимной связи между элементами цепи СВЧ, построенной на отрезках линий передачи открытого типа, и некоторых других.

Принцип зондирования

Обоснование принципа зондирования (зондирующие сигналы, тестируемая цепь и т.д.)

Каким принципиально способом возможно и необходимо аппаратно анализировать цепи СВЧ? Чтобы понять, что этот вопрос не лишен содержания, обратимся к более низким частотам, на которых цепи, в основном, состоят из локальных элементов, соединенных проводниками. В этом случае существует не одна возможность аппаратного изучения структуры и свойств процессов, протекающих в цепи: можно, подключаясь к узлам цепи, измерять их потенциалы относительно некоторой точки (например, «земли») по «высокой» частоте и по ПТ; можно измерять активную и реактивную составляющую сопротивлений между парами точек, а также разность потенциалов между этими точками; можно измерять токи в ветвях цепи, подключаясь к ним трансформаторно или через размыкание ветвей; можно подавать на узлы цепи выходной РЧ сигнал от генератора и измерять потенциалы в узлах; и т.д.

Вследствие вышеперечисленных особенностей диапазона СВЧ и фактически вынужденного портового представления тракта и его элементов, множество возможностей аппаратного анализа устройств СВЧ суживается до одной-единственной: необходимо подавать на порты исследуемого устройства сигналы специальной формы от некоторого генератора и, принимая в то же время сигналы с тех же и других портов этого устройства, анализировать их параметры. Такой принцип аппаратного анализа цепей назовем *принципом зондирования*, подаваемые на исследуемую цепь сигналы – *зондирующими сигналами* (иногда в литературе, по аналогии с физиологией, их называют *стимулирующими*), соответствующие генераторы – *генераторами зондирующих сигналов*, саму исследуемую цепь – *зондируемой*, или *тестируемой*, или *измеряемой*, и т.д.

Зондирование однопортовой цепи

Рассмотрим принцип зондирования несколько более конкретно. Пусть сначала тестируемая цепь – *однопортовая* (иногда такие цепи называют *нагрузками*). На рис. 1.3.2 условно показано взаимодействие волн при зондировании такой цепи (такого рода иллюстрации иногда называют *оптической аналогией* зондирования).

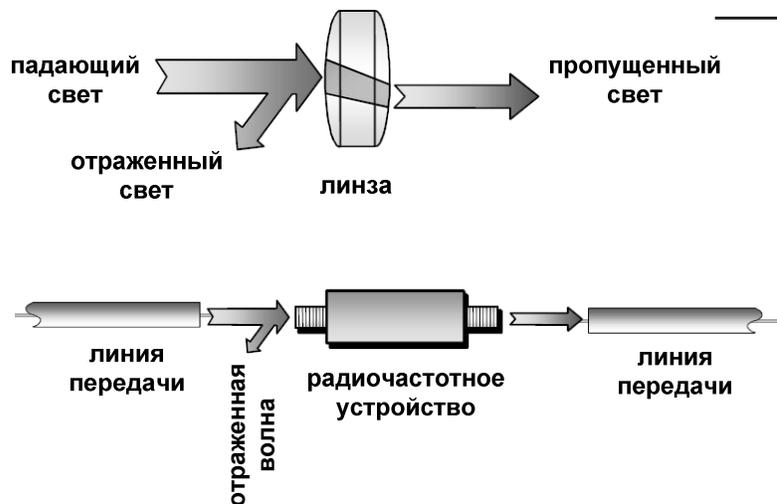


Рис. 1.3.1. Оптическая аналогия зондирования двухпортовой цепи

Падающая волна (зондирующий сигнал) от генератора подается на единственный порт цепи по аналогии со световым лучом, падающим на образец; в результате электродинамического взаимодействия внутри «черного ящика» цепи часть мощности волны рассеивается внутри цепи, остальная часть выходит в виде волны, распространяющейся в обратном направлении; эта волна интерпретируется как отраженная от того же порта. Часть этой волны ответвляется специальным

устройством в приемник, с которого начинается обработка принятой волны, приводящая к оценке *коэффициента отражения* тестируемой цепи и эквивалентных ему характеристик.

Классический рефлектометр

Фактически мы описали устройство и действие *классического рефлектометра*, упрощенная скелетная схема которого приведена на рис. 1.3.2.

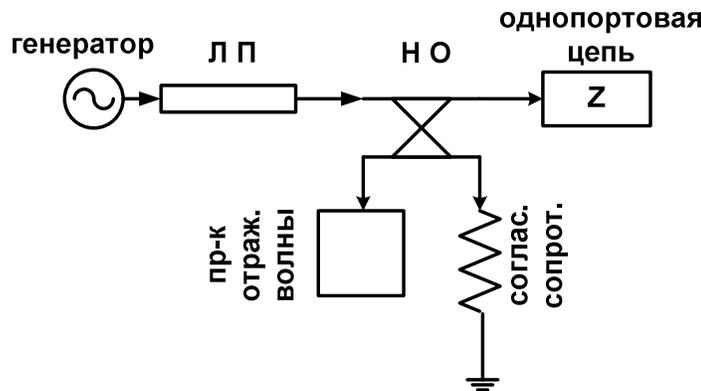


Рис. 1.3.2. Упрощенная структурная схема рефлектометра

Зондирование двухпортовой цепи, два способа двустороннего зондирования (илл. Балло, с. 23)

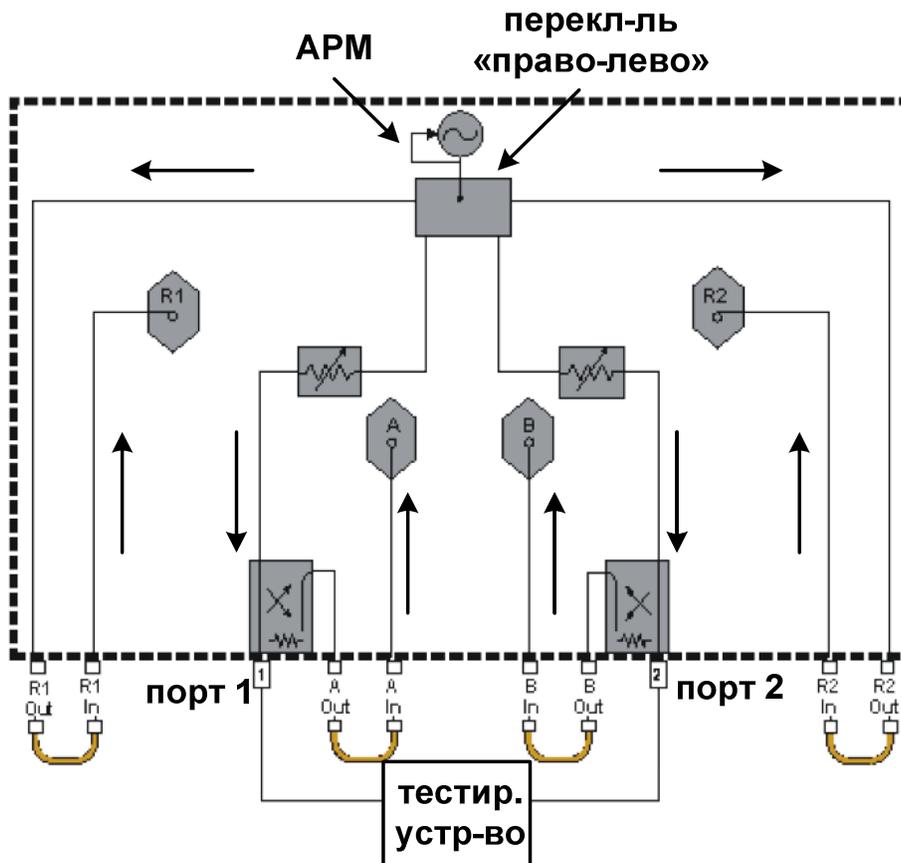
Теперь рассмотрим случай, когда тестируемая цепь имеет два порта. Реализация принципа зондирования теперь состоит из двух частей. Сначала зондирующий сигнал подается на первый порт тестируемого устройства (ТУ). Оптическая аналогия этого опыта показана на рис. 1.3.1.

Падающая волна (зондирующий сигнал) от генератора подается на первый порт цепи по аналогии со световым лучом, падающим на образец, в результате электродинамического взаимодействия внутри «черного ящика» цепи часть мощности падающей волны рассеивается внутри цепи, остальная мощность делится на две части: одна выходит через первый порт и движется в обратном направлении, она называется *отраженной волной*, т.к. интерпретируется как волна, отраженная от первого порта; другая выходит через второй порт и движется в прямом направлении (то есть том же, что и падающая волна от генератора), она называется *пропущенной волной*. Часть отраженной волны ответвляется устройством направленного ответвления в первый приемник, с которого начинается обработка отраженной волны, приводящая к оценке коэффициента отражения цепи со стороны первого порта и эквивалентных ему характеристик. Часть пропущенной волны (или вся она) ответвляется во второй приемник, с которого начинается обработка пропущенной волны, приводящая к оценке коэффициента пропускания цепи от первого порта до второго тестируемой цепи и эквивалентных ему характеристик.

Если бы рассматриваемое двухпортовое ТУ было симметричным, то были бы равны коэффициенты отражения от первого порта и от второго порта, а также коэффициенты пропускания от первого порта ко второму и от второго порта к первому. В этом случае можно было бы ограничиться зондированием со стороны первого порта и не зондировать ТУ со стороны второго порта. Однако, во-первых, существуют несимметричные двухпортовые цепи; во-вторых, даже по замыслу, по конструкции симметричная цепь на самом деле в некоторой степени несимметрична. Это объясняется случайными отклонениями геометрических факторов, электромагнитных параметров проводников и диэлектриков, портов (разъемов) и т.д. Поэтому часто необходимо измерить те же характеристики – коэффициент отражения и коэффициент пропускания – также и со стороны второго порта ТУ. Это можно сделать одним из двух способов: или «перевернуть» ТУ, то есть подключить второй порт к генератору, а первый – к приемнику пропускания; или поставить переключатели каналов генератора, приемника

отражения и приемника пропускания. Упрощенная скелетная схема зондирования, реализующего второй способ, показана на рис. 1.3.3.

Скелетная схема анализатора цепей



Упрощенная скелетная схема зондирования двухпортовой цепи

Прохождение сигналов при положении переключателя «лево» и «право»

Зондирующий сигнал вырабатывается панорамным (по частоте) генератором, в качестве которого чаще всего выступает синтезатор частот. Для стабилизации мощности генератора он охвачен петлей *автоматической регулировки мощности* (АРМ). В положении переключателя «лево» зондирующий сигнал по одной ветви поступает на приемник сравнения **R1**, а по другой – через управляемый аттенюатор и НО-1 – на левый (первый) порт тестируемого устройства (ТУ). Отраженный от ТУ сигнал, несущий информацию о коэффициенте отражения первого порта ТУ (о параметре S_{11}), поступает на НО-1 и через его ответвительный порт – на приемник отраженного сигнала **A**. Пропущенный ТУ сигнал, несущий информацию о коэффициенте пропускания ТУ (о параметре S_{21}), поступает на НО-2 и через его ответвительный порт – на приемник пропущенного сигнала **B**.

В положении переключателя «право» зондирующий сигнал по одной ветви поступает на приемник сравнения **R2**, а по другой – через управляемый аттенюатор и НО-2 – на правый (второй) порт ТУ. Отраженный от ТУ сигнал, несущий информацию о коэффициенте отражения второго порта ТУ (о параметре S_{22}), поступает на НО-2 и через его ответвительный порт – на приемник **B**, который в данном случае играет роль приемника отраженного от второго порта ТУ сигнала. Пропущенный от второго порта ТУ к первому сигнал, несущий информацию об обратном коэффициенте пропускания ТУ (о параметре S_{12}), поступает на НО-1 и через его

ответственный порт – на приемник А, который в данном случае играет роль приемника обратно пропущенного сигнала.

Проблема тестирования устройств с более чем двумя портами

В случае если ТУ имеет более двух портов, стратегия зондирования ясна: сигнал генератора по очереди подается на один из выбранных портов (не обязательно всех), а выходные сигналы с портов (включая порт, на который подается сигнал) подаются на соответствующие приемники.

Конкретизация задачи анализа цепи

Системно-информационная интерпретация зондирования, операторная формулировка

Посмотрим на принцип зондирования с более общей системно-информационной точки зрения (рис. 1.3.4). Пусть имеется объект общего вида, способный взаимодействовать с электромагнитными колебаниями. На «входы» объекта поступают электромагнитные волны (*входные воздействия*), в результате чего с «выходов» объекта исходят электромагнитные сигналы (*выходные отклики*). Возможны три задачи:

- Известны входные воздействия и характеристики объекта, необходимо найти выходные отклики. Это *прямая задача прохождения сигналов* через объект.
- Известны характеристики объекта и выходные отклики, необходимо найти входные воздействия. Это *обратная задача прохождения сигналов* через объект.
- Известны входные воздействия и выходные отклики, необходимо найти характеристики объекта. Это *задача идентификации* объекта.



Схематическое представление взаимодействия электромагнитных колебаний с радиотехническим объектом

Сказанное можно выразить в математических терминах [1.12]. Представим входные воздействия как векторы (точки, элементы) $\mathbf{q}_{\text{вх}}$ некоторого векторного пространства \mathbf{V} , то есть $\mathbf{q}_{\text{вх}} \in \mathbf{V}$, а выходные отклики – как векторы (точки, элементы) $\mathbf{q}_{\text{вых}}$ векторного пространства \mathbf{W} , то есть $\mathbf{q}_{\text{вых}} \in \mathbf{W}$. Тогда связь между входными воздействиями и выходными откликами можно выразить как

$$\mathbf{q}_{\text{вых}} = \mathbf{T} \mathbf{q}_{\text{вх}}, \quad (1.3.1)$$

где \mathbf{T} – *системный оператор*.

Чтобы задачи были полностью определены, необходимо указать область $D_{\text{вх}} \subset \mathbf{V}$ *допустимых входных воздействий* и область $D_{\text{вых}} \subset \mathbf{W}$ *допустимых выходных откликов*.

Названные выше задачи теперь имеют вид:

- Прямая задача прохождения: по заданным $\mathbf{q}_{\text{вх}} \in D_{\text{вх}}$ и \mathbf{T} найти $\mathbf{q}_{\text{вых}} \in D_{\text{вых}}$ на основе (1.3.1).
- Обратная задача прохождения: по заданным $\mathbf{q}_{\text{вых}} \in D_{\text{вых}}$ и \mathbf{T} найти $\mathbf{q}_{\text{вх}} \in D_{\text{вх}}$ согласно уравнению

$$\mathbf{q}_{\text{вх}} = \mathbf{T}^{-1} \mathbf{q}_{\text{вых}},$$

где \mathbf{T}^{-1} – обратный оператор.

- Задача идентификации объекта: по заданным $\mathbf{q}_{\text{вх}} \in D_{\text{вх}}$ и $\mathbf{q}_{\text{вых}} \in D_{\text{вых}}$ определить существенные для интересующих нас приложений свойства оператора \mathbf{T} и, возможно, отнести систему к определенному классу.

Совокупность системного оператора \mathbf{T} и областей $D_{\text{вх}}$ и $D_{\text{вых}}$ называется *математической моделью системы*. Классификацию систем проводят на основании существенных свойств их математических моделей.

Мы видим, что реализация принципа зондирования в аппаратном анализе цепей с общей системно-информационной точки зрения есть задача идентификации радиотехнических цепей.

Планирование измерения по тестированию устройства

Для конкретизации этой задачи требуется еще до проведения непосредственно измерений ответить на многочисленные вопросы, в частности:

- Какие характеристики тестируемого устройства необходимо определить (в зависимости от типа устройства)?
- Какой минимальный набор зондирующих сигналов требуется для определения этих характеристик?
- В каком состоянии и как обустроено должно быть испытываемое устройство во время его зондирования?
- С какой подробностью и точностью должны фиксироваться входные и выходные сигналы для получения характеристик устройства с необходимыми точностями, разрешениями и т.п.?

Не существует однозначных, раз и навсегда заданных ответов на эти вопросы. Приемлемые ответы зависят от планируемого использования испытываемых устройств, внешних требований к ним, их типа и структуры и многого другого. Следовательно, не существует однозначного решения задачи о структуре и алгоритме действия анализатора цепей.

Принцип панорамности

До сих пор принцип зондирования пояснялся на примере некоторого одного сигнала, подаваемого на порт ТУ. В действительности, принцип зондирования почти всегда применяется совместно с *принципом панорамности*, состоящего в том, что зондирование осуществляется последовательно набором сигналов с различными значениями некоторого параметра.

Для пояснения идеализируем ситуацию и считаем, что сигналы генератора монохроматичные, то есть игнорируем конечность полосы сигнала, помехи, шумы и т.д. Аналоговый монохроматический сигнал имеет вид:

$$s(t) = A_0 \cos(\omega t + \varphi), \quad (1.4.1)$$

то есть имеет аргументом время t и три параметра: амплитуду A_0 (или мощность $P = \frac{1}{2} A_0^2$), круговую частоту ω (или циклическую частоту $f = \omega / 2\pi$), начальную фазу φ .

Назовем панораму *простой*, если варьируется только один параметр сигнала, и *сложной*, если варьируется два или более параметров; *непрерывной*, если параметр меняется непрерывно, и *дискретной*, если задается только последовательность изолированных значений параметра; *односторонней*, если параметр пробегает значения в одном направлении (если не считать быстрого сброса к исходному значению) и *двусторонней*, если он пробегает значения сначала в одном, затем в другом направлении; *симметричной*, если у двусторонней панорамы пробеги в одном и другом направлениях одинаковы по времени и набору значений; *эквидистантной*, если

в простой дискретной панораме значения варьируемого параметра отстоят друг от друга на равные расстояния, и *неэквидистантной* – в противном случае.

По причинам теоретической и технической простоты чаще всего используется простая дискретная односторонняя эквидистантная панорама. В этом случае на порт ТУ поступает последовательность *цугов* сигнала вида (1.4.1) одинаковой длительностью T каждый, но с изменяющимся значением одного параметра от цуга к цугу. Очевидно, можно варьировать или частоту f , или мощность P ; варьировать начальную фазу φ бессмысленно, т.к. информативны только разности фаз в разных точках цепи; более того, нет смысла даже контролировать начальные фазы цугов.

При вариации частоты f (*частотная панорама*) генератор зондирующих сигналов вырабатывает последовательность сигналов

$$s_k(t) = A_0 \cos(2\pi f_k t + \varphi), \quad k = 0, 1, 2, \dots, K,$$

каждый длительностью T_f , причем

$$f_k = f_0 + k \cdot \Delta f, \quad k = 0, 1, 2, \dots, K,$$

$K + 1$ – число *частотных точек*,

Δf – *шаг перестройки частоты*,

$f_0 = f_{\text{нач}}$ – *начальная частота* диапазона перестройки (*частотной панорамы*),

$f_{\text{кон}} = f_0 + K \cdot \Delta f$ – *конечная частота* этого диапазона,

При вариации мощности P (*мощностная панорама*) генератор зондирующих сигналов вырабатывает последовательность сигналов

$$s_q(t) = A_{0q} \cos(2\pi f_0 t + \varphi), \quad q = 0, 1, 2, \dots, Q,$$

каждый длительностью T_p , причем

$$A_{0q} = A_{00} + q \cdot \Delta A_0, \quad q = 0, 1, 2, \dots, Q,$$

$Q + 1$ – число *мощностных точек*,

ΔA_0 – *шаг перестройки мощности*,

$A_{00} = A_{\text{нач}}$ – *начальная амплитуда* диапазона перестройки амплитуд (*мощностной панорамы*),

$A_{0\text{кон}} = A_{00} + Q \cdot \Delta A_0$ – *конечная амплитуда* этого диапазона.

Внимательный читатель заметит, что данное выше определение эквидистантной панорамы несовершенно. Действительно, при вариации мощности мы предполагали равномерную сетку амплитуд; но можно было взять равномерную сетку по мощности (то есть по квадрату амплитуд), или по логарифму мощности (то есть равномерную сетку в децибельном масштабе), или по какой-либо другой функции амплитуды и т.д. Ясно, что это разные сетки значений. Если их все считать эквидистантными, то определение эквидистантной панорамы будет таким: панорама эквидистантна, если она простая, дискретная и значения варьируемого параметра или любой однозначной монотонной функции от него отстоят друг от друга на равные расстояния. Такой подход делает эквидистантную панораму более гибкой, что особенно важно для мощностной панорамы.

Несмотря на удобство эквидистантной панорамы, часто применяют неэквидистантную панораму, особенно в мощностном варьировании, необходимость в котором возникает при тестировании нелинейных устройств. Иногда набор частотных или мощностных точек задают таблично. Такая необходимость может возникнуть, например, при тестировании полосового

фильтра, когда желательно подробнее изучить рабочую часть полосы, в отличие от остальной ее части; или при необходимости более подробного изучения рабочего диапазона входных мощностей усилителя, в отличие от области сжатия усиления.

Замечание. Хорошо известно, что идеализированная модель (1.4.1) представляет монохроматический сигнал (то есть имеет нулевую полосу частот), только если задана на бесконечном интервале времени. Рассматриваемые выше цуги имеют спектры с конечной полосой частот. Вообще говоря, это может осложнить интерпретацию результатов измерения, например, при частотной панораме. Но на данном этапе рассмотрения проблемы аппаратного анализа цепей мы отложим этот вопрос на будущее, предположив, что длительности цугов достаточно велики, чтобы их полоса, находящаяся к ним в обратной зависимости, была пренебрежима в сравнении с требуемым частотным разрешением анализа цепи.

Обсудим теперь необходимость панорамирования и даваемые им преимущества. Основные тенденции развития радиотехнической аппаратуры – уменьшение габаритов, расширение диапазонов рабочих частот – требуют увеличения объема измерения параметров и характеристик СВЧ элементов трактов. Старые методы измерения параметров цепей с помощью измерительных линий и мостовых измерителей полных сопротивлений могли работать только на фиксированных частотах, что существенно увеличивало трудоемкость измерений и чрезвычайно усложняло отработку и настройку СВЧ узлов. Автоматизация этих измерений и создание панорамной измерительной аппаратуры не только увеличивают производительность, но и повышают качество измерений, так как позволяют визуально наблюдать частотную характеристику измеряемого параметра в широком диапазоне частот в наиболее удобной форме (графики, цифры).

Панорамное наблюдение амплитудных и фазовых характеристик СВЧ устройств в широкой полосе частот позволяет значительно сокращать время их проверки, эффективно осуществлять их широкополосную настройку, способствует прогрессу проектирования СВЧ аппаратуры. В связи с этим особенное значение имеет разработка панорамных измерителей КСВ и ослаблений, измерителей S-параметров одно- и двухпортовых цепей.

Использование сигналов со сложными видами модуляции в радиосвязи, радиолокации, радионавигации, радиоскопии, радиоастрономии, радиоспектроскопии выдвигает определенные требования к устройствам их приема, передачи и формирования. В зависимости от назначения устройства и места включения его в аппаратуру необходимо контролировать модуль коэффициента передачи, электрическую длину, коэффициент отражения со стороны входа и выхода, или полное сопротивление – составляющие S-параметров. Развитие техники панорамного измерения S-параметров значительно стимулировало внедрение вычислительной техники в область радиоизмерений, так как только на базе измерителей S-параметров могут быть созданы автоматизированные прецизионные системы для измерения параметров СВЧ узлов. Измерители S-параметров обеспечивают автоматическое с высокой точностью измерение:

- комплексных коэффициентов передачи;
- КСВ;
- амплитудно-частотных характеристик;
- фазочастотных характеристик.

Они незаменимы в системах автоматического контроля измерения, управления технологическими процессами, при расчете и конструировании узлов и элементов СВЧ трактов.

Структура анализатора цепей

Как уже отмечалось, не существует однозначного решения задачи о структуре и алгоритме действия анализатора цепей. Тем не менее, в соответствии с логическими концентрерами **1.1 – 1.4:** особенностями диапазона СВЧ, портовым представлением тестируемых цепей, принципами зондирования и панорамности – структура анализатора цепей проясняется в основных чертах. АЦ должен иметь по крайней мере один генератор зондирующих сигналов; последние должны в

определенной последовательности направляться на порты ТУ (или определенным способом воздействовать на цепь, не оборудованную разъемами), а также, как будет пояснено ниже, в опорные каналы; сигналы с выходных портов ТУ (среди которых может быть и входной порт) должны направляться и ответвляться в приемники, где, возможно, частотно преобразуются, усиливаются, детектируются, оцифровываются; аналогичные преобразования совершаются над сигналами в опорных каналах; затем происходит математическая обработка выходных сигналов с целью вычисления необходимых функционалов: отношений амплитуд и разностей фаз измерительных и опорных сигналов, так называемых S-параметров тестируемой цепи и т.д.; кроме того, полученная информация приводится к виду, позволяющему отображать ее на экране монитора, сохранять, транслировать через внешний интерфейс и т.д.

Функциональные блоки АЦ (перечень)

• Структурная схема рис. 1.3.3 подчеркивает принципиальные элементы АЦ: генератор зондирующих колебаний (ГЗК), расщепители мощности, устройства направленного ответвления обратного сигнала, приемники с детекторами, компьютер с дисплеем. Она включает следующие основные блоки (слово «блок» здесь понимается не в конструктивном, а в функциональном смысле):

- Блок источников зондирующих сигналов
- Блок ответвления, разделения и перенаправления сигналов
- Блок приемников
- Блок координации и управления прибором (виртуальный)
- Блок измерительной калибровки и коррекции данных измерения
- Блок математической обработки сигналов и расчетов
- Блок отображения результатов анализа
- Блок внутренней автоматики.

Блок источников

Блок источников зондирующих сигналов включает один или два генератора зондирующих сигналов, выполненных, в современных условиях, в виде синтезаторов частот. Этот блок может включать систему стабилизации выходной мощности генераторов и другие необходимые устройства. Блок обеспечивает формирование и передачу в различных режимах зондирующих сигналов двух видов:

– цугов монохроматических колебаний в определенных (устанавливаемых) частотных точках или (а) с дискретным частотным панорамированием с устанавливаемыми частотным диапазоном и шагом, (б) с дискретным мощностным панорамированием с устанавливаемыми мощностным диапазоном и шагом; такой вид зондирующих сигналов необходим для характеристики анализируемых линейных и нелинейных цепей в частотной области;

– или (и) последовательности коротких радиоимпульсов эталонной формы с СВЧ заполнением на точной временной сетке; возможно поимпульсное частотное панорамирование заполнения; такой вид зондирующих сигналов необходим для характеристики анализируемых линейных и нелинейных цепей во временной области.

Главные требования к зондирующим сигналам: высокая кратковременная частотная стабильность, определяемая фазовыми шумами генераторов и системой управления ими; стабильность мощности цугов и импульсов.

Блок ответвления, разделения и перенаправления сигналов не локализован пространственно, его элементы находятся в различных трактах. Как следует из названия блока, его элементы выполняют функции:

– ответвления части падающей волны, до ее попадания в порт ТУ, в опорный канал, с целью дальнейшего обеспечения относительных измерений амплитуд и, возможно, фаз; эту функцию выполняют расщепители (сплиттеры), или делители мощности, или направленные ответвители, имеющие взаимные преимущества и недостатки; в двухпортовом АЦ таких ответвителей необходимо два: при зондировании первого порта и при зондировании второго порта;

– разделения падающей (идущей к порту ТУ) и отраженной (идущей от порта ТУ) волн на входе ТУ, с тем чтобы направить отраженную компоненту в приемник отражения; в связи со специальным требованием к так называемой направленности, эту функцию выполняют направленные ответвители или отражательные направленные мосты, имеющие взаимные преимущества и недостатки; в двухпортовом АЦ таких разделителей необходимо два: при зондировании первого порта и при зондировании второго порта;

– перенаправления падающей волны с первого порта ТУ на второй и наоборот, для обеспечения последовательного зондирования двухпортовой цепи с обоих портов; эту функцию выполняет переключатель, к которому предъявляется ряд требований (высокая скорость переключения, большой ресурс, высокая повторяемость коэффициента передачи и его одинаковость для обоих положений переключателя); этим требованиям удовлетворяют твердотельные, но не механические переключатели;

– разделения сигнала единого гетеродина на приемники в случае применения в них супергетеродинирования.

Блок приемников содержит 4 приемника: 2 приемника опорных каналов и 2 приемника измерительных каналов (по одному на зондирование каждого порта ТУ). Иногда производитель АЦ ограничивается одним опорным приемником, что хотя и дает некоторую экономию оборудования и стоимости, ограничивает возможности измерительной калибровки прибора, подвергает измерение дополнительной ошибке за счет неидеальной повторяемости переключателя и т.д. Применяются приемники одного из двух принципиально различающихся типов:

– первый тип – детекторный приемник с широкополосным диодным детектированием, его входная полоса частот должна охватывать весь диапазон панорамного изменения частоты зондирующих сигналов; важнейшие преимущества такого детектора и приемника с таким детектором – экономия стоимости, простота реализации и настройки, удобство тестирования устройств с изменением частоты; важнейшие недостатки – относительно высокая мощность последетекторных шумов вследствие сверхширокополосности, что сильно ухудшает чувствительность и динамический диапазон приемника (последний можно несколько расширить, увеличивая выходную мощность генератора зондирующих сигналов), отсутствие в последетекторном сигнале информации о фазе радиочастотной несущей в додетекторном сигнале; последний недостаток не позволяет применять этот тип детектирования в так называемых векторных анализаторах цепей (ВАЦ), то есть приборах, измеряющих, помимо амплитуд, и фазы сигналов; этот тип детектирования применяется в так называемых скалярных анализаторах цепей (САЦ);

– второй тип – супергетеродинный приемник с постоянной промежуточной частотой (ПЧ), с узкополосным детектированием в тракте ПЧ; такие приемники включают гетеродин, перестраиваемый синхронно с изменением частоты генератора зондирующих сигналов, смеситель и полосовой фильтр в тракте ПЧ (заметим, что в этом приемнике детектора в узком смысле слова вообще нет: после узкополосного фильтра включен АЦП, тракт разветвляется: в одной ветви производится цифровое амплитудное детектирование (корень квадратный из суммы квадратов отсчетов на интервале разрешения), в другой – цифровое фазовое детектирование (арктангенс из отношения средних квадратур)); вместо четырех гетеродинов можно использовать один, разветвляя его сигнал на четыре приемника (в структурной схеме АЦ тогда добавятся устройства разветвления, например, три расщепителя или делителя мощности); важнейшие преимущества такого приемника: (а) большой выигрыш в мощности последетекторных шумов, за счет чего резко улучшается чувствительность, увеличивается динамический диапазон; (б)

сохранение в последетекторном сигнале информации о фазе радиочастотной додетекторной несущей; (в) улучшение очистки от гармоник и побочных гармонических сигналов в принимаемом сигнале, важнейший недостаток: значительно большая стоимость приемников, в сравнении с детекторными; хотя этот тип приемников, в принципе, можно применять и в ВАЦ, и в САЦ, последнее не имеет смысла, т.к. к САЦ не предъявляются высокие требования по точности, чувствительности и динамическому диапазону, в частности, из-за вынужденного несовершенства калибровки в отсутствие фазовой информации и стремления к удешевлению прибора; часто на выходах приемников устанавливают аналого-цифровые преобразователи (АЦП) и дальнейшая обработка сигналов ведется в цифровом виде.

Блок координации и управления прибором (виртуальный) с помощью процессора, операционной системы и сигнальных процессоров выполняет различные наборы функций в зависимости от совершенства, многофункциональности, универсальности и сложности анализатора, в частности:

- управляет генераторами, гетеродинами, приемниками, переключателями, фильтрами, аттенюаторами и т.д.;
- синхронизирует алгоритмы и протоколы;
- формирует частотные, временные, мощностные сетки;
- инициирует и завершает выполнение компьютерных программ;
- контролирует соблюдение количественных условий и параметрических установок.

Блок измерительной калибровки и коррекции данных измерения с помощью процессора и операционной системы может работать в трех режимах: (1) механической калибровки (измерение калибровочных стандартов (эталонов); (2) электронной (автоматической) калибровки; (3) поверки прибора (верификации);

в режиме (1) он выполняет функции:

- принимает сообщение о выборе механической калибровки;
- начинает работу по сигналу оператора;
- предлагает на выбор виды и типы механической измерительной калибровки;
- предлагает и проверяет исполнение последовательности калибровочных измерений;
- фиксирует (сохраняет) результаты калибровочных измерений в стандартной протокольной форме;
- по результатам основных и калибровочных измерений рассчитывает и фиксирует скорректированные данные основных измерений;

в режиме (2) он выполняет функции:

- принимает сообщение о выборе электронной калибровки;
- начинает работу по сигналу оператора;
- проверяет «законность» и правильность подключения блока электронной калибровки;
- фиксирует (сохраняет) результаты калибровочных измерений в стандартной протокольной форме;
- по результатам основных и калибровочных измерений рассчитывает и фиксирует скорректированные данные основных измерений;

в режиме (3) он выполняет функции:

- принимает сообщение о выборе верификации;

- начинает работу по сигналу оператора;
- предлагает на выбор виды верификационных стандартов;
- предлагает и проверяет исполнение последовательности измерений с верификационными стандартами;
- фиксирует (сохраняет) результаты измерений с верификационными стандартами в стандартной протокольной форме.

Блок математической обработки сигналов

Виртуальный блок математической обработки сигналов и расчетов выполняет часть функций внутреннего процессора АЦ или внешнего компьютера, сопряженного с АЦ, в зависимости от выбранной конфигурации:

- приводит выходные данные приемников к необходимому формату;
- рассчитывает отношения сигнальных и опорных отсчетов;
- рассчитывает функционалы выборочных данных, соответствующие трассовым и временным усреднениям, а также другим операциям.

Блок отображения результатов

Управляет работой монитора (в варианте внутренних процессора и монитора или внешнего компьютера):

- разбивает поле отображения на зоны;
 - выбирает вид, форму и режим отображения в каждой зоне (диаграмма Вольперта-Смита, частотная панорама амплитуды или фазы S-параметра в линейном или логарифмическом формате, мощностная панорама и т.п.);
 - формирует экранные надписи;
- обслуживает систему маркеров и граничных линий.

Виртуальный блок внутренней автоматики реализует некоторую форму внутренней измерительной автоматики определенной сложности: от наиболее простой формы – *состояния воспроизведения* (устанавливает АЦ в предконфигурированное состояние со всеми необходимыми инструментальными параметрами) до наиболее сложной – *тестового программирования* (запись нажатий кнопок, усложненные программы, различные интерфейсы с персональными измерительными услугами и т.п.).

Типичные параметры и характеристики тестируемых устройств

Большое разнообразие тестируемых устройств и характеристик

Приведенные выше классификация цепей и перечень тестируемых устройств показывают, насколько велико разнообразие цепей, узлов, трактов, устройств, и позволяют догадываться, что число подлежащих измерению анализатором цепей параметров и характеристик велико. Это число еще увеличивается, если учесть разветвленные семейства родственных параметров и характеристик. Например, отражение от порта тестируемого устройства характеризуется семейством, включающим: (комплексный) коэффициент отражения (параметр рассеяния S_{11}), задаваемый модулем и фазой; модуль коэффициента отражения в децибельной мере (возвратные потери); коэффициент стоячей волны (КСВ); коэффициент бегущей волны (КБВ); модуль входного импеданса.

Цели тестирования

Кроме того, по мере развития радиоэлектронных приложений появляются все новые устройства и цепи и все новые функции как старых, так и новых устройств. Отталкиваясь от

назначения тестируемых устройств и необходимости измерения их параметров и характеристик, можно назвать следующие целевые области измерения:

А) Проверка выполнения технических требований (спецификаций) к устройству.

Б) Выяснение степени искажения выполняемых устройством функций, в частности, степени искажения передачи сигналов.

В) Оптимизация параметров и характеристик устройств в процессе их конструирования и подбора.

Г) Описание «неизвестных» устройств.

Д) Получение данных для математического моделирования устройства.

Измерения анализатором цепей (АЦ) можно разделить в зависимости от класса тестируемого устройства (см. 1.7.) (линейное – нелинейное, активное – пассивное, стационарное – нестационарное, детерминированное – стохастическое, управляемое – не управляемое, сосредоточенное – распределенное, инерционное – безынерционное, частотно-селективное – неселективное, и т.д.), области анализа (частотная, временная), сканируемого параметра (частота, мощность). На этом этапе изложения ограничимся перечнем типичных измеряемых параметров и характеристик, не вдаваясь в методологию измерения и интерпретации, анализ эффективности измерения и т.п.

Во многих измерительных задачах необходимо измерение только амплитудных параметров (модулей $|S_{ij}|$ параметров рассеяния). Например, для определения степени согласования может быть необходимо измерить модуль коэффициента отражения ($|S_{11}|$ или $|S_{22}|$) (КСВ) от порта устройства, или для определения усилительной способности активного устройства необходимо измерить модуль его коэффициента пропускания ($|S_{21}|$ или $|S_{12}|$), или для выяснения заградительной способности фильтра необходимо измерить модуль его коэффициента передачи (режекции) в полосе заграждения. Если при этом требуемые динамический диапазон и точность измерения не велики, достаточно использовать *скалярный анализатор цепей* (САЦ). В противном случае необходим *векторный анализатор цепей* (ВАЦ), который за счет применения супергетеродина в приемниках обеспечивает узкую полосу тракта ПЧ и, следовательно, высочайший динамический диапазон, а за счет измерения фазы – изошренную калибровку и, следовательно, высочайшую точность измерения.

Но часто измерительная задача поставлена так, что кроме амплитудных параметров, необходимо измерение и фазовых параметров и характеристик. Это, например, необходимо:

- для полной характеристики линейной цепи;
- при конструировании согласующих цепей, поскольку при этом необходимы значения полного импеданса;
- при математическом моделировании устройств;
- для характеристики устройств и трактов во временной области при частотном панорамировании;
- для описания векторных уклонений выполняемых устройствами функций;
- и т.д.

Типичные измеряемые параметры и характеристики

1) Характеристики отражения от порта устройства:

- Частотные зависимости модуля $\rho(f)$ и фазы $\varphi_{\text{отр}}(f)$ коэффициента отражения $\Gamma = \rho e^{i\varphi_{\text{отр}}}$ от порта.

- Частотные зависимости действительной $R(f)$ и мнимой $X(f)$ частей входного импеданса $Z = R + jX$, однозначно связанные с $\rho(f)$, $\varphi_{\text{отр}}(f)$.
- Частотные зависимости действительной $G(f)$ и мнимой $B(f)$ частей входного адмиттанса $Y = G + jB$, однозначно связанные с $\rho(f)$, $\varphi_{\text{отр}}(f)$.
- Частотная зависимость возвратных потерь $L_{\text{возвр}}(f)$, однозначно связанных с ρ :

$$L_{\text{возвр}} = -20 \lg \rho.$$
- Частотная зависимость КСВ, однозначно связанного с ρ : $\text{КСВ} = \frac{1+\rho}{1-\rho}$.
- Частотная зависимость коэффициента бегущей волны (КБВ), однозначно связанного с ρ :

$$\text{КБВ} = \frac{1-\rho}{1+\rho}.$$
- Временные (пространственные) зависимости тех же параметров $\rho, \varphi_{\text{отр}}, R, X, G, B, L_{\text{возвр}}, \text{КСВ}, \text{КБВ}$ (при фиксированном опорном сечении тракта).
- Интерпретация изломов и скачков временной зависимости $\varphi_{\text{отр}}$.
- Классификация скачков временной зависимости $X(t)$ мнимой части входного импеданса в терминах «емкостной переход – индуктивный переход».

2) Характеристики пропускания с одного порта на другой:

- Частотные зависимости модуля $\tau(f)$ (пропускания, амплитудно-частотной характеристики) и фазы $\phi(f)$ (фазовой задержки, внесенной фазы, фазо-частотной характеристики) коэффициента пропускания $T = \tau e^{j\phi}$ с одного порта на другой.
 - Частотная зависимость внесенных потерь $L_{\text{проп}}(f)$, дБ, однозначно связанных с τ :

$$L_{\text{проп}} = -20 \lg \tau.$$
 - Частотная зависимость направленности и коэффициента ответвления направленного ответвителя или направленного моста.
 - Частотная зависимость групповой задержки $GD(f)$ от одного порта к другому, однозначно связанной с $\phi(f)$:

$$GD = -\frac{1}{2\pi} \cdot \frac{d\phi}{df}.$$

3) Характеристики нелинейности устройства: зависимости от входной мощности:

- Зависимости модуля $\rho(P_{\text{вх}}; f)$ и фазы $\varphi_{\text{отр}}(P_{\text{вх}}; f)$ коэффициента отражения от входной мощности $P_{\text{вх}}$ на заданной частоте f .
- Зависимости модуля $\tau(P_{\text{вх}}; f)$ и фазы $\phi(P_{\text{вх}}; f)$ коэффициента пропускания от входной мощности $P_{\text{вх}}$ на заданной частоте f .
- Отклонение $\tau(P_{\text{вх}}; f)$ от постоянной величины (насыщение) по заданному критерию как функция от $P_{\text{вх}}$ на различных частотах f .
- Отклонение $\phi(P_{\text{вх}}; f)$ от линейной зависимости по заданному критерию как функция от $P_{\text{вх}}$ на различных частотах f .

- Нелинейные искажения (протечка, интермодуляция, гармонические искажения, амплитудно-фазовая конверсия и др.) по заданным критериям в функции входной мощности на различных частотах f .

Классификация тестируемых устройств

Прежде чем установить набор подлежащих тестированию устройств, перечислить измеряемые характеристики этих устройств, необходимые для этого аппаратные и алгоритмические методы измерений и их техническое воплощение, необходимо дать классификацию тестируемых устройств (цепей). Всякая классификация может быть выполнена на различную глубину, в зависимости от цели, ради которой она проводится. В нашем случае классификация цепей должна охватывать лишь те признаки, которые существенны для структуры анализатора цепей, технических решений его компонент, алгоритмов обработки сигналов и управления прибором, эффективности его работы и т.д. Ниже предлагается вариант такой классификации.

1. Цепь может быть *линейной*, или *нелинейной*. Этот важнейший классификационный признак основан на отношении цепи к фундаментальному принципу суперпозиции:

$$\mathbf{T} \cdot (\alpha \mathbf{q}_{\text{вх1}} + \beta \mathbf{q}_{\text{вх2}}) = \alpha \mathbf{T} \mathbf{q}_{\text{вх1}} + \beta \mathbf{T} \mathbf{q}_{\text{вх2}}$$

выражающему линейность системного оператора \mathbf{T} (α, β – произвольные скаляры). Строго говоря, все физические системы, встречающиеся в радиотехнике, в той или иной степени нелинейны. Термин «линейная» означает всего лишь, что степень нелинейности цепи такова, что ею можно пренебречь с точки зрения принятых нами критериев, которые сами могут нами меняться в разных приложениях и ситуациях. В свою очередь, констатируя нелинейность в рамках этих критериев, мы можем углубить классификацию, например, разделив нелинейности на два вида: *существенные* и *малые*. Заметим также, что нелинейность не является имманентным свойством цепи, ее степень и само наличие могут меняться в зависимости от условий работы цепи, уровня сигналов и т.д. Например, усилитель можно считать линейным в области «малых» входных сигналов, слабо нелинейным – в области небольших «сжатий» усиления, существенно нелинейным – в области предельно больших сигналов. Другой пример – устройство может демонстрировать нелинейность пропускания на одной паре портов, и линейность – на другой.

2. Цепь называется *стационарной (инвариантной во времени)*, если ее выходные отклики не меняются при сдвиге в целом входных воздействий во времени. То есть цепь стационарна, если из равенства

$$\mathbf{q}_{\text{вых}}(t) = \mathbf{T} \mathbf{q}_{\text{вх}}(t)$$

следует, что

$$\mathbf{q}_{\text{вых}}(t \pm t_0) = \mathbf{T} \mathbf{q}_{\text{вх}}(t \pm t_0)$$

при любом t_0 . С известной осторожностью можно считать, что стационарная система имеет постоянные во времени параметры. Обратное, вообще говоря, неверно, из-за трудности интерпретации выражения «все параметры». Например, рассмотрим линейную цепь, у которой параметры, в том числе – ширина импульсной характеристики, постоянны во времени, но форма импульсной характеристики меняется во времени; очевидно, это не стационарная цепь.

3. Нестационарные цепи, в свою очередь, можно разбить на несколько категорий. В частности, нестационарная цепь может иметь детерминированное временное поведение параметров (то есть в виде заданной функции времени; условимся такую цепь называть *детерминированной*), или стохастическое поведение параметров (то есть в виде случайных процессов некоторой структуры; условимся такую цепь называть *стохастической*). Более сложный случай возникает, когда одни параметры нестационарной цепи ведут себя детерминированно, другие – стохастически.

4. Параметры нестационарной цепи (или часть из них) могут управляться извне (детерминированно или стохастически) – такие цепи называются *параметрическими*, или их изменение может быть предопределено. Примером параметрической цепи является смеситель в блоке преобразователя частоты (конвертора) супергетеродинного приемника: мгновенные характеристики смесителя в тракте ВЧ-ПЧ управляются сигналом гетеродина.

5. Элементы цепи могут быть *сосредоточенными* или *распределенными*. Вся цепь также может классифицироваться по этому признаку. Понятно, что отнесение цепи или элемента к тому или иному классу зависит от соотношения характерного размера с длиной волны. Четкой границы здесь нет: объект считается сосредоточенным, если его размер много меньше длины волны, и распределенным в ином случае.

6. Цепь может считаться *безынерционной* или *инерционной*. Строго говоря, все физические системы, с точки зрения отношения их реакций ко входным сигналам, инерционны. Термин «безынерционная» означает всего лишь, что инерционность системы, выраженная через такие параметры, как «мертвое время», ширина импульсной характеристики и т.п., пренебрежима с точки зрения принятых нами критериев, которые сами могут нами меняться в разных приложениях и ситуациях.

7. В пределах заданного частотного диапазона цепь может считаться *частотно селективной* или *частотно не селективной*. Например, такие цепи, как нагрузка холостого хода (ХХ), нагрузка короткого замыкания (КЗ), резистор часто для упрощения анализа считаются неселективными в пределах некоторого частотного диапазона. Строго говоря, всякая цепь частотно селективна в любом частотном диапазоне. Выражение «частотно не селективна» означает всего лишь, что вариации параметров цепи с частотой в данном диапазоне пренебрежимы с точки зрения принятых нами критериев, которые сами могут нами меняться в разных приложениях и ситуациях.

8. В свою очередь, частотно селективные цепи можно классифицировать по типу частотной зависимости важнейших параметров. Например, коэффициент пропускания двухпортовой линейной стационарной цепи может меняться по частоте как у фильтра нижних частот (ФНЧ), фильтра верхних частот (ФВЧ), полосового фильтра (ПФ), различных видов запирающих фильтров, вырезающего фильтра и т.п.

9. Общеизвестно деление цепей на *активные* и *пассивные*. Устройство СВЧ называется *пассивным*, если в его состав не входят активные преобразующие или усиливающие элементы, например, транзисторы. В противном случае устройство СВЧ называется активным

10. С электродинамической точки зрения устройства СВЧ делятся на *взаимные* и *невзаимные*. К сожалению, четкие и достаточно строгие общие интерпретации свойства взаимности возможны только на базе не введенных пока понятий, таких как матрица рассеяния, тензоры диэлектрической и магнитной проницаемости, а также на базе таких свойств, как лемма Лоренца в дифференциальной или интегральной форме и т.п. Часто поясняют свойство взаимности на примерах цепей. Например, если в радиолинии, включающей передающую и приемную антенны и атмосферный канал, приемную антенну сделать передающей, а передающую – приемной, параметры радиолинии не изменятся.

11. Цепи можно разделить на *частото-неизменяющие* и *частото-преобразующие*. К первым относятся цепи, на выходе которых несущие (опорные) частоты те же, что на входе. Типичный пример частото-преобразующей цепи – преобразователь частоты (конвертор), состоящий из смесителя, гетеродина и выходного фильтра. При подаче на вход конвертора узкополосного сигнала с несущей f_0 на выходе отфильтровывается узкополосный сигнал с несущей, равной модулю разности частот f_0 и гетеродина f_r . Другие примеры: делитель частоты, умножитель частоты и т.д.

Примеры тестируемых цепей

Анализаторы цепей (АЦ) как измерители параметров рассеяния (S -параметров) (скалярные АЦ – измерители модулей S -параметров, векторные АЦ – измерители модулей и фаз S -параметров) способны производить:

- Измерение амплитудных характеристик – усиления, ослабления, развязки, потерь на отражение, неравномерности АЧХ.
- Измерение фазовых характеристик – фазового сдвига коэффициента передачи, фазового сдвига коэффициента отражения, группового времени запаздывания, нелинейности фазочастотной характеристики, непостоянства группового времени запаздывания.
- Измерение комплексных характеристик – коэффициента передачи, коэффициента отражения, полного сопротивления, полной проводимости.
- Измерение параметров сигналов – отношения уровней, разности фаз.

Необходимость измерения амплитуд параметров рассеяния очевидна. Знание фазовых соотношений в радиотехнических цепях позволяет:

- создавать узкополосные фильтры;
- выделять полосы модуляции;
- улучшать свойства генераторов СВЧ, антенн, усилителей и других устройств.

Измерение фазы имеет большое значение в устройствах и системах определения пространственного положения объектов и повышения достоверности передачи сообщений в условиях естественных и организованных помех.

Количество различных видов устройств (цепей) СВЧ, которые можно анализировать с помощью АЦ, огромно и увеличивается по мере развития приложений радиоэлектроники и самих АЦ. Чтобы как-то упорядочить примерный перечень возможных тестируемых цепей, в [1.2] предложено располагать его по степени активности и по степени интегрированности устройства. В этом случае примерный перечень имеет вид [1.2]:

А) Пассивные устройства (в порядке убывания интегрированности):

- Антенные переключатели (дуплексеры, триплексеры и т.д.).
- Блоки частотной развязки (диплексеры).
- Фильтры.
- Направленные ответвители.
- Мосты (в том числе ответвительные).
- Расщепители, делители мощности.
- Объединители.
- Развязывающие устройства.
- Циркуляторы.
- Аттenuаторы.
- Адаптеры (переходные устройства).
- Нагрузки на отрезках линий передачи: короткозамкнутые, холостходные, с произвольным КСВ.
- Сосредоточенные элементы. Конкретные конструкции в диапазоне СВЧ в ряде случаев используют сосредоточенные элементы, то есть элементы с физическими размерами, много меньшими длины волны во всем диапазоне рабочих частот (резисторы, конденсаторы,

индуктивности, диоды и т.д.). Эти элементы на СВЧ имеют паразитные параметры, которые нельзя игнорировать:

– Резистор кроме основного параметра R характеризуется постоянной времени τ , учитывающей влияние индуктивности соединительных проводников L_s и емкости корпуса C_p , равной большей из величин L_s/R или $C_p R$.

– Вследствие влияния потерь в проводниках и диэлектрике конденсатор характеризуется, кроме основного параметра C , тангенсом угла потерь $\operatorname{tg}\delta$, равным большей из величин $r_s\omega C$ или $1/R_p\omega C$. Еще один паразитный параметр – индуктивность соединительных проводников.

– Катушка индуктивности из-за влияния сопротивления проводников и других потерь, емкости корпуса дополнительно характеризуется добротностью Q , равной меньшей из величин $\omega L/r_s$ или $R_p/\omega C_p$.

– Кроме основных параметров диод характеризуется паразитными параметрами: емкостью корпуса C_p и перехода C_i (последняя в ряде случаев может быть полезным параметром), индуктивностью выводов L_p , сопротивлением потерь R_s .

Рассмотренные выше параметры достаточно полно характеризуют элементы и позволяют предсказать их поведение в схемах, работающих на низких и высоких частотах (до единиц и десятков мегагерц). Но представление элементов в СВЧ диапазоне в виде эквивалентной схемы, состоящей из элементов с сосредоточенными параметрами, затруднительно, поэтому их представляют как многополюсники и анализируют с помощью векторных анализаторов цепей.

Б) Устройства, имеющие признаки пассивного и активного (в порядке убывания интегрированности):

- Антенны.
- Переключатели.
- Коммутаторы (мультиплексеры).
- Смесители.
- Отборщики.
- Умножители.
- Диоды.

В) Активные устройства (в порядке убывания интегрированности):

- Интегральные схемы.
- Монолитные интегральные схемы.
- Трансмиссионно-рефлексионные модули.
- Приемо-передатчики.
- Приемники.
- Настройщики.
- Конвертеры.
- Адаптеры оцифровки видеоизображений.
- Усилители.
- Генераторы, управляемые напряжением (ГУН) или током (ГУТ).

- Генераторы.
- Модуляторы.
- Атенюаторы, регулируемые напряжением.
- Транзисторы.

Аппаратный анализ столь различных устройств требует различных зондирующих совокупностей (например, один монохроматический сигнал, с частотной панорамой или без нее, с мощностной панорамой, или без нее; два одновременных монохроматических сигнала, амплитудно или частотно модулированный сигнал; импульсный сигнал и т.д.), различных типов измеряемых характеристик, различной архитектуры измерителя, различных протокола и алгоритмов обработки сигналов, различных дополнительно подключаемых измерительных приборов и т.д.

Возможно тестировать пассивные и активные устройства, а также те, что имеют признаки того и другого; линейные и нелинейные устройства; одни и те же устройства в их линейной области и нелинейной области; определять характеристики устройств в частотной и временной областях.

Последовательность действий при тестировании устройств (см. Осн анал цепей, с. 66) (подготовка, калибровка, измерение, с. 70-74, уход за кабелями и разъемами, с. 67-69)

Подготовка: включение, прогрев, подключение и проверка разъемов и кабелей, подключение адаптеров, подключение тестируемого устройства для проверки функционирования прибора.

Калибровка: Отключение тестируемого устройства, измерительная калибровка, верификация, сохранение состояния прибора.

Измерение: Подключение ТУ, измерение.

Факторы эффективности

Некоторые метрологические понятия

Измерение физической величины это нахождение ее значения экспериментальным путем с помощью специальных технических средств – *средств измерений*. По способу получения числового значения измеряемой величины все измерения делят на *прямые, косвенные, совокупные и совместные*. *Прямые измерения* основаны на методе сравнения измеряемой величины с мерой этой величины или на методе непосредственной оценки значения измеряемой величины по отсчетному устройству средства измерений, шкала которого проградуирована в единицах измеряемой величины. *Косвенные измерения* являются более сложным видом измерений, результат которых получают после прямых измерений величин, связанных с измеряемой величиной известной зависимостью. Косвенные измерения в ряде случаев позволяют получить более точные результаты, чем прямые, а иногда являются единственно возможными для данной физической величины. *Совокупные измерения* представляют собой неоднократные, обычно прямые измерения одной или нескольких одноименных величин (при различных сочетаниях этих величин) с получением общего результата измерений путем решения системы уравнений, составляемых по частным результатам измерений. *Совместные измерения* состоят в одновременном измерении двух или нескольких не одноименных величин с последующим вычислением результата путем решения системы полученных при измерениях уравнений. *Средство измерений* – техническое устройство, используемое при измерениях и имеющее нормированные метрологические характеристики. К средствам измерений относятся меры, измерительные приборы и измерительные установки (системы). Кроме того, к средствам измерений относят измерительные преобразователи, измерительные принадлежности, служащие для расширения диапазона измерений, повышения точности измерений, дистанционной передачи результатов измерений, питания средств измерений высокостабильной электрической энергией и т. д. Иногда необоснованно расширяют применение понятия средства измерений на

самостоятельно функционирующие, имеющее конкретное техническое назначение сложные устройства, хотя и решающие измерительные задачи. Например, радиолокационная станция является также измерительным устройством, но не средством измерения. Средство измерений не предназначено для выполнения самостоятельных функций, а выполняет их совместно с источниками измеряемых физических величин. Этими источниками являются другие технические устройства, явления, тела и т. д.

Не следует относить к средствам измерений и многочисленные виды испытательных установок, несмотря на то что многие из них также имеют нормированные метрологические характеристики. Например, барокамеры, термокамеры, вибростенды, центрифуги должны обеспечивать определенный диапазон и поддержание условий испытаний с заданной точностью. Но указанные средства не измеряют, а воспроизводят с теми или иными точностями и диапазонами условия испытаний (в том числе и условия измерений).

Мерой называется средство измерений в виде тела или устройства, предназначенное для хранения и (или) воспроизведения физической величины заданного размера. К мерам относятся *нормальные элементы, измерительные магазины емкостей, генераторы стандартных сигналов* и др. В нашем случае к мерам относятся калибровочные элементы (прецизионные нагрузки холостого хода, короткого замыкания, нагрузки с заданным КСВ и т.д.).

Измерительная установка представляет собой комплекс средств измерений и вспомогательных устройств, обеспечивающий получение измерительной информации на исследуемом объекте в заданном объеме и заданных условиях. Обычно в измерительную установку входят приборы, меры, а также устройства их сопряжения, питания, обеспечения режимов измерений. Такая установка может быть с ручным управлением или автоматизированная. В нашем случае измерительная установка включает непосредственно АЦ, набор калибровочных мер, соединительные кабели, оборудованные разъемами, переходные соединители (адаптеры), дополнительные подключаемые по мере необходимости извне устройства, например, усилители, аттенюаторы и т.п.

Измерительная система – понятие, обязанное усложнению измерительных задач, решаемых в интересах проверки работоспособности многопараметрических электронных систем в условиях их функционирования. В настоящее время, как правило, измерительные системы создаются *автоматизированными* и называются *автоматизированными измерительными системами* (АИС), *автоматизированными системами контроля* (АСК) и т. д. Объединение измерительных приборов в измерительную систему облегчается принятым в большинстве случаев управлением от (внешнего или внутреннего) компьютера, согласованным интерфейсом и открытостью матобеспечения к перепрограммированию.

Все большую роль в средствах измерений, особенно автоматизированных, играют *измерительные преобразователи (датчики)*, предназначенные для преобразования измерительного сигнала в форму, удобную для передачи, дальнейшего преобразования, хранения и обработки. Измерительные преобразователи имеют нормированные метрологические характеристики, конструктивно они, как правило, оформлены в самостоятельное средство измерений, но иногда являются встроенной составной частью измерительного прибора. Основное требование, предъявляемое к измерительным преобразователям: их унификация и стандартизация с целью сопряжения со средствами измерений, использования в измерительных системах, встраивания в объект измерения. Многие датчики выполняются не только для передачи преобразованной информации в отсчетное устройство, как, например, в случае дистанционного измерения давления, но и для передачи измерительного сигнала в соответствующие каналы управления.

По метрологическому иерархическому принципу признаку все средства измерений делятся на рабочие, образцовые и эталоны.

Рабочие средства измерений – не предназначены для воспроизведения и хранения единиц физических величин, а также для передачи их размеров другим средствам измерений.

Образцовые средства измерений – меры, измерительные приборы (установки) или измерительные преобразователи, применяемые для поверки по ним других средств измерений (рабочих или образцовых меньшей точности). Образцовые средства измерений обычно специально разрабатываются и аттестуются. В радиоизмерительной технике, особенно при высокоточных измерениях, иногда не удается создать образцовые приборы, и в качестве их используются специально отобранные и аттестованные приборы из числа рабочих.

Эталоны – средства измерений (обычно комплекс средств измерений), предназначенные для воспроизведения и (или) хранения единицы физической величины с целью передачи ее размера низшим по точности средствам измерений (как правило, образцовым). Эталон должен быть официально утвержден Госстандартом в качестве такового. Эталон может быть *государственным* – исходным для страны. При этом государственный эталон является *первичным*, воспроизводящим или хранящим единицу физической величины, *специальным*, заменяющим первичный эталон для воспроизведения единицы в особых условиях. Для обеспечения нужд отраслевых метрологических служб применяются *вторичные эталоны*, размер единицы которым передается от первичного (специального) эталона.

В некоторых случаях, когда не представляется возможным создать эталонное средство измерений, допускается применение так называемого *группового эталона*, состоящего из группы однотипных средств измерений (например, образцовых или даже рабочих). Размер единицы физической величины, хранимой групповым эталоном, определяется в соответствии с принятым алгоритмом обработки значений физической величины, воспроизводимых отдельными средствами измерений, входящими в состав группового эталона. В настоящее время применяются достаточно сложные алгоритмы обработки, вплоть до применения методов «самообучения».

Анализатор цепей (АЦ) СВЧ предназначен для *совокупных косвенных измерений S*-параметров (параметров рассеяния) различных устройств и цепей СВЧ в лабораторных условиях. Он является средством измерения, более точно – измерительным прибором.

Всем средствам измерений присущи основные свойства: метрологические, эксплуатационные, информационные и др. Наиболее важными являются метрологические свойства (характеристики) средств измерений. К ним относятся: точность и диапазон измерений, чувствительность, разрешающая способность, стабильность, метрологическая надежность, неинформативные параметры средства измерений (например, диапазон частот, в котором возможно измерять значения нужного параметра) и т.п.

Все свойства, и в частности, метрологические так или иначе отражаются на эффективности средства измерения. *Эффективность* – общий условный итог качества достижения поставленной перед средством измерения цели. Количественная оценка эффективности возможна, если сформулирован комплексный критерий эффективности, такой, что большему значению критерия соответствует большая эффективность. Если комплексный критерий эффективности сформулирован, появляется возможность не только сравнивать различные измерительные приборы, но и оптимизировать их. Комплексный критерий учитывает веса различных свойств средства измерения, в частности, метрологических. Несмотря на сугубую условность и гибкость количественного исчисления весов, очень редко удается подобрать приемлемый комплексный критерий эффективности радиоэлектронной системы, в частности, измерительного прибора. Уменьшение же комплексности критериев увеличивает число *частных критериев*, а в пределе разрушает саму идею исчисления эффективности и оптимизации на ее основе.

Тем не менее, по опыту оптимизации различных радиоэлектронных систем, можно утверждать, что возможно сформулировать ряд частных критериев невысокой комплексности, позволяющих сравнивать измерительные приборы по отдельным направлениям и оптимизировать приборы по этим направлениям. Примерами таких направлений являются комплексы точность-разрешение-оперативность, цена-многофункциональность, точность-цена-надежность и т.д.

ВАЦ или САЦ

При выборе архитектуры и схемотехнического решения АЦ перед разработчиком стоит ряд дилемм, в частности:

- измерение амплитуд и фаз, или только амплитуд;
- использовать широкополосные детекторные приемники, или узкополосные супергетеродинные приемники;
- использовать простую однопортовую калибровку, или сложную многопортовую калибровку;
- использовать дорогие и сложные направленные ответвительные приборы с максимальной направленностью, или недорогие с умеренной направленностью;
- использовать анализ в частотной области, во временной области, в обеих областях;
- использовать для получения анализа во временной области импульсные зондирующие сигналы, или только частотно-сканирующие сигналы.

На первый взгляд, решение этих дилемм порождает многочисленные типы и подтипы АЦ, например, скалярный АЦ (САЦ) с широкополосным или узкополосным детектором, с простой или сложной калибровкой, с прецизионными сверхширокополосными направленными ответвителями (НО) или с направленными мостами с «вшитыми» диодными детекторами и т.д. На самом деле, в результате логических, технических, маркетинговых соображений и использования здравого смысла, для АЦ диапазона СВЧ в подавляющем большинстве случаев остаются только два основных типа и несколько мало различающихся подтипов.

Первый основной тип – **скалярный анализатор** (измеряет только модули параметров рассеяния) с широкополосными детекторными приемниками, с простой однопортовой калибровкой (т.н. нормированием), с ответвительными устройствами умеренной направленности (например, с направленными мостами с «вшитыми» диодными детекторами), с анализом в частотной области. Сложные многопортовые методы калибровки в САЦ недоступны, т.к. требуют измерения не только модулей, но и фаз параметров рассеяния. Использование в САЦ узкополосных супергетеродинных приемников, хотя и снизило бы уровень шума и расширило динамический диапазон измерения модулей параметров рассеяния, но увеличило бы цену САЦ до сравнимых с векторным АЦ (ВАЦ) величин, что резко сузило бы рынок таких приборов. Использование дорогих прецизионных НО нецелесообразно по той же причине, а также потому, что уменьшение измерительной ошибки за счет дефицита направленности на фоне сравнительно больших (из-за ущербности калибровки) ошибок за счет многих других факторов – не имеет смысла.

Второй основной тип – **векторный анализатор** (измеряет модули и фазы параметров рассеяния) с узкополосными супергетеродинными приемниками, с набором методов измерительной калибровки по выбору, в том числе т.н. «полные» калибровки, рассчитанные на различные наборы калибровочных эталонов, а также (возможно, опционально) с «электронной» калибровкой, с прецизионными НО, с использованием частотного сканирования зондирующих сигналов, с анализом в частотной и (по крайней мере, опционально) во временной областях. Такой прибор имеет значительную стоимость, но обеспечивает высокие (а по некоторым факторам – наивысшие) измерительные точности, динамический диапазон, универсальность, чувствительность и другие факторы эффективности.

Фактически гипотетический спектр типов АЦ выродился в два типа: недорогой малоэффективный, с узкой областью применения, и дорогой высокоэффективный универсальный.

Несомненно, что в ближайшее время, с развитием субнаносекундной импульсной техники и увеличением быстродействия АЦП появится третий основной тип АЦ – импульсный анализатор цепей (ИАЦ), с импульсными зондирующими сигналами (в виде короткого импульса

определенной формы или короткого скачка), с прямым анализом во временной области и косвенным – в частотной области. В настоящее время такие приборы выпускаются или разрабатываются с ограничением верхней частоты спектра анализа до сотен МГц.

Точностные характеристики

Вокруг понятия *точность измерения* группируется комплекс параметров и характеристик, так или иначе характеризующих степень приближения результатов измерений к истинным значениям измеряемых величин. Важнейшая из таких характеристик – *погрешность измерения* (*погрешность прибора*) – отражает отклонение результата измерения от *истинного значения* измеряемой величины. Рассмотрим сначала случай измерения скалярной величины A ; (абсолютная) погрешность измерения

$$\Delta A = A_{\text{изм}} - A_{\text{ист}},$$

где $A_{\text{изм}}$, $A_{\text{ист}}$ – измеренное и истинное значения.

Заметим прежде всего, что независимо от типа, характера и т.п. погрешности измерения, она не может стать известной точно как величина. Действительно, мы исходим из того, что никакая физическая величина никаким измерительным прибором не может быть измерена абсолютно точно (следовательно, всегда $\Delta A \neq 0$). Но если в результате какого-то измерения получено значение $A_{\text{изм}}$ и каким-то способом найдено точное значение погрешности ΔA , то это означает, что стало известно точное значение измеряемой величины:

$$A_{\text{ист}} = A_{\text{изм}} + \Delta A,$$

что невозможно.

Далее, необходимо различать *погрешность измерения* и *погрешность прибора* (в нашем случае погрешность АЦ), первое понятие шире.

В общем плане погрешность измерений обуславливаются многими факторами, в частности:

- погрешностью метода измерения (*методическая погрешность*); в нашем случае методическая погрешность обязана неточности перехода от электродинамического представления линий передачи и их сочленений к модели эквивалентных линий и устройств; неточности портового представления цепей СВЧ; неточности предположения о линейности элементов трактов; некоторой неадекватности модели матрицы рассеяния; и т.п.;
- несовершенством средства измерения; этот многосторонний фактор будет постепенно раскрываться на протяжении книги;
- влиянием условий проведения измерений; в нашем случае это, например, значение и изменчивость температуры, влажности и давления окружающего воздуха; вибрации различного рода и т.д.;
- влиянием каналов связи объекта измерений со средством измерений; в нашем случае это линии передачи, разъемы и т.п. соединительных трактов между АЦ и тестируемым устройством (ТУ);
- субъективными погрешностями, промахами, недочетами оператора, управляющего, настраивающего, калибрующего средство измерения, фиксирующего и интерпретирующего его показания.

Последние два фактора не относятся к погрешности средства измерений, но влияют на погрешность результата измерений.

По способу числового выражения погрешности разделяются на:

- *абсолютные*, выражаемые в единицах измеряемой величины;

- *относительные*, выражаемые отношением абсолютной погрешности к истинному значению измеряемой величины или к значению, близкому к истинному;

- *приведенные*, выражаемые отношением абсолютной погрешности к некоторому условно принятому значению измеряемой физической величины, например, к значению верхнего предела измерений.

По характеру и причинам появления погрешности разделяются на *систематические, случайные и промахи*.

Систематические погрешности представляют собой такие составляющие погрешности средства измерений, которые при неоднократном использовании средства измерений остаются постоянными или изменяются закономерно. Эти погрешности поддаются изучению и учету, в силу чего результат измерений может быть уточнен путем внесения *поправок*. Более того, стремятся создать средство измерений таким образом, чтобы исключить влияние систематических погрешностей.

Вообще в измерительных приборах *систематические погрешности* обусловлены:

- неточной градуировкой средств измерений;
- неправильной установкой прибора;
- несовершенством метода измерений;
- применением приближенных формул (в частности, в тех случаях, когда измерительный прибор обрабатывает измерительную информацию);
- и т. д.

Конкретно в частотно-панорамных и мощностно-панорамных измерительных системах СВЧ диапазона, в частности, в анализаторах цепей, систематические погрешности обусловлены *рассогласованиями* в точках сочленений трактов, *протечками* из одних точек в другие, изменениями коэффициентов передачи в трактах прибора с частотой и т.д.

Вообще в измерительных приборах *исключение (или уменьшение) систематических погрешностей* бывает возможно:

- установкой нуля отсчета;
- калибровкой измерительных приборов перед измерением.

Конкретно в анализаторах цепей СВЧ диапазона *калибровка* разделяется на *измерительную и производственную*. Измерительная калибровка в этих приборах и коррекция данных на ее основе носит изошренный характер, особенно в векторных анализаторах цепей. Эти вопросы будут в дальнейшем подробно рассмотрены в томе 3 серии.

Случайные погрешности изменяются непредвиденным (случайным) образом при проведении повторных измерений одной и той же физической величины. В отличие от систематических случайные погрешности невозможно исключить из результата измерений. Но многократные измерения одной и той же величины позволяют уменьшить случайную составляющую погрешности путем нахождения среднего ее значения.

Промахи представляют собой погрешности, существенно превышающие присущие данному средству измерений систематические и случайные погрешности. Они возникают или из-за неисправностей средства измерений, или из-за грубых ошибок оператора. Не существует абсолютно надежного алгоритма идентификации измерения как содержащего промах. Адекватная постановка этой задачи – двухальтернативное испытание статистических гипотез на основе математической статистики (*оптимальное обнаружение события*). При такой постановке задачи в рамках критерия Неймана-Пирсона можно выработать решающее правило, при котором всегда существует вероятность «ложной тревоги» (вероятность принятия решения «промах» при его отсутствии на самом деле) и вероятность пропуска промаха (вероятность принятия решения

«нет промаха», когда на самом деле он есть). При обнаружении промаха (то есть принятии решения «промах») соответствующий результат измерений должен быть исключен как неверный.

Погрешности, возникающие в нормальных условиях работы средств измерений [температура окружающей среды $20 \pm 5^{\circ}\text{C}$, атмосферное давление 100 ± 4 кПа (750 ± 30 мм рт. ст.), относительная влажность воздуха $65 \pm 15\%$], называются *основными*. В технических условиях на радиоизмерительные приборы обычно указываются также *дополнительные погрешности*, представляющие собой дополнительное изменение основной погрешности за счет изменения внешних условий относительно нормальных. Так, довольно часто указывается дополнительная погрешность за счет изменения температуры относительно нормальной.

Более глубокую, чем в теории ошибок интерпретацию видов ошибок, открывающую путь к оптимизации измерений, дает *математическая статистика*, точнее – ее часть *теория оценок*, кратко изложенная в главе 8. С ее точки зрения окончательный результат измерения $A_{\text{рез}}$ (не обязательно равный величине $A_{\text{ист}}$) есть *оценка* величины $A_{\text{ист}}$; оценка, за редкими тривиальными исключениями есть *случайная* величина; систематическая ошибка есть *смещение оценки* (то есть искажение ее математического ожидания), а случайная ошибка есть мера неопределенности оценки как случайной величины; например, среднеквадратичная абсолютная ошибка есть среднеквадратичное отклонение этой случайной величины. Теория оценок дает методы формирования и оптимизации структуры оценки по выборочным данным, то есть по результатам отдельных измерений, из которых формируется $A_{\text{рез}}$; это открывает путь к оптимизации обработки сигналов в измерительном приборе и, в какой-то степени – к оптимизации алгоритмов измерительной калибровки и других процедур.

Поскольку истинное значение измеряемой величины недостижимо (а при априорной известности этого значения нет необходимости в измерении), в теории ошибок, а более широко – в математической статистике иногда приходится в некоторые выражения подставлять вместо истинного значения $A_{\text{ист}}$ величину $A_{\text{рез}}$ или даже $A_{\text{изм}}$. Если ошибка подставляемой величины по отношению к $A_{\text{ист}}$ достаточно мала, такая подстановка в большинстве случаев вносит в некотором смысле пренебрежимую ошибку в выражение. Например, покажем, что если в выражение для относительной ошибки результата измерения

$$\delta_A = \frac{\Delta A}{A_{\text{ист}}} = \frac{A_{\text{рез}} - A_{\text{ист}}}{A_{\text{ист}}}$$

подставить $A_{\text{рез}}$ вместо $A_{\text{ист}}$, то есть заменить δ_A на

$$\delta'_A = \frac{\Delta A}{A_{\text{рез}}},$$

то относительная ошибка $\delta_{\delta'_A}$ выражения за счет этой подстановки мала. Действительно, записав

$$\delta_{\delta'_A} = \frac{\delta'_A - \delta_A}{\delta_A},$$

и подставив предыдущие равенства, получаем:

$$\delta_{\delta'_A} = \frac{\delta'_A - \delta_A}{\delta_A} = \delta'_A (1 - \delta'_A + \delta'^2_A - \dots) \approx \delta'_A.$$

Таким образом, при замене в выражении для δ_A величины $A_{\text{ист}}$ на величину $A_{\text{рез}}$, в оценку ошибки δ_A (а не в оценку измеряемой величины A !) вносится относительная ошибка порядка δ_A , что вполне приемлемо.

Вспомним теперь, что объектами измерения анализатора цепей являются комплексные величины – параметры рассеяния S_{ij} тестируемых устройств (скалярный анализатор измеряет только модули $|S_{ij}|$ этих параметров, векторный анализатор измеряет эти параметры в полном объеме). Понятие ошибки комплексной величины можно ввести по-разному, в зависимости от принятого представления комплексной величины. Например, при «естественном представлении» в форме $z = x + iy$ истинное значение измеряемой величины будет $z_{\text{ист}} = x_{\text{ист}} + iy_{\text{ист}}$, измеренное – $z_{\text{изм}} = x_{\text{изм}} + iy_{\text{изм}}$, абсолютная ошибка измерения будет комплексной величиной $\Delta z = \Delta x + i\Delta y = (x_{\text{изм}} - x_{\text{ист}}) + i(y_{\text{изм}} - y_{\text{ист}})$, а относительная ошибка – комплексной величиной сложной структуры:

$$\delta z = \frac{(x_{\text{изм}} - x_{\text{ист}}) + i(y_{\text{изм}} - y_{\text{ист}})}{x_{\text{ист}} + iy_{\text{ист}}}.$$

К сожалению, такое введение абсолютной и относительной ошибок не годится для практического использования по ряду причин. Во-первых, для комплексных величин не существует отношения «больше-меньше» и так введенные ошибки не позволяют сравнивать измерения по точности. Во-вторых, конкретное измерение параметра рассеяния осуществляется не путем канального разделения измерений действительной и мнимой частей, а путем канального разделения измерения амплитуды и фазы (аргумента) величины, то есть используется «экспоненциальное представление» комплексной величины $z = |z|e^{i\arg z}$. В-третьих, для характеристики точности измерения амплитуды удобно использовать относительную ошибку в линейном, или, чаще – в децибельном масштабе, а для характеристики точности измерения фазы удобно использовать абсолютную ошибку. В-четвертых, даже если для некоторых теоретических выкладок понадобится выражения для относительной ошибки комплексной величины в виде отношения двух комплексов, то и в этом случае гораздо удобнее использовать их выражения в «экспоненциальном» виде.

Таким образом, мы приходим к двум системам измеряемых величин и их ошибок; первая система: модули параметров рассеяния $|S_{ij}|$, $i, j = 1, \dots, N$, их абсолютные ошибки $\Delta|S_{ij}|$ и относительные ошибки $\delta|S_{ij}|$; вторая система: фазы (аргументы) параметров рассеяния $\varphi_{ij} = \arg S_{ij}$ и их абсолютные ошибки $\Delta\varphi_{ij}$.

В дальнейшем, при подробном изложении вопросов точности аппаратного анализа цепей мы рассмотрим два популярных подхода к интерпретации случайных ошибок амплитуды и фазы параметра рассеяния: в виде доверительного интервала при фиксированной доверительной вероятности и в виде среднеквадратичной ошибки; и два популярных подхода к интерпретации систематических ошибок амплитуды и фазы параметра рассеяния: в виде интервала неопределенности между возможными минимальным и максимальным значениями и в рамках квазистатистического подхода.

Диапазон измерений

Область значений измеряемой величины, в пределах которой нормированы погрешности средства измерений. Формально, по способу измерений векторные АЦ измеряют амплитуды и фазы (а скалярные АЦ – амплитуды) параметров рассеяния тестируемых устройств во всем диапазоне возможных значений этих параметров; например, для пассивных тестируемых устройств амплитуды – в диапазоне $0 \leq |S_{ij}| \leq 1$, $i, j = 1, \dots, N$. Однако, согласно определению,

диапазон измерений амплитуд, на самом деле, уже. Дело в том, что при приближении истинного значения модуля параметра рассеяния к нулю уровень входного сигнала на соответствующем приемнике также стремится к нулю, ошибка измерения отношения этого уровня к уровню опорного сигнала возрастает, а при нулевом значении модуля измерение становится невозможным, что соответствует бесконечной ошибке. Поэтому, если заданы *максимальные допустимые ошибки* измерения модулей $\delta_{\text{макс доп}} |S_{ij}|$, $i, j = 1, \dots, N$, то существуют *минимальные измеримые значения* этих модулей $|S_{ij}|_{\text{мин}}$, $i, j = 1, \dots, N$, и диапазоны измерения этих модулей (если нет других ограничивающих факторов, кроме названного) суть $|S_{ij}|_{\text{мин}} \leq |S_{ij}| \leq 1$, $i, j = 1, \dots, N$.

Аргументация относительно диапазона измерения фаз более сложна, но вывод тот же: *диапазон измерения фаз* параметров рассеяния уже диапазона $0 \leq \arg S_{ij} \leq 2\pi$, $i, j = 1, \dots, N$.

Чем шире диапазоны измерений амплитуд и фаз параметров рассеяния, тем выше эффективность АЦ.

Глубина частотной панорамы

Обычно АЦ допускает установку оператором диапазона перестраиваемой частоты зондирующего сигнала на дискретном множестве от некоторой начальной частоты $f_{\text{нач}}$ до некоторой конечной частоты $f_{\text{кон}}$. Предельный диапазон, образуемый минимально возможной $f_{\text{нач}}$, равной $f_{\text{мин}}$, и максимально возможной $f_{\text{кон}}$, равной $f_{\text{макс}}$ – важный фактор эффективности. Отношение $f_{\text{макс}} / f_{\text{мин}}$ (или его децибельную меру $10 \lg(f_{\text{макс}} / f_{\text{мин}})$) можно назвать *глубиной частотной панорамы*. Вообще говоря, чем ниже $f_{\text{мин}}$ и чем выше $f_{\text{макс}}$, тем, при прочих равных условиях, выше эффективность АЦ; расширяется множество устройств, которые способен тестировать АЦ и т.д. Однако, расширение предельного частотного диапазона усиливает требования к генератору зондирующих сигналов (может даже потребоваться применение двух или более переключаемых генераторов), блокам преобразования частот, системе отображения результатов измерения. Кроме того, при сохранения частотного шага расширение предельного частотного диапазона увеличивает максимальное число частотных точек, а это, в свою очередь, снижает оперативность прибора и увеличивает нагрузку на блоки координации и управления прибором, измерительной калибровки и коррекции данных, математической обработки сигналов и расчетов, отображения результатов, внутренней автоматике; при сохранении же числа частотных точек ухудшается частотное разрешение прибора, что означает огрубление частотного анализа и при испытании некоторых устройств может даже стать недопустимым.

Не существует какого-то единого решения для оптимального частотного диапазона АЦ СВЧ; обычно принимают $f_{\text{мин}}$ порядка десяти МГц, а $f_{\text{макс}}$ – порядка нескольких ГГц, то есть реализуют глубину частотной панорамы порядка или более трех декад. Наиболее критичен выбор максимальной частоты $f_{\text{макс}}$, т.к. он может сильно повлиять на выбор компонент (например, направленных ответвителей, смесителей и т.д.) и даже схемотехнической архитектуры, а это, в свою очередь, на точностные характеристики и себестоимость производства прибора.

Глубина мощностной панорамы

Как указывалось выше, для тестирования нелинейных и активных устройств и изучения мешающих нелинейных явлений АЦ должен иметь возможность варьировать выходную мощность генератора зондирующих сигналов, в том числе – в режиме сканирования этой мощности. Обычно АЦ допускает установку оператором диапазона перестраиваемой мощности зондирующего сигнала на дискретном множестве от некоторой начальной мощности $P_{\text{нач}}$ до

некоторой конечной мощности $P_{\text{кон}}$. Предельный диапазон, образуемый минимально возможной $P_{\text{нач}}$, равной $P_{\text{мин}}$, и максимально возможной $P_{\text{кон}}$, равной $P_{\text{макс}}$ – важный фактор эффективности. Отношение $P_{\text{макс}} / P_{\text{мин}}$ (или его децибельную меру $10\lg(P_{\text{макс}} / P_{\text{мин}})$) можно назвать *глубиной мощностной панорамы*. Вообще говоря, чем ниже $P_{\text{мин}}$ и чем выше $P_{\text{макс}}$, тем, при прочих равных условиях, выше эффективность АЦ: расширяется множество устройств, которые способны тестировать АЦ, возрастает глубина изучения нелинейных характеристик и т.д. Однако, расширение предельного мощностного диапазона усиливает требования к генератору зондирующих сигналов, блокам преобразования частот, системе отображения результатов измерения и может потребовать включения дополнительных каскадов усиления.

Динамический диапазон измерений

Это важный параметр эффективности АЦ. В п. 3. указано, что (статический) диапазон измерений модуля параметра рассеяния уже возможных значений этого модуля (для пассивных тестируемых устройств диапазон возможных значений $0 \leq |S_{ij}| \leq 1$, $i, j = 1, \dots, N$). Сейчас поставим более тонкий вопрос: насколько уже диапазон измерения, чем диапазон значений? Более точно: пусть нормированы погрешности измерения модуля $|S_{ij}|$; каковы крайние измеримые значения $|S_{ij}|_{\text{мин}}$, $i, j = 1, \dots, N$; $|S_{ij}|_{\text{макс}}$, $i, j = 1, \dots, N$ в рамках этих погрешностей.

Чтобы почувствовать, насколько актуален этот вопрос, приведем пример. Пусть измеряется коэффициент пропускания S_{21} полосового фильтра в некоторой области $(f_{\phi 1}, f_{\phi 2})$ частот. Для простоты предположим, что в области пропускания $f_1 \leq f \leq f_2$ коэффициент пропускания фильтра имеет максимальное значение 1 (0 дБ), а в области заграждения $(f_{\phi 1} \leq f < f_1) \cup (f_2 < f \leq f_{\phi 2})$ имеет минимальное значение (-G), дБ (например, G=100). Возможны два режима измерения: «долгий» и «быстрый». В долгом режиме коэффициент пропускания S_{21} измеряется отдельно в области пропускания $f_1 \leq f \leq f_2$ и отдельно – в области заграждения $(f_{\phi 1} \leq f < f_1) \cup (f_2 < f \leq f_{\phi 2})$, причем перед измерением в каждой области отдельно устанавливается выходная мощность $P_{\text{вых}}$ генератора зондирующих сигналов: меньшая для области пропускания, большая – для области заграждения. Достижимый в долгом режиме диапазон измерения назовем *статическим*. В быстром режиме коэффициент пропускания S_{21} измеряется одновременно, то есть в течение сканирования частоты (возможно – многократного) в диапазоне $(f_{\phi 1}, f_{\phi 2})$. Достижимый в этом режиме диапазон измерения следует назвать *динамическим*. Ясно, что измерение можно назвать успешным, только если динамический диапазон измерения $G_{\text{изм}}$ больше или равен перепаду (в децибелах) между минимальным (-G), дБ и максимальным 0 дБ значениями измеряемого коэффициента пропускания фильтра.

Не вдаваясь в детали, можно считать, что динамический диапазон измерения параметров рассеяния ограничивается сверху допустимыми рабочими мощностями зондирующих сигналов, и снизу – допустимыми (пороговыми) отношениями сигнал-шум. В свою очередь, допустимые рабочие мощности ограничиваются нелинейными искажениями и ошибками вследствие них, а также предельными входными мощностями некоторых элементов, а уровень шума зависит от полосы в последетекторном (последесмесительном) тракте. Поскольку отношение сигнал-шум зависит от (некогерентного) усреднения по нескольким сканированиям частоты, а при наличии когерентного усреднения в тракте ПЧ – от степени этого усреднения, существуют, по крайней мере, два понятия динамического диапазона: динамический диапазон до усреднения и динамический диапазон после усреднения.

Простой вывод из приведенных рассуждений: ВАЦ, с его узкополосным супергетеродинным приемником, должен иметь значительно больший динамический диапазон, чем САЦ, с его

широкополосным детекторным приемником, поскольку уровень шума в последетекторном тракте прямо зависит от полосы на входе детектора.

Разрешение по измеряемой величине

Мера различимости измерителем различных значений измеряемой величины. Задачи разрешения, восходящие к оптике, обычно ставятся как пороговые: предполагается, что существует такое (пороговое) различие двух значений измеряемой величины, что при большем различии эти значения воспринимаются как разные, а при меньшем – как одинаковые; это пороговое различие и есть мера *разрешающей* способности. Для построения *решающего правила* разрешения необходим *критерий* разрешения, связывающий решение с разностью или отношением (в общем случае – с функцией) оценок значений. Не существует единственного критерия разрешения, но при любом разумном критерии разрешающая способность связана с точностью измерения, степенью сглаживания частотных и временных зависимостей и т.д. Задачи разрешения, восходящие к оптике, обычно ставятся как пороговые: предполагается, что существует такое (пороговое) различие двух значений измеряемой величины, что при большем различии эти значения воспринимаются как разные, а при меньшем – как одинаковые; это пороговое различие и есть мера *разрешающей* способности. Для построения *решающего правила* разрешения необходим *критерий* разрешения, связывающий решение с разностью или отношением (в общем случае – с функцией) оценок значений. Не существует единственного критерия разрешения, но при любом разумном критерии разрешающая способность связана с точностью измерения, степенью сглаживания частотных и временных зависимостей и т.д.

Разрешение по частоте, мощности, времени

Другой смысл и характер носят разрешающие способности по аргументам, в функции которых представляются оценки измеряемых величин; в нашем случае эти аргументы: частота и мощность зондирующего сигнала – при анализе в частотной области, время задержки – при анализе во временной области. Этот вид разрешения однозначно задается при установке шага по частоте, по мощности, по времени, соответственно, и разрешающая способность просто равна этому шагу. При сглаживании по частоте, мощности или времени разрешающая способность уменьшается.

Преобразовательная чувствительность

Так предлагается назвать параметр, который обычно называется просто *чувствительностью*, то есть способность реагировать на изменения входного сигнала; она оценивается отношением изменения выходного сигнала к вызвавшему его изменению входного сигнала. Например, чувствительность отклонения луча осциллографа измеряется в миллиметрах на вольт. Поскольку АЦ не сводится просто к преобразователю, сопряженному с измерителем, целесообразно характеризовать чувствительности отдельных блоков и устройств АЦ. Свои чувствительности имеют система управления частотой генератора зондирующих сигналов, детекторы, смесители, АЦП, монитор и т.д.

Пороговая чувствительность

Этот параметр носит другой смысл и характер. Пороговую чувствительность АЦ можно характеризовать введенным при обсуждении понятия «диапазон измерений» минимальным измеримым значением $|S_{ij}|_{\text{мин}}$, $i, j = 1, \dots, N$ (пороговым значением) модуля $|S_{ij}|$ параметра рассеяния при заданной максимальной допустимой ошибке $\delta_{\text{макс доп}} |S_{ij}|$, $i, j = 1, \dots, N$. Заметим, что, вообще говоря, пороговые значения модулей коэффициентов пропускания $|S_{21}|, |S_{12}|$ отличаются от пороговых значений модулей коэффициентов отражения $|S_{11}|, |S_{22}|$. Пороговые чувствительности АЦ по параметрам рассеяния тестируемых устройств зависят от уровня шума и установленных степеней сглаживания и усреднения.

Быстродействие (оперативность)

Это важный фактор эффективности АЦ; от его величины зависит сама применимость данной модели АЦ в ряде приложений, например, при настройке и тестировании радиотехнических устройств на поточной производственной линии. Обычно *быстродействие* характеризуется интервалом времени, необходимым для производства единичного измерения. В случае ВАЦ удобно в качестве величины, характеризующей быстродействие, принять время, необходимое для измерения параметра рассеяния в одной частотной точке при анализе в частотной области. Это время включает ряд независимых и взаимозависимых составляющих, таких как время переключения генератора зондирующих сигналов с одной частоты на другую; время сбора данных измерения; время, необходимое для расчетов и выполнения процессорных процедур и т.д.

Полезно различать быстродействие собственно измерений, быстродействие с учетом калибровки, быстродействие с учетом калибровки и подготовительных процедур.

Стабильность

Отражает постоянство во времени метрологических показателей средства измерений. Часто эта характеристика представляется обратной величиной – *нестабильностью* показателей во времени. Общая нестабильность АЦ складывается из нестабильности (в частности, частотной) генератора зондирующих сигналов, переключателей, разъемов и т.д.

Помехозащищенность и электромагнитная защищенность

Помехозащищенность – способность электронного средства измерений (в частности, АЦ) сохранять в процессе измерений свои характеристики при наличии внешних радиопомех. *Электромагнитная совместимость* средства определяется гарантией достаточно низкого уровня его влияния на окружающие электромагнитные устройства, обеспечивающей сохранение свойств последних. Требования к помехозащищенности и электромагнитной совместимости измерительного средства определяются нормативными документами, а выполнимость этих требований зависит от состава, расположения и характеристик окружающих устройств.

Надежность

Свойство средства измерений функционировать при сохранении метрологических и других показателей в заданных пределах и режимах работы. Обычно надежность характеризуется свойствами *безотказности, ремонтпригодности, долговечности, сохраняемости*.

Для средств измерений показатели безотказности характеризуются средней *наработкой на отказ* (среднее значение наработки средства измерений между отказами) и *вероятностью безотказной работы за заданный промежуток времени*.

В качестве показателя *ремонтпригодности* для средств измерений обычно нормируется *среднее время восстановления*.

Долговечность средств измерений оценивается *гамма-процентным ресурсом* (наработка, в течение которой средство измерений не достигнет предельного состояния с вероятностью γ процентов) и *сроком службы* (календарной продолжительностью эксплуатации средства измерений до предельного состояния). *Предельное состояние* наступает тогда, когда обычные виды ремонта не позволяют поддерживать работоспособность прибора на требуемом уровне. В технических требованиях на средство измерений задаются обычно оба показателя долговечности, поскольку гамма-процентный ресурс определяется наработкой прибора (независимо от календарного времени эксплуатации), а срок службы – календарным временем эксплуатации (независимо от наработки). Обычно в технических требованиях под сроком службы понимается среднее его значение.

Средства измерений до поступления на эксплуатацию могут длительное время находиться на хранении. В связи с этим в технические требования часто включается в качестве показателя сохраняемости *средний срок сохраняемости* или *гамма-процентный срок сохраняемости*.

Эргономичность

Комплекс факторов, определяющих удобство работы оператора, в свою очередь влияющих на его производительность, вероятность ошибок и промахов и т.д.

Энергопотребление

Этот фактор эффективности особенно критичен при использовании АЦ на наземных, надводных, воздушных подвижных носителях: автомобилях, кораблях, самолетах и т.д.

Вес и габариты

Эти факторы не требуют комментариев.

Рабочие условия

Рабочие условия (условия измерений) содержатся в технических условиях на средство измерений и указывают на возможность проведения измерений с допустимыми изменениями метрологических характеристик. Рабочие условия измерений определяются назначением и степенью устойчивости метрологических характеристик данного средства измерений. Для унификации применяемых средств измерений рабочие условия измерений нормируются соответствующими государственными стандартами. Оценка условий измерений производится путем определения пределов изменения влияющих величин, то есть величин, оказывающих влияние на результаты измерений.

К влияющим величинам обычно относят:

- температуру, давление, влажность окружающей среды;
- частоту (диапазон частот), при которой производится измерение;
- механические нагрузки при транспортировании;
- напряжение и частоту питающей сети;
- напряженность магнитного (электрического) поля, в котором находится средство измерений;
- и т. д.

Наиболее часто в технических условиях на средство измерений указываются допустимые пределы рабочей температуры, относительная влажность (при той или иной температуре), напряжение и частота источников питания (с допустимыми отклонениями).

Поколения АЦ

Все более часто в последнее время радиотехнические средства измерений при характеристике их технических и конструктивных особенностей относят к тому или иному поколению. Это позволяет в какой-то степени оценивать прогресс эффективности средств измерения, улавливать и использовать тенденции развития и даже, в некоторой степени, прогнозировать развитие.

К *первому поколению* относят приборы с ручным управлением, применением электровакуумных приборов и объемного монтажа элементов (резисторов, конденсаторов и др.).

Ко *второму поколению* относят приборы с ручным управлением, полупроводниковыми приборами с применением как объемного, так и печатного монтажа элементов.

Третье поколение характерно применением микросхем и микросборок, одно- и двухслойных печатных плат, полуавтоматическим управлением (автоматическая установка нуля, самокалибровка).

К *четвертому поколению* относят средства измерений, имеющие автоматическое управление от встроенного микропроцессора и использующие микросхемы и микросборки большой степени интеграции, а также многослойные (двухслойные) печатные платы.

Можно уверенно констатировать, что современные АЦ вполне преодолели рубеж четвертого поколения и находятся на пути к пятому поколению. Обычно критерием отнесения средства измерений к соответствующему поколению являются наиболее характерные его черты. Вступая на зыбкую почву предположений, можно ожидать, что для *пятого поколения* характерной чертой будет полная готовность средств измерения к объединению в измерительные системы, с межкомпьютерной трансляцией данных и результатов, автоматическим формированием, выдачей и распечаткой отчетов с широким выбором номенклатуры и форматов, хранение установленных параметров, данных и результатов измерений в выбираемых номенклатурах и форматах и т. д.

Поскольку в основном радиотехнические приложения суть эксплуатация эффектов взаимодействия радиоизлучения с радиотехническими цепями и веществом, для АЦ наиболее гармоничным было бы сочетание в рамках измерительной системы указанного уровня с анализатором сигналов, способным анализировать характеристики и параметры сигналов в широком классе последних: детерминированных и случайных, узкополосных и широкополосных, в широком диапазоне частот и т.д.

Следует отметить еще одно обстоятельство. К настоящему времени средства измерений, применяемые в электронике и получившие общее наименование *радиоизмерительных приборов (РИП)*, достигли достаточно высокого уровня развития и в своем большинстве имеют наивысшие точности по сравнению со средствами измерений других физических величин – электрических, механических, тепловых и т.д. Более того, развитие средств измерений физических величин неэлектрического характера (длины, массы, времени, температуры, силы света и др.) за последние годы все более приближается по принципиальному и конструктивному выполнению к РИП: неэлектрическая величина с помощью измерительных датчиков преобразуется в электрическую (иногда в достаточно широком диапазоне частот), и в дальнейшем процесс усиления, преобразования и регистрации (отображения) измерительной информации основывается на принципах и методах электронных средств измерений. Кстати, отсюда следует все возрастающая необходимость применения анализаторов цепей к измерению элементов высокочастотных трактов измерителей неэлектрических величин. Это направление развития в ближайшем будущем должно получить широкое распространение, поскольку позволяет автоматизировать процесс измерений, унифицировать и стандартизировать различные виды средств измерений, обеспечить метрологическую, конструктивную, информационную, энергетическую и эксплуатационную совместимость средств измерений между собой и с объектами измерений. В связи с этим происходит непрерывное повышение точности и расширение диапазона измерений РИП. Особенно быстро возрастали требования к точности частотно-временных измерений. Наряду с требованиями к точности приборов все больше предъявляются требования к их *быстродействию, производительности*. В связи с этим широкое распространение получили цифровые приборы, позволяющие наряду с высокими показателями точности обеспечить быстродействие, высокую производительность измерений, программное управление приборами, сопряжение с универсальными ЭВМ.

Появление новых принципов и методов измерений, расширение диапазонов измерений привело к расширению номенклатуры применяемых измерительных приборов, которая сейчас классифицируется на более чем 100 подгрупп и видов. Основными видами являются измерители напряжений и токов, параметров электрических цепей и трактов, частоты, формы сигналов, их спектров, а также измерительные генераторы. Получили развитие приборы для измерения характеристик случайных процессов, логические анализаторы (для определения правильности структуры цифровых меток, потока цифровых данных в контрольных точках электронных цифровых устройств, в том числе и цифровых измерительных приборов) и др.

Очень быстрыми темпами развивается направление анализаторов сигналов в реальном масштабе времени. Эти приборы все большее распространение находят не только как анализаторы спектра сигналов, но и как *измерители амплитудных и фазочастотных характеристик радиоэлектронных устройств*, функций корреляции, когерентности, плотностей вероятности случайных процессов, полей и последовательностей потоков информации. В то же

время качественные показатели анализаторов цепей в частотной и временной областях не в последнюю очередь зависят от развития техники анализа высокочастотных сигналов в реальном времени).

Большое значение имеет тенденция к *виртуализации* приборов. К *виртуальным* принято относить приборы многофункциональные, позволяющие по выбору оператора осуществлять функции то одного, то другого прибора. С таким прибором можно проводить измерения по определенной программе, задаваемой оператором или извне контроллером. Все задачи по восприятию программы измерений, перестраиванию субмодулей и отдельных схем прибора в необходимое сочетание для выполнения различных функций, организации измерений выполняет встроенный микропроцессорный блок.

Важна и достаточная надежность приборов четвертого поколения, достигаемая высокой степенью интеграции применяемых БИС и СБИС, почти полным исключением из конструкции приборов малонадежных механических узлов (переключателей, электромеханических реле, перестраиваемых элементов и т. д.).

Таким образом, мы приходим к важному выводу, определяющему направление развития электронных измерительных приборов, в частности, анализаторов цепей СВЧ: Нужно стремиться создавать радиоизмерительные приборы, совместимые между собой и приспособленные к объединению (агрегатированию) в автоматизированные измерительные комплексы для решения большого числа измерительных задач. Подобные автоматизированные измерительные комплексы (АИК), создаваемые (иногда – просто собираемые) на основе объединения стандартных, серийных средств измерений, должны управляться через стандартную информационную магистраль внешним компьютером, который регулирует сбор, обработку измерительной информации, передачу и хранение полученных данных и результатов.

Аппаратный анализ цепей во временной области

Рефлектометрия во временной области (РВО)

РВО – технология, позволяющая измерять импедансы в функции расстояния. Важные применения этого вида измерения – конструирование оснастки (обнаружение позиции отражения в оснастке и коэффициент отражения) и калибровочных эталонов в оснастке (определяется их качество). В частности, можно различать емкостные и индуктивные переходы, индцировать ЛП с импедансом, отличным от R_0 . В пределах пространственного разрешения мы можем видеть отражения от разъемных запусков, независимо от отражений от калибровочных стандартов.

Используя стробирование во временной области, можно развязать различные части оснастки и видеть эффекты на этих частях в частотной области. Например, можно отдельно посмотреть на частотную зависимость коэффициента отражения от разъемного запуска без остальной части эталона.

Другое применение РВО – определение положений дефектов на коаксиальном кабеле (например, в сотовых и кабельно–телевизионных применениях): прецизионное определение положения кабельного дефекта, например, завитка, плохого соединения, к.з., х.х., т.е. чего-то такого, что приводит к отражению части сигнала.

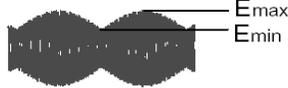
Два вида анализа во временной области: физический, математический.

ДОПОЛНИТЕЛЬНЫЕ ПОДРОБНОСТИ АНАЛИЗА ЦЕПЕЙ (по Балло)

Параметры отражения (илл. из Балло, с. 6, с. 8)

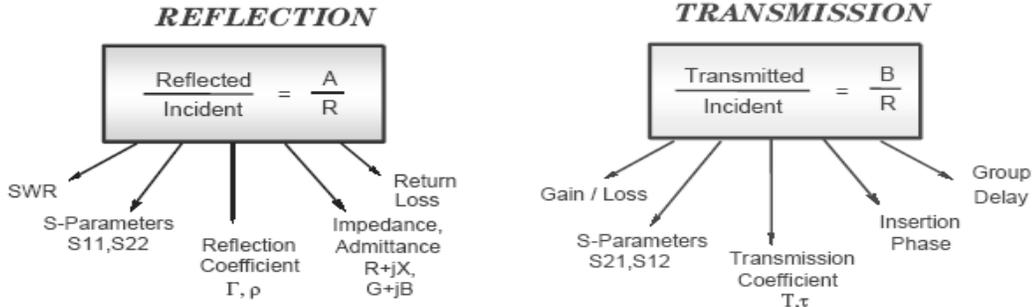
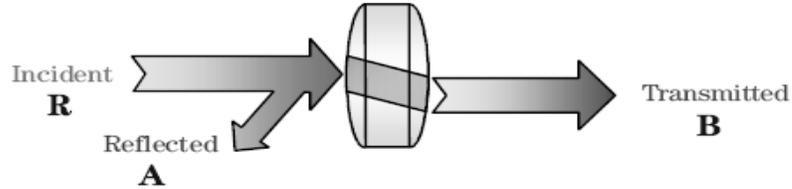
Reflection Coefficient $\Gamma = \frac{V_{\text{reflected}}}{V_{\text{incident}}} = \rho \angle \Phi = \frac{Z_L - Z_0}{Z_L + Z_0}$

Return loss = $-20 \log(\rho)$, $\rho = |\Gamma|$



Voltage Standing Wave Ratio
 $VSWR = \frac{E_{\text{max}}}{E_{\text{min}}} = \frac{1 + \rho}{1 - \rho}$

	No reflection ($Z_L = Z_0$)	Full reflection ($Z_L = \text{open, short}$)
ρ	0	1
RL (dB)	∞	0
VSWR	1	∞



Параметры передачи (илл. Балло, с. 8)

Transmission Parameters

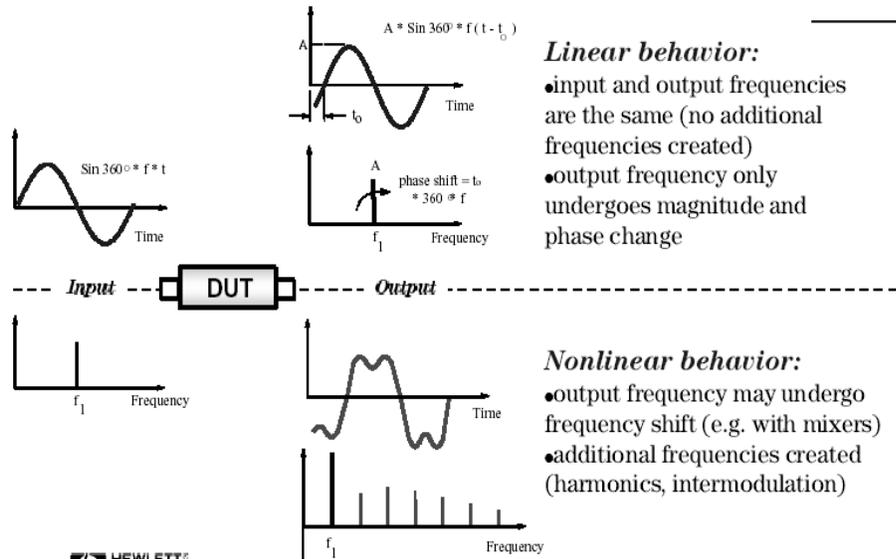


Transmission Coefficient = $T = \frac{V_{\text{Transmitted}}}{V_{\text{Incident}}} = \tau \angle \phi$

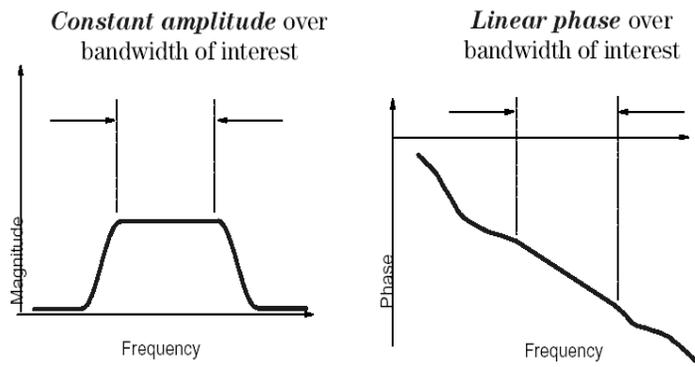
Insertion Loss (dB) = $-20 \text{ Log} \left| \frac{V_{\text{Trans}}}{V_{\text{Inc}}} \right| = -20 \log \tau$

Gain (dB) = $20 \text{ Log} \left| \frac{V_{\text{Trans}}}{V_{\text{Inc}}} \right| = 20 \log \tau$

Линейное или нелинейное поведение цепи (илл. Балло, с. 9)

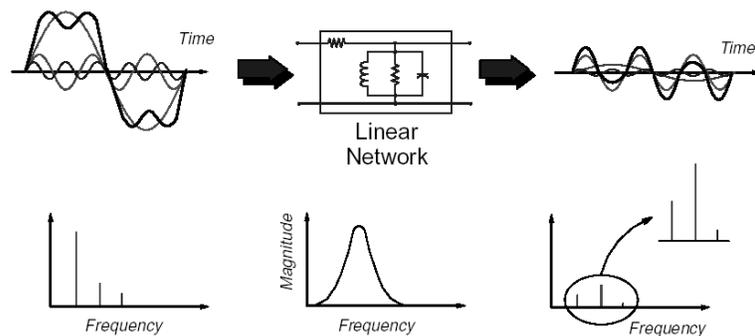


Линейные искажения (илл. Балло, с. 9)



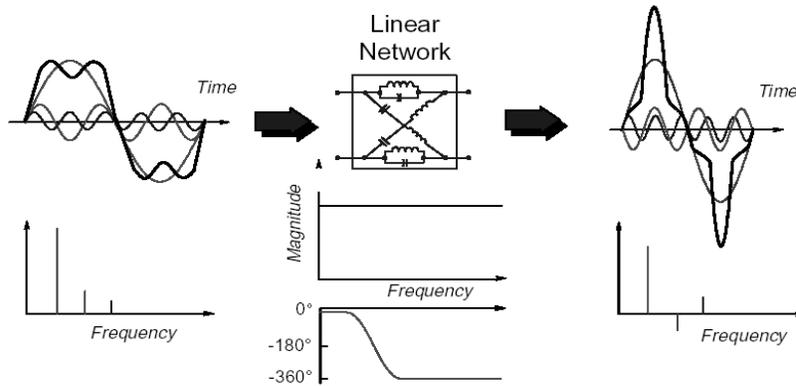
Вариации амплитуды с частотой (илл. Балло, с. 10)

$$F(t) = \sin \omega t + 1/3 \sin 3\omega t + 1/5 \sin 5\omega t$$



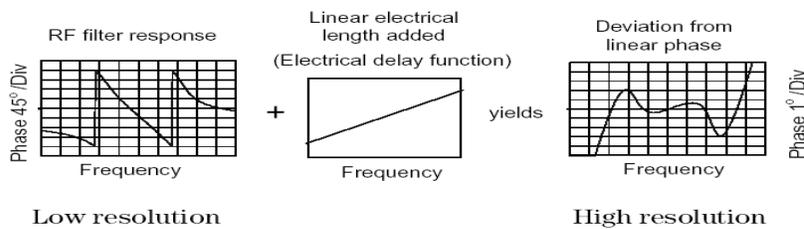
Вариации фазы с частотой (илл. Балло, с. 10)

$$F(t) = \sin \omega t + 1/3 \sin 3\omega t + 1/5 \sin 5\omega t$$

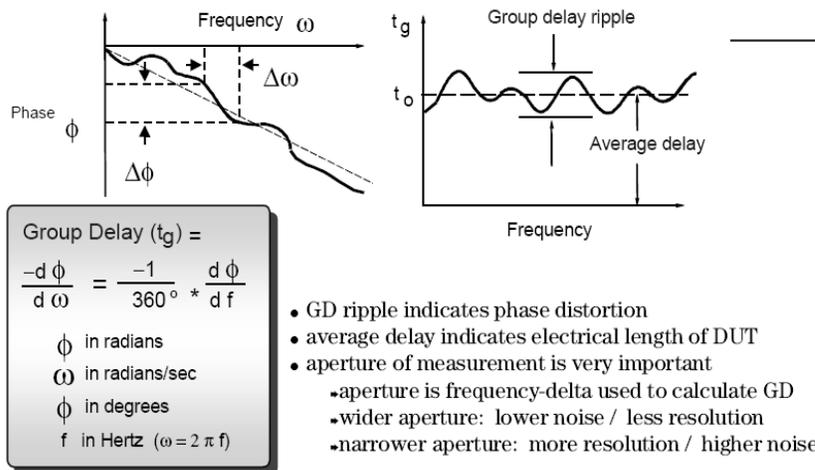


Отклонение от линейной фазы (илл. Балло, с. 11)

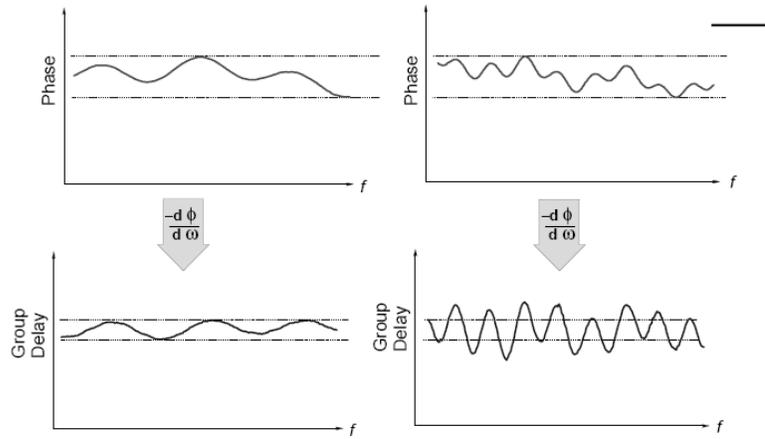
Use electrical delay to remove linear portion of phase response



Групповая задержка (илл. Балло, с. 11)



Зачем измерять групповую задержку? (илл. Балло, с. 12)

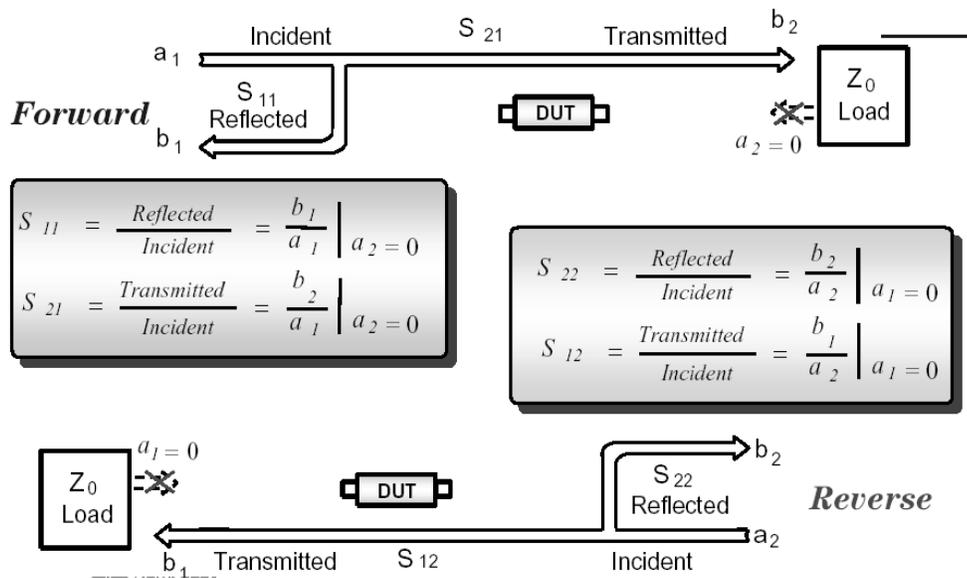


Same p-p phase ripple can result in different group delay

Почему измеряется как отклонение от линейной фазы, так и групповая задержка? В зависимости от прибора оба могут быть важны. Установление размаха фазовых пульсаций не является достаточным для полной характеристики прибора, поскольку наклон фазовых пульсаций зависит от числа пульсаций на единицу частоты. Групповая задержка учитывает это, поскольку она дифференциальна к фазовому отклику. Групповая задержка часто более легко интерпретируемая характеристика фазовых искажений.

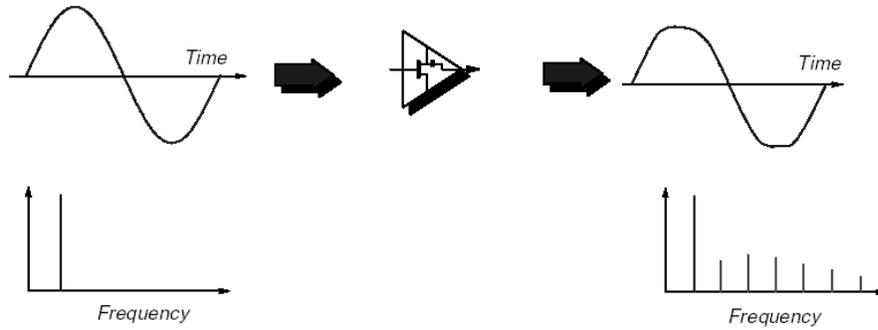
Картинки сверху показывают, что при том же размахе фазовых пульсаций можно получить существенно различные отклики групповых задержек. Отклик справа с большей вариацией групповой задержки означает большее искажение сигнала.

Измерение S-параметров (илл. Балло, с. 14)

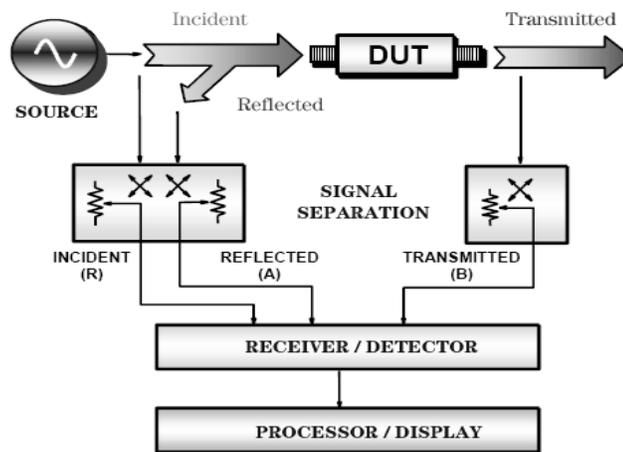


Нелинейные искажения (насыщение, перекрестные искажения, интермодуляция, гистерезис, амплитудно-фазовая конверсия) (илл. Балло, с. 15)

Saturation, crossover, intermodulation, and other nonlinear effects can cause signal distortion



Обобщенная блок-схема АЦ (илл. Балло, с. 17)

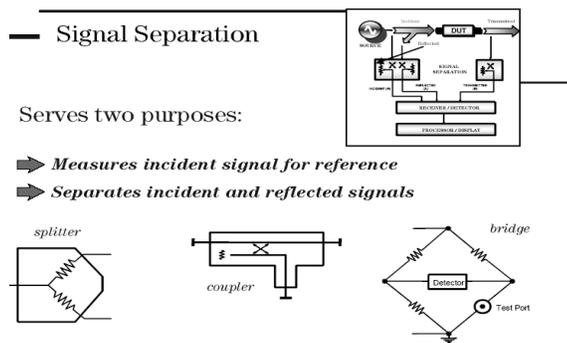


Источник зондирующих сигналов (см. выше)

Источник сигнала обеспечивает возбуждение нашей тестовой системы. Мы можем или качать частоту источника, или качать его уровень мощности. Традиционно анализатор цепей использует отдельный источник. Эти источники или основаны на ГУНе с открытой петлей обратной связи, что дешевле, или более дорогие синтезаторы частот, которые обеспечивают более высокое качество, особенно для измерения узкополосных устройств.

Интенсивный фазовый шум ГУНа с открытой петлей значительно деградирует точность измерения, когда тестируется узкополосная компонента при малой амплитуде качания частоты. Большинство анализаторов цепей фирмы HP, продаваемых сегодня, имеет интегрированный синтезированный источник.

Устройства разделения сигналов (дивайдеры и сплиттеры, делители и расщепители)

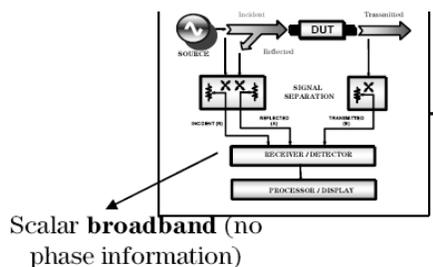
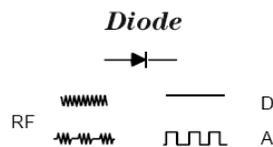


Оборудование разделения сигналов должно обеспечивать две функции. Первая, измерять часть падающего сигнала для получения опорной величины к рациионированию (получению относительной величины). Это можно сделать расщепителем или направленным ответвителем. Расщепители обычно резистивны. Они не направленные приборы (подробнее о направленности далее) и могут быть очень широкополосны. Плата за это в том, что обычно они имеют 6 дБ или более потерь в каждой ветви. Направленные ответвители имеют очень низкие вносимые потери (по главной ветви) и хорошие развязку и направленность.

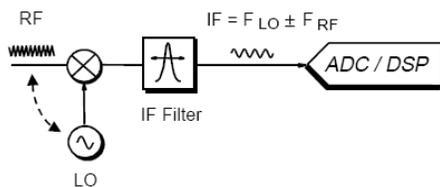
Вторая функция сигнально-разделительного оборудования, разделить падающую (вперед идущую) и отраженную (обратную) волны на входе тестируемого прибора. Снова, ответвители идеальны в том, что они направлены, имеют малые потери и высокую обратную развязку. Однако, из-за трудностей изготовления истинно широкополосных ответвителей, часто используются мосты. Более подробно об устройствах сигнального разделения в Приложении.

Приемники (широкополосные и узкополосные) (илл. Балло, с. 20, с. 21, с. 21)

Detector Types



Tuned Receiver



Scalar **narrowband** (magnitude only)

Vector (magnitude and phase)

Есть два основных пути обеспечения детектирования сигнала в анализаторах цепей. Диодное детектирование преобразует уровень радиочастотного сигнала в пропорциональный постоянный уровень. Если сигнал амплитудно модулирован, диод убирает радиочастотное заполнение из модуляции. Диодное детектирование по природе скалярно, так как фазовая информация радиочастотной несущей теряется.

Подстраиваемый приемник использует гетеродин для преобразования вниз на более низкую промежуточную частоту. Гетеродин замкнут или на радиочастотный, или на промежуточный сигнал, так что приемник анализатора цепей уже подстроен к радиочастотному сигналу, присутствующему на входе. Сигнал ПЧ пропускается через полосовой фильтр, который обужает полосу приемника и сильно улучшает чувствительность и динамический диапазон. Современные анализаторы используют аналого-цифровой преобразователь (АЦП) и цифровую обработку

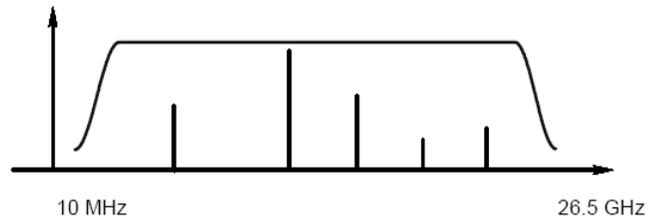
сигнала (ЦОС) для выделения амплитудной и фазовой информации из сигнала ПЧ. Подход с подстраиваемым приемником можно использовать в скалярном или векторном анализаторе цепей.

Сравнение методов приема (илл. Балло, с. 22)

Широкополосное детектирование



- Easy to make **broadband**
- **Inexpensive** compared to tuned receiver
- Good for measuring frequency-translating devices
- Improve dynamic range by increasing power
- **Medium** sensitivity / dynamic range

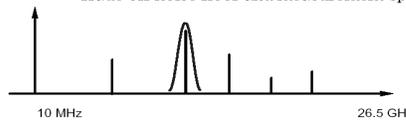


- Легко сделать **широкополосным**
- **Дешево** в сравнении с подстраиваемым приемником
- Хорош для тестирования частотно-преобразующих приборов
- Улучшает динамический диапазон путем увеличения мощности
- **Средние** чувствительность и динамический диапазон

Узкополосное детектирование

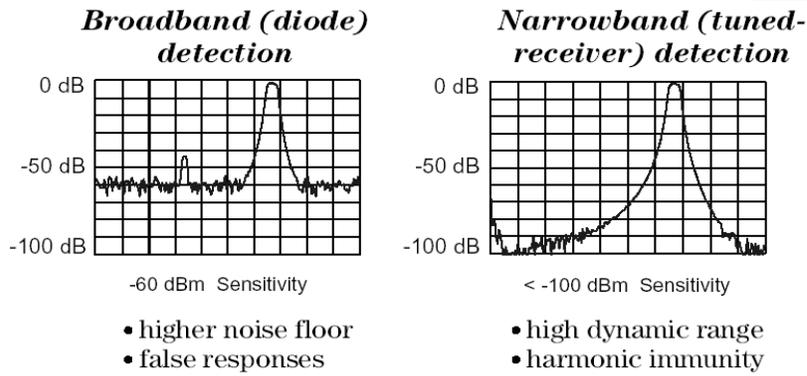


- **Best** sensitivity / dynamic range
- Provides harmonic / spurious signal **rejection**
- Improve dynamic range by increasing **power**, decreasing IF **bandwidth**, or **averaging**
- Trade off noise floor and measurement speed



- Лучшая чувствительность/динамический диапазон
- Обеспечение **вырезания** гармоник/побочных сигналов
- Улучшение динамического диапазона увеличением **мощности**, уменьшением **полосы ПЧ**, или **усреднением**
- Дихотомия шумового пола и скорости измерения

Динамический диапазон



Dynamic range = maximum receiver power - receiver noise floor

Два типа переключателей

Есть два типа переключателей передачи, которые могут использоваться в установке с S-параметрами: твердотельный и механический. Твердотельные переключатели имеют преимущество бесконечного времени жизни (в предположении, что они не повреждаются слишком большой мощностью от тестируемого прибора). Однако, они дают большие, чем механические, потери, поэтому они снижают максимальную выходную мощность анализатора. Механические переключатели имеют очень низкие потери и позволяют достичь более высоких выходных мощностей. Их главный недостаток в том, что с течением времени они изнашиваются (после 5 миллионов циклов, или около того). Когда используют анализатор с механическими переключателями, измерения обычно делают в одновипповом режиме, поэтому переключатель не работает непрерывно.

Процессор и дисплей для расчетов и отображения результатов (илл. Балло, с. 24)

Последний главный блок оборудования в анализаторе это секция дисплей/процессор. Здесь данные прохождения и отражения форматируются так, чтобы сделать легко интерпретируемыми измерительные результаты. Большинство анализаторов имеют такие функции, как линейные и логарифмические качания, линейные и логарифмические форматы, полярные системы, диаграммы Смита и т.д. Другие общие свойства – трассовые маркеры, граничные линии, тестирование срывов пропускания. Многие из анализаторов НР имеют специальные измерительные свойства, приспособленные для частных рынков или применений. Один пример: подстраиваемый анализатор НР 8730А.

Модели ошибок

Виды ошибок (случайные, систематические, дрейфовые)

Systematic errors

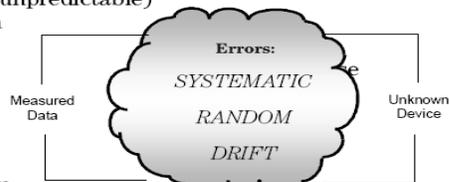
- due to **imperfections** in the analyzer and test setup
- are assumed to be **time invariant** (predictable)
- can be characterized (during calibration process) and **mathematically removed** during measurements

Random errors

- **vary** with time in random fashion (unpredictable)
- ✗ **cannot be removed** by calibration
- main contributors:
 - **instrument noise** (source noise, IF noise floor, etc.)
 - **switch** repeatability
 - **connector** repeatability

Drift errors

- are due to instrument or test-system performance changing **after** a calibration has been done
- are primarily caused by **temperature variation**
- can be removed by further calibration(s)



Случайные ошибки (флуктуации зондирующего сигнала, шумы, другие флуктуации)

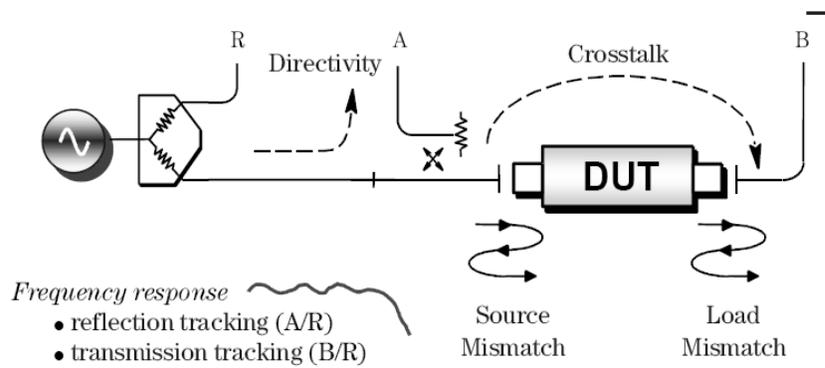
- Варьируют во времени случайным образом (непредсказуемы)
- Не могут быть исключены калибровкой
- Главные составляющие:
 - инструментальный шум (источники шума, шумовой пол на ПЧ и т.д.)
 - повторяемость переключателя
 - повторяемость разъема

Борьба со случайными ошибками (уменьшение коэффициента шума цепей, усреднение и сглаживание, оптимальные оценки)

Дрейфовые ошибки

- Из-за изменений инструментальной и тестовой характеристик после сделанной калибровки
- Из-за первичных проявлений температурных вариаций
- Могут убираться путем дальнейших калибровок

Модели факторов систематических ошибок (12-факторная модель, другие модели) (илл. Балло, с. 30)



***Six forward and six reverse error terms yields
12 error terms for two-port devices***

Показанные здесь главные систематические ошибки связаны с измерениями цепей. Ошибки, связанные с протечками сигнала, это направленность и перекрестная помеха. Ошибки, относящиеся к отражению сигнала, это согласование на источнике и на нагрузке. Ошибки, относящиеся к частотному отклику приемника, это сопровождение прохождения и сопровождение отражения. Полная двухпортовая модель ошибок включает все шесть этих составляющих для прямого направления и те же шесть (с другими данными) в обратном направлении, итого 12 составляющих ошибки. Вот почему мы часто относим двухпортовую калибровку к двенадцати-членной коррекции ошибок.

Уравнения для измеренных значений параметров рассеяния (см. свое изложение) (потокосные графы: прямой и обратный, их решение, уравнения для S_{ijm})

Тип ВАЦ

Далее по умолчанию: ВАЦ двухпортовый, для измерения S-параметров (с переключателем), с двумя измерительными НО, с конфигурируемым измерительным блоком (6 выведенных дополнительных пар разъемов), с АРМ источника, на коаксиальных линиях и коаксиальных разъемах, два измерительных приемника с преобразованием частоты, два опорных приемника, 2 аттенюатора источника, 1 переключатель, 1 источник-синтезатор, 1 гетеродин-синтезатор.

Объект тестирования

Устройство двухпортовое, пассивное, линейное, без переноса частоты, взаимное, с аналоговыми входом и выходом. Подключается коаксиальными ЛП с обычными разъемами.

Источники (факторы) систематических ошибок

Пока учитываются «классические» источники (факторы): с каждой стороны зондирования 2 протечки, 2 согласования, 2 трекинга.

2 протечки: за счет небесконечной направленности и за счет неидеальной изоляции между 1-ым и 2-ым портами ВАЦ;

2 согласования: за счет рассогласования источника и рассогласования выходного порта ВАЦ;

2 трекинга: трекинг (нуль отсчета) отражения и трекинг пропускания.

Потом будут учтены другие факторы ошибок: разъемы, дополнительные кабели, адаптеры и другие.

Фиксируемые параметры

Ошибки и соотношения для них параметрически зависят от частоты источника, мощности источника, положения аттенуаторов, полосы ПЧ, режима видео-усреднения, времени стояния на одной частотной точке, параметров эталонов конкретного кита, . Значения этих параметров образуют вектор параметров $\bar{\alpha}$.

Виды описания систематических ошибок измерения S-параметров

- Доверительный интервал со 100-процентной доверительной вероятностью, строится с помощью максимальной абсолютной ошибки.
- Доверительный интервал с заданной доверительной вероятностью.
- Распределение вероятностей (плотность вероятностей квазираспределения) ошибки S-параметра.
- Среднеквадратичное отклонение квазираспределения ошибки S-параметра.

Объекты измерения

Поскольку измеряемые ВАЦем величины S_{ij} ($i, j = 1, 2$) – комплексные, то объекты измерения – модули $|S_{ij}|$ и фазы $\arg S_{ij}$ S-параметров.

Обозначения ошибок S-параметров

$\Delta^+ |S_{ij}|$ – максимальная верхняя абсолютная ошибка модуля $|S_{ij}|$ S-параметра;

$\Delta^- |S_{ij}|$ – максимальная нижняя абсолютная ошибка модуля $|S_{ij}|$ S-параметра;

$\Delta^+ \arg S_{ij}$ – максимальная верхняя абсолютная ошибка фазы $\arg S_{ij}$ S-параметра;

$\Delta^- \arg S_{ij}$ – максимальная нижняя абсолютная ошибка фазы $\arg S_{ij}$ S-параметра;

Физический смысл и обозначения факторов ошибок

Перечисленные в п. 1.3. факторы создают *двенадцатичленную модель ошибок*, включающую (комплексные) *факторы ошибок*. Дадим их определения при комплексной амплитуде сигнала источника $\dot{U}_{ист} = 1$.

При зондировании слева направо:

E_{DF} – ошибка направленности: обусловленная дефицитом направленности левого НО комплексная амплитуда протечки сигнала источника в отражательный порт этого НО. Модуль $|E_{DF}|$ удовлетворяет соотношению

$$|E_{DF}| = C_1 / D_1, \quad (1к)$$

где C_1 – коэффициент связи НО по напряжению (см. раздел «НО»),

D_1 – коэффициент направленности НО по напряжению (см. раздел «НО»).

E_{SF} – ошибка согласования источника: комплексный коэффициент отражения источника.

E_{LF} – ошибка согласования нагрузки: комплексный коэффициент отражения нагрузочного порта ВАЦ.

E_{RF} – трекинг (нуль отсчета) отражения (мультипликативная ошибка):

$$E_{RF} = 1 + E'_{RF}. \quad (2к)$$

E_{TF} – трекинг (нуль отсчета) пропускания (мультипликативная ошибка):

$$E_{TF} = 1 + E'_{TF}. \quad (3к)$$

E_{XF} – протечка (изоляция) в нагрузочный порт ВАЦ: комплексная амплитуда протечки в правый порт.

При зондировании справа налево:

те же 6 составляющих, с заменой второго индекса с F на R .

Обозначения составляющих ошибок (из файла «Специфические калибровочные стандарты»).

E_{DF}, E_{DR} -Directivity	E_{LF}, E_{LR} -Load Match
E_{SF}, E_{SR} -Source Match	E_{TF}, E_{TR} -Trans. Tracking
E_{RF}, E_{RR} -Refl. Tracking	E_{XF}, E_{XR} -Isolation

Пояснения: Первая буква индекса указывает на вид составляющей ошибки, вторая буква – откуда зондирующий сигнал, E – составляющая ошибки.

Обозначение вида составляющей ошибки:

D – направленность (от directivity)

S – согласование источника (от source match)

R – отражательное сопровождение (от reflection tracking)

L – согласование нагрузки (от load match)

T – сопровождение пропускания (от trans. tracking)

X – изоляция

Откуда зондирующий сигнал (вторая буква)

F – прямое зондирование (от forward)

R – обратное зондирование (от reverse)

Обозначения измеряемого и истинного S-параметра

Третий индекс S-параметра: M – измеряемый (от measure), A – истинный (от actuality)

Исходная система уравнений

Исходная система уравнений для анализа докалибровочных и послекалибровочных (остаточных) ошибок получается из уравнений оценок S-параметров тестируемого устройства:

$$S_{ijM} = f(S_A, \vec{E}; \vec{\alpha}), \quad i, j = 1, 2, \quad (4к)$$

где S_{ijM} – оценки S-параметров тестируемого устройства (M – от measure);

S_A – матрица истинных S-параметров тестируемого устройства (от Actuality);

$\vec{E} = (E_{DF}, E_{SF}, E_{LF}, E_{RF}, E_{TF}, E_{XF})$ – вектор факторов ошибок;

$\vec{\alpha}$ – вектор параметров (см. п. 1.4).

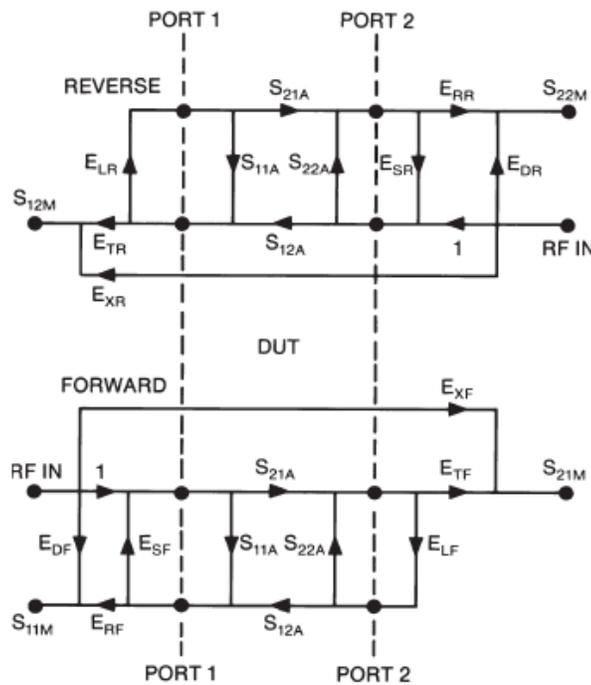


Figure 1. Agilent 8510 full 2-port error model

Исходную систему уравнений (4к) можно получить, решив потоковые графы, соответствующие зондированию слева и справа и учитывающие факторы ошибок в полной модели ошибок. Эти графы приведены на рис. 1к, 2к. Решая эти графы методом неприкасающихся ветвей относительно оценок S_{ijM} , получаем искомую систему:

$$S_{11M} = E_{DF} + \frac{S_{11A}E_{RF}(1 - S_{22A}E_{LF}) + S_{21A}E_{LF}S_{12A}E_{RF}}{1 - S_{11A}E_{SF} - S_{22A}E_{LF} - S_{21A}E_{LF}S_{12A}E_{SF} + S_{11A}E_{SF}S_{22A}E_{LF}} \quad (5к)$$

$$S_{21M} = E_{XF} + \frac{S_{21A} E_{TF}}{1 - S_{11A} E_{SF} - S_{22A} E_{LF} - S_{21A} E_{LF} S_{12A} E_{SF} + S_{11A} E_{SF} S_{22A} E_{LF}}. \quad (6к)$$

$$S_{22M} = E_{DR} + \frac{S_{22A} E_{RR} (1 - S_{11A} E_{LR}) + S_{12A} E_{LR} S_{21A} E_{RR}}{1 - S_{22A} E_{SR} - S_{11A} E_{LR} - S_{12A} E_{LR} S_{21A} E_{SR} + S_{22A} E_{SR} S_{11A} E_{LR}}. \quad (7к)$$

$$S_{12M} = E_{XR} + \frac{S_{12A} E_{TR}}{1 - S_{22A} E_{SR} - S_{11A} E_{LR} - S_{12A} E_{LR} S_{21A} E_{SR} + S_{22A} E_{SR} S_{11A} E_{LR}}. \quad (8к)$$

Вид уравнений (5к)-(8к) можно упростить, если обозначить

$$\det \mathbf{S}_A = S_{11A} S_{22A} - S_{21A} S_{12A} \quad (9к)$$

– комплексный определитель матрицы \mathbf{S}_A . Тогда

$$S_{11M} = E_{DF} + \frac{S_{11A} E_{RF} - \det \mathbf{S}_A \cdot E_{RF} E_{LF}}{1 - S_{11A} E_{SF} - S_{22A} E_{LF} + \det \mathbf{S}_A \cdot E_{SF} E_{LF}}; \quad (10к)$$

$$S_{22M} = E_{DR} + \frac{S_{22A} E_{RR} - \det \mathbf{S}_A \cdot E_{RR} E_{LR}}{1 - S_{22A} E_{SR} - S_{11A} E_{LR} + \det \mathbf{S}_A \cdot E_{SR} E_{LR}}; \quad (11к)$$

$$S_{21M} = E_{XF} + \frac{S_{21A} E_{TF}}{1 - S_{11A} E_{SF} - S_{22A} E_{LF} + \det \mathbf{S}_A \cdot E_{SF} E_{LF}}; \quad (12к)$$

$$S_{12M} = E_{XR} + \frac{S_{12A} E_{TR}}{1 - S_{22A} E_{SR} - S_{11A} E_{LR} + \det \mathbf{S}_A \cdot E_{SR} E_{LR}}. \quad (13к)$$

Максимальная ошибка модуля S-параметра

Максимальная верхняя абсолютная ошибка модуля $|S_{ij}|$ S-параметра $\Delta^+ |S_{ij}|$ найдется как

$$\Delta^+ |S_{ij}| = \left| \varphi(\mathbf{S}_A, \vec{E}_{\text{остат}}; \vec{\alpha}) \right| - |S_{ijA}|, \quad (14к)$$

где $\varphi(\bullet, \bullet; \bullet)$ получается из $f(\bullet, \bullet; \bullet)$ путем синфазного сложения комплексных слагаемых и замены \vec{E} на $\vec{E}_{\text{остат}}$. Вектор $\vec{E}_{\text{остат}}$ получается из вектора \vec{E} путем замены первоначальных (комплексных) факторов ошибок на модули факторов, скорректированные при калибровке.

Замечание: это означает, что если калибровка неполная, т. е. корректируются не все факторы ошибок, то в функции $\varphi(\dots)$ вектор $\vec{E}_{\text{остат}}$ включает факторы по итогу: те, что калибровались, берутся в остаточном виде, те, что не калибровались, берутся в начальном виде. В свою очередь, это означает, что нижеследующие формулы годятся для любого анализатора (скалярный, векторный, любая модель) и любого типа калибровки (полная-неполная, двупортовая-однопортовая, механическая-электронная, с хорошими или плохими эталонами, вообще без калибровки и т. д.), надо только брать те значения модулей факторов, которые получаются экспериментально после калибровки.

Максимальная ошибка фазы S-параметра

Максимальная верхняя абсолютная ошибка фазы $\arg S_{ij}$ S-параметра $\Delta^+ \arg S_{ij}$ найдется как

$$\Delta^+ \arg S_{ij} = \arcsin \frac{\Delta^+ |S_{ij}|}{|S_{ijA}|}. \quad (15к)$$

2. Максимальные абсолютные ошибки амплитуд S-параметров

2.1. Первый порядок факторов ошибок, результаты в линейном масштабе

Согласно (14к), в первом порядке факторов ошибок находим:

$$\Delta^+ |S_{11}| = |E_{DF}| + |S_{11A}| \cdot |E'_{RF}| + |S_{11A}|^2 \cdot |E_{SF}| + |S_{21A}| \cdot |S_{12A}| \cdot |E_{LF}|; \quad (16к)$$

$$\Delta^+ |S_{22}| = |E_{DR}| + |S_{22A}| \cdot |E'_{RR}| + |S_{22A}|^2 \cdot |E_{SR}| + |S_{21A}| \cdot |S_{12A}| \cdot |E_{LR}|; \quad (17к)$$

$$\Delta^+ |S_{21}| = |E_{XF}| + |S_{21A}| \cdot (|E'_{TF}| + |S_{11A}| \cdot |E_{SF}| + |S_{22A}| \cdot |E_{LF}|); \quad (18к)$$

$$\Delta^+ |S_{12}| = |E_{XR}| + |S_{12A}| \cdot (|E'_{TR}| + |S_{22A}| \cdot |E_{SR}| + |S_{11A}| \cdot |E_{LR}|). \quad (19к)$$

Отметим, что $\Delta^+ |S_{11}|$ и $\Delta^+ |S_{22}|$ выражаются квадратично через $|S_{11A}|$ и $|S_{22A}|$ соответственно, и зависят от одного параметра $|S_{21A}| \cdot |S_{12A}|$; в то же время $\Delta^+ |S_{21}|$ и $\Delta^+ |S_{12}|$ выражаются линейно через $|S_{21A}|$ и $|S_{12A}|$ соответственно, и зависят от двух параметров $|S_{11A}|$ и $|S_{22A}|$.

2.2. Первый порядок факторов ошибок, результаты в децибельном масштабе

Ошибки, определенные по формулам (16к)-(19к), имеют линейный масштаб. Ошибки модулей S-параметров в децибельном масштабе найдем как

$$\delta_{ijdB}^{\pm} = 20 \lg \frac{|S_{ijA}| \pm \Delta^+ |S_{ij}|}{|S_{ijA}|} = 20 \lg \left(1 \pm \frac{\Delta^+ |S_{ij}|}{|S_{ijA}|} \right).$$

При выполнении условия

$$\Delta^+ |S_{ij}| \ll |S_{ijA}|,$$

имеем

$$\delta_{ijdB}^{\pm} \approx \pm 8,6858 \frac{\Delta^+ |S_{ij}|}{|S_{ijA}|}.$$

3. Максимальные абсолютные ошибки фаз S-параметров

3.1. Первый порядок факторов ошибок, результаты в радианах

Согласно (15к), в первом порядке факторов ошибок и с учетом приближения $\arctg(x) \approx x$, $x \ll 1$, находим:

$$\Delta^+ \arg S_{11} = |E'_{RF}| + |S_{11A}| \cdot |E_{SF}| + \frac{|E_{DF}| + |S_{21A}| \cdot |S_{12A}| \cdot |E_{LF}|}{|S_{11A}|}, \quad \Delta^+ |S_{11}| \square |S_{11A}|;$$

$$\Delta^+ \arg S_{22} = |E'_{RR}| + |S_{22A}| \cdot |E_{SR}| + \frac{|E_{DR}| + |S_{21A}| \cdot |S_{12A}| \cdot |E_{LR}|}{|S_{22A}|}, \quad \Delta^+ |S_{22}| \square |S_{22A}|;$$

$$\Delta^+ \arg S_{21} = |E'_{TF}| + |S_{11A}| \cdot |E_{SF}| + |S_{22A}| \cdot |E_{LF}| + \frac{|E_{XF}|}{|S_{21A}|}, \quad \Delta^+ |S_{21}| \square |S_{21A}|;$$

$$\Delta^+ \arg S_{12} = |E'_{TR}| + |S_{22A}| \cdot |E_{SR}| + |S_{11A}| \cdot |E_{LF}| + \frac{|E_{XR}|}{|S_{12A}|}, \quad \Delta^+ |S_{12}| \square |S_{12A}|.$$

Заметим, что при нарушении неравенств, приведенных справа от формул, эти формулы не справедливы, т. к. тогда неправомерно ограничиваться первым порядком по факторам ошибок и заменять $\arctg(x)$ на x .

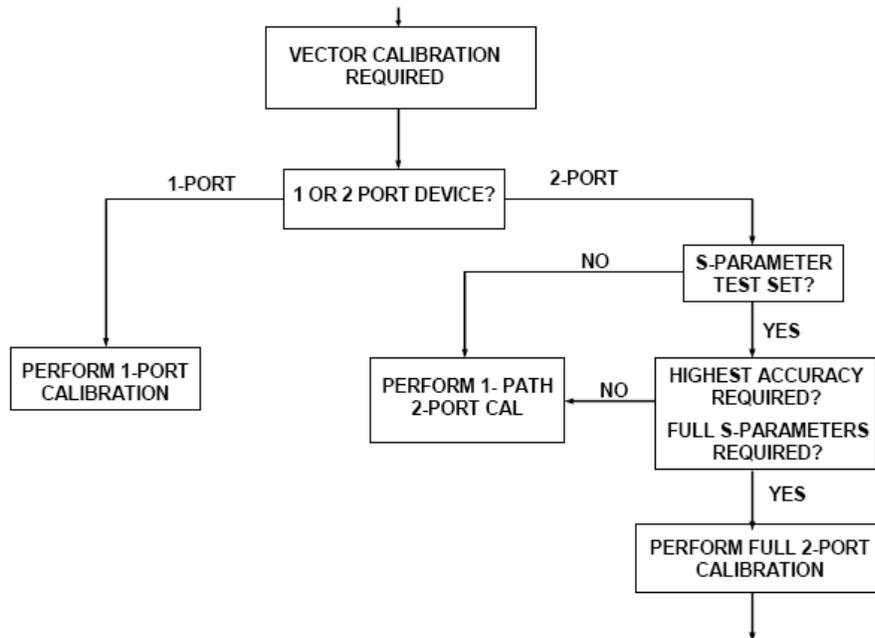
3.2. Первый порядок факторов ошибок, результаты в градусах

Для перевода результатов п. 3.1. в градусы, их нужно помножить на коэффициент $180/\pi$.

Измерительная калибровка

Компьютерный калибровочный диалог (илл. Осн анал цепей, с. 123)

Choosing the Proper Calibration (cont.)

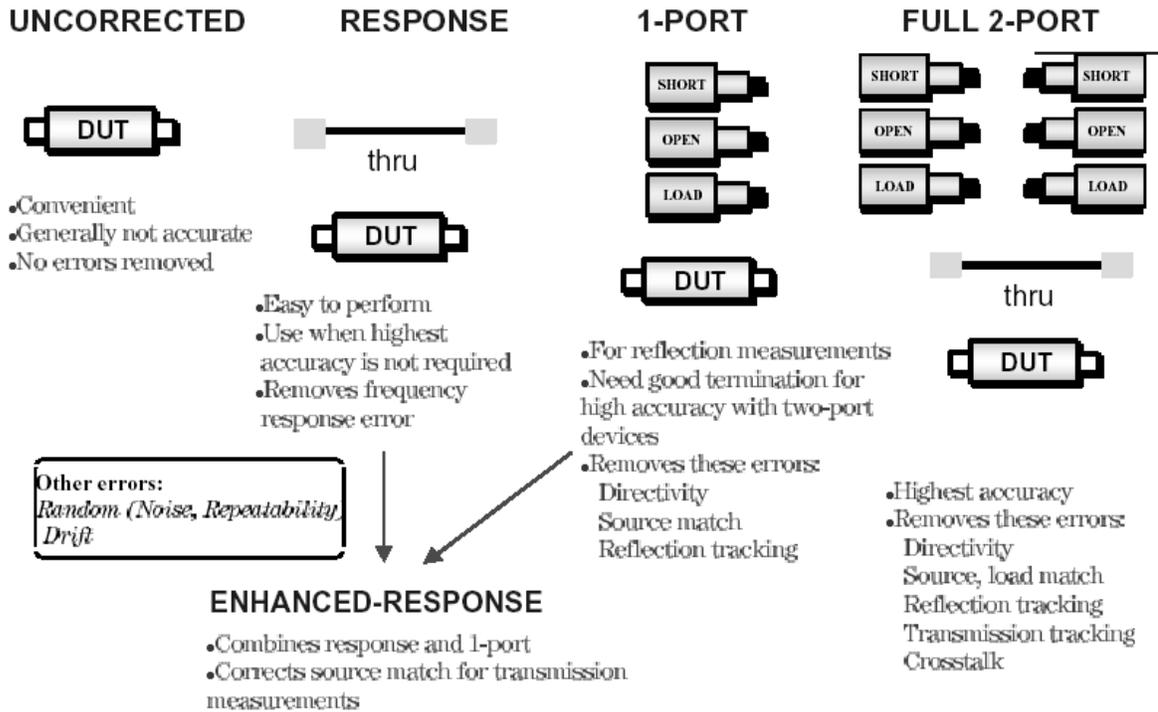


Калибровочные эталоны (SOLT), кит, файл описаний (илл. Осн анал цепей, с. 121)



The Cal Kit contains the known Standards

Виды калибровок (илл. Балло, с. 36)



Механическая и электронная калибровка

Коррекция данных

Уравнения коррекции

$$S_{11A} = \frac{\xi_{11} A_R - \xi_{21} \xi_{12} E_{LF}}{D}$$

$$S_{21A} = \frac{\xi_{21} (A_R - \xi_{22} E_{LF})}{D}$$

$$S_{12A} = \frac{\xi_{12}(A_F - \xi_{11}E_{LR})}{D} \quad S_{22A} = \frac{\xi_{22}A_F - \xi_{21}\xi_{12}E_{LR}}{D}$$

$$D = A_F A_R - \xi_{21}\xi_{12}E_{LF}E_{LR}$$

$$A_F = 1 + \xi_{11}E_{SF} \quad A_R = 1 + \xi_{22}E_{SR}$$

$$\xi_{11} = \frac{S_{11M} - E_{DF}}{E_{RF}} \quad \xi_{21} = \frac{S_{21M} - E_{XF}}{E_{TF}}$$

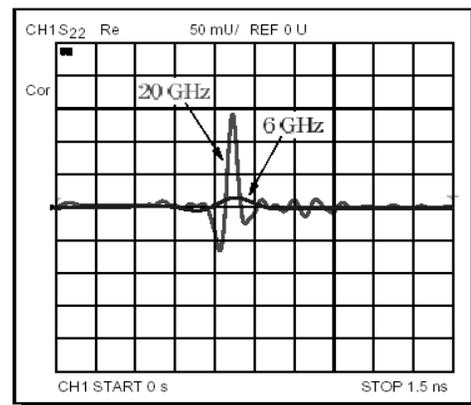
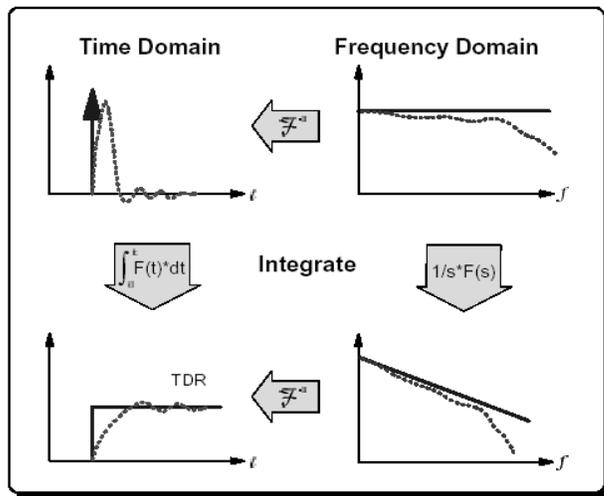
$$\xi_{12} = \frac{S_{12M} - E_{XR}}{E_{TR}} \quad \xi_{22} = \frac{S_{22M} - E_{DR}}{E_{RR}}$$

Верификация точностей (измерительные линии, специальные нагрузки)

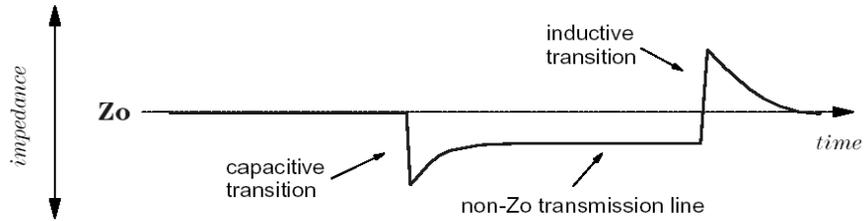
Маркеры

Рефлектометрия во временной области (PBO)

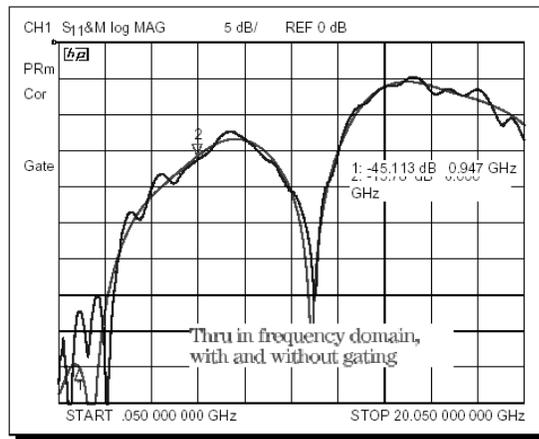
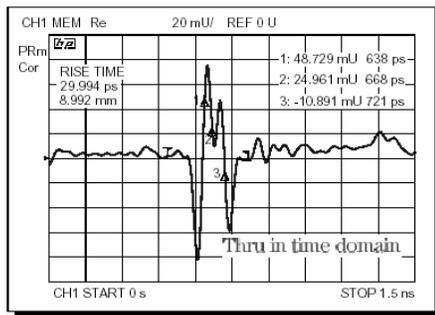
- start with broadband frequency sweep (often requires microwave VNA)
- use inverse-Fourier transform to compute time-domain
- resolution inversely proportionate to frequency span



- what is TDR?
 - time-domain reflectometry
 - analyze impedance versus time
 - distinguish between inductive and capacitive transitions
- with gating:
 - analyze transitions
 - analyzer standards

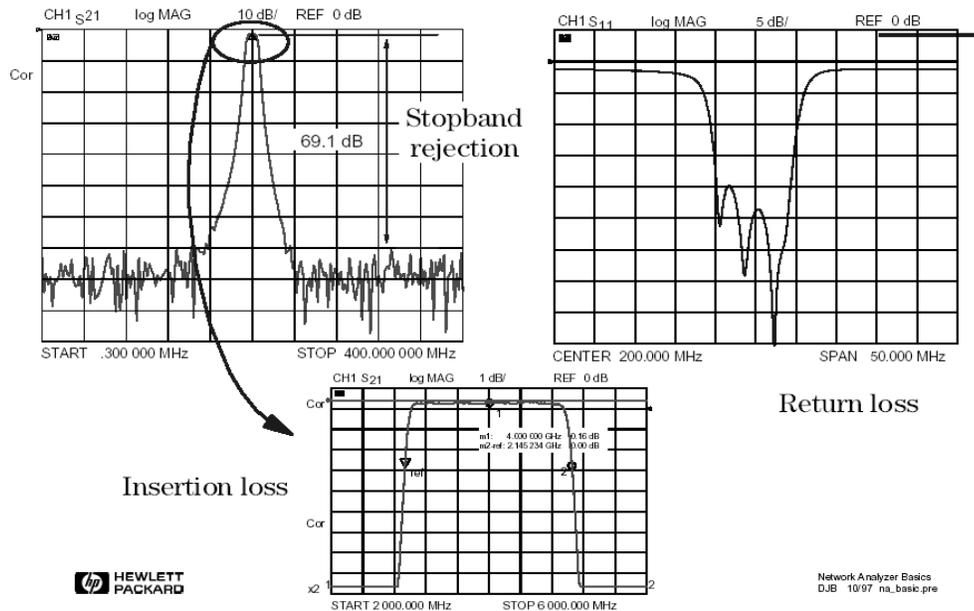


- TDR and gating can **remove** undesired reflections (a form of error **correction**)
- Only useful for **broadband** devices (a load or thru for example)
- Define **gate** to only include DUT
- Use two-port calibration



Типовые измерения

Тестирование полосового фильтра в частотной области (илл. Балло, с. 50)



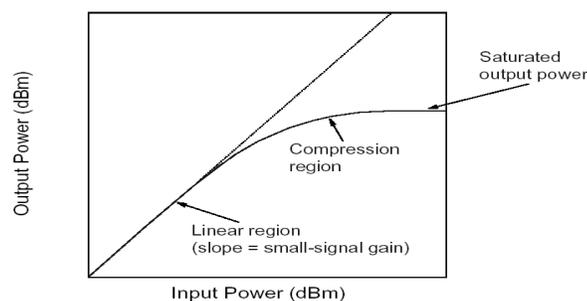
Выше показаны частотные отклики фильтра. Слева сверху отклик пропускания в логарифмическом масштабе амплитуд (внесенные потери), справа – отклик отражения (возвратные потери).

Наиболее общие характеристики измеряемого фильтра это внесенные потери и полоса, показаны на нижнем рисунке с растянутой вертикальной шкалой. Другой общий параметр, который мы можем измерить, есть внеполосная режекция. Это мера того, как хорошо фильтр пропускает сигналы внутри полосы, и в то же время вырезает все другие сигналы вне той же полосы. Возможность тестирующей системы измерять внеполосную режекцию прямо зависит от того, как задан ее динамический диапазон.

Кривая возвратных потерь очень типична, показывает высокое отражение (около 0 дБ) в полосе заграждения, и разумное согласование в полосе пропускания. Большинство пассивных фильтров работают таким образом. Существует специальный класс фильтров, поглощающих как в полосе пропускания, так и в полосе заграждения. Эти фильтры демонстрируют хорошее согласование в широком частотном диапазоне.

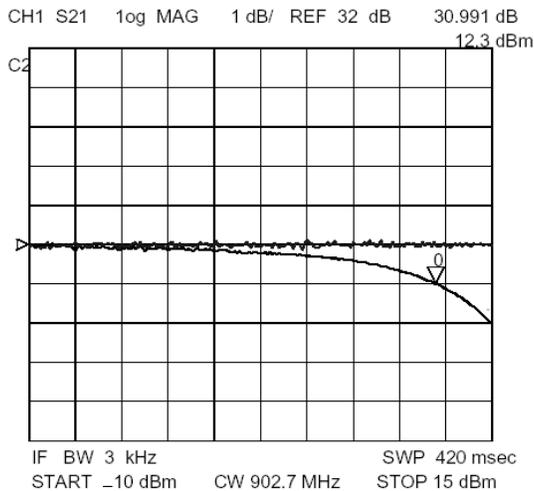
Для очень узкополосных приборов, таких как кристаллические фильтры, анализатор цепей должен качать частоту достаточно медленно, чтобы позволить фильтру соответственно реагировать. Если оставить скорость качания такой же большой, могут появиться большие ошибки. Это может случиться с приборами, которые электрически очень длинные. Большое время задержки прибора может быть результатом настройки приемника на более высокие частоты, чем те, что следуют с прибора, что также может обусловить значительные измерительные ошибки.

Тестирование усилителя с помощью мощностной панорамы (илл. Балло, с. 51, с. 51)



Многие анализаторы имеют возможность производить качание мощности, наряду с качанием частоты. Качание мощности помогает характеризовать нелинейные свойства усилителя. Выше показана кривая выходной мощности усилителя в функции входной мощности на одной частоте. Усиление усилителя на любом частном мощностном уровне есть наклон этой кривой. Заметим, что усилитель имеет линейную область действия, где усиление постоянно и не зависит от уровня мощности. Усиление в этой области обычно называют «усилением малого сигнала». Когда входная мощность увеличивается, усиление начинает уменьшаться, и говорят, что усилитель сжимается. В этих нелинейных условиях выход усилителя уже не синусоидальный – часть выходной мощности находится в гармониках, вместо того, чтобы полностью присутствовать только на основной частоте. Когда входная мощность увеличивается еще больше, усилитель становится насыщенным, и выходная мощность остается постоянной. В этой точке усиление, по существу, равно нулю, поскольку дальнейшее увеличение входной мощности не изменяет выходную мощность. Насыщенная выходная мощность может быть считана прямо из приведенной кривой.

Для измерения насыщенной выходной мощности усилителя анализатор должен быть способен обеспечить качание мощности с достаточной выходной мощностью, чтобы вести усилитель из линейной области в насыщение. Для достижения этого может понадобиться предусиление на входе тестируемого усилителя.

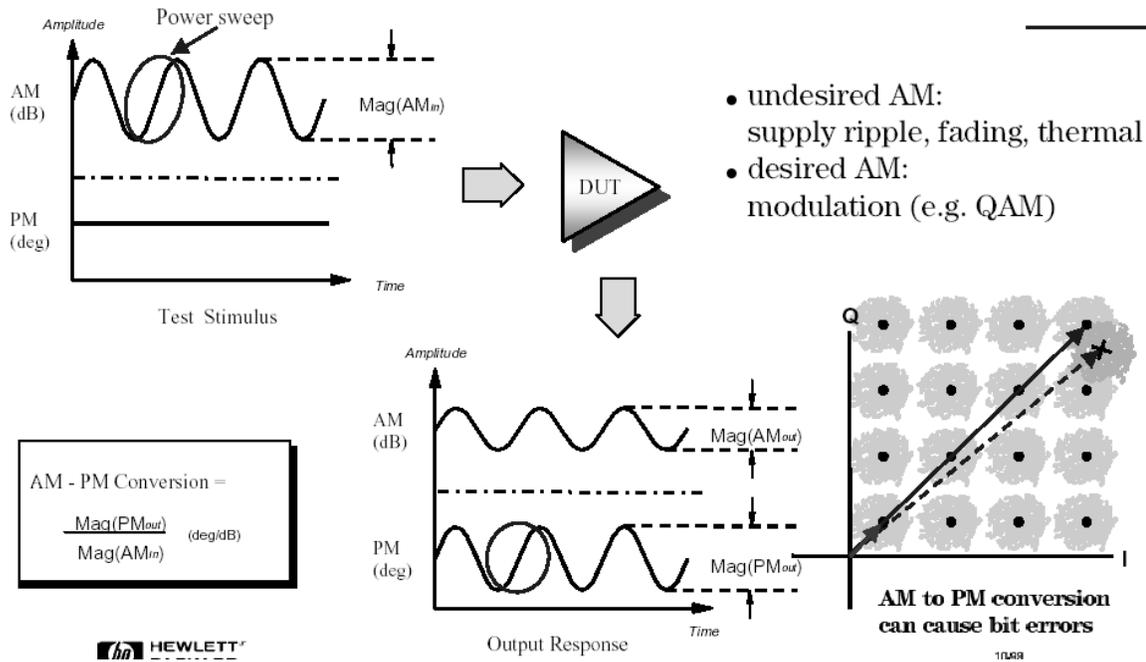


- **1 dB compression:** input power resulting in 1 dB *drop* in gain
- Ratioed measurement
- Output power available (non-ratioed measurement)

- 1 дБ сжатия: входная мощность приводит к уменьшению усиления на 1 дБ
- относительное измерение
- доступность выходной мощности (неотнoсительное измерение)

Тестирование амплитудно-фазовой конверсии с помощью мощностной панорамы (илл. Балло, с. 52, с. 53)

Measure of phase deviation caused by amplitude variations

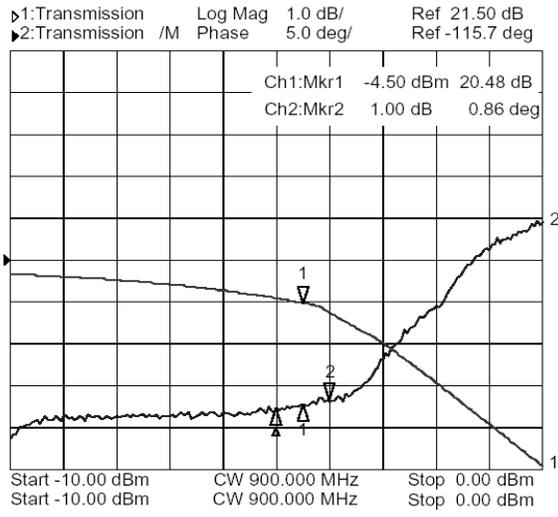


Другое общее измерение, которое помогает охарактеризовать нелинейное поведение усилителей, это измерение амплитудно-фазовой конверсии, которое измеряет величину нежелательной фазовой девиации (ФМ), индуцируемой амплитудными вариациями (АМ), присущими системе. В системах связи эта нежелательная фазовая модуляция обусловлена непреднамеренными амплитудными вариациями, подобными дрожанию мощности питания, температурному дрейфу или многолучевому федингу, или преднамеренному амплитудному изменению, который происходит в результате использованного типа модуляции, подобного QAM или вспышечной модуляции.

Амплитудно-фазовая конверсия есть, в частности, критический параметр в системах, где эксплуатируется фазовая (угловая) модуляция, поскольку нежелательные фазовые искажения служат причиной деградации аналоговых сигналов, или увеличивают частоту битовых ошибок (BER) в цифровых системах. Примеры общих модуляционных типов, использующих фазовую модуляцию, это ЧМ, QPSK, 16QAM. Хотя легко измерять BER цифровых систем связи, одно это измерение не обеспечивает проникновение в лежащие под ним явления, которые обуславливают битовые ошибки. Амплитудно-фазовая конверсия есть один из основных вкладов в BER, поэтому важно количественно определить этот параметр в системах связи.

Диаграмма I/Q, приведенная выше, показывает как амплитудно-фазовая конверсия может привести к битовым ошибкам. Пусть желаемое состояние изменилось от маленького сплошного вектора к большому сплошному вектору. При амплитудно-фазовой конверсии настоящий большой вектор может быть таким, как показано точечной линией. Это произошло благодаря фазовому сдвигу, который образовался из изменения уровня мощности. Для сигнала с модуляцией 64QAM, как показано (нарисован только один квадрант), мы видим, что шумовые кружки, которые окружают каждое действительное состояние, перекрываются, что означает, что статистически присутствует некоторые битовые ошибки.

Амплитудно-фазовая конверсия обычно определяется как изменение выходной фазы на 1 дБ добавки выходной мощности усилителя, выраженное в градусах на дБ (град/дБ). Идеальный усилитель должен не иметь взаимодействия между его фазовым откликом и уровнем входного сигнала.



- Use transmission setup with a power sweep
- Display phase of S21
- AM - PM = 0.86 deg/dB

АНАЛИЗАТОРЫ СПЕКТРА

Предисловие

Цепи и сигналы – два аспекта передачи и преобразования информации в тракте любого радиотехнического устройства, каждый из них не имеет смысла без другого. Поэтому два главных аппаратных исследования, связанных с этими передачей и преобразованием – анализ цепей и анализ сигналов. Число параметров и характеристик сигналов, интересных для практики и теории, бесконечно. Очевидно, первое (но не единственное!), с чего нужно начинать анализ сигналов, это спектральный анализ – при характеристике в частотной области, и форма сигнала – при характеристике во временной области.

Фактически спектральный анализ электромагнитных колебаний восходит к 17 веку, когда был открыт *спектральный прибор* – призма, реагирующий на частоту света, и начало изучаться и использоваться спектральное разложение света призмой. Еще раньше производились генерация и анализ спектральных составляющих звука на базе таких спектральных приборов, как струна, мембрана, воздушный резонатор. Однако переход к осознанному количественному спектральному анализу стал возможен только с работ Ж.-Б.Ж. Фурье, фактически открывшего спектрально-временной дуализм прикладных функций. С появлением радиотехники сразу возникла необходимость анализа сигналов, в частности, спектрального анализа. В качестве спектрального прибора использовался резонатор, в качестве вспомогательных явлений – интерференция и преломление в диспергирующей среде. Из двух принципиальных разновидностей аппаратного спектрального анализа – параллельного и последовательного действия – в универсальных спектроанализаторах выжила вторая, в то время как для специализированных часто используется первая.

Главной особенностью задачи спектрального анализа сигналов СВЧ является чрезвычайное разнообразие структур сигналов, используемых в различных радиоэлектронных приложениях, и разнообразие значений их параметров. Это чрезвычайно усложняет структуру, архитектуру и функциональную вариативность спектроанализатора, имеющего претензии на универсальность. Такие параметры прибора, как ширина обзора частотной области, шаг частотной дискретизации, частотное разрешение, степень видеосреднения, степень сглаживания измеряемых характеристик и др. должны иметь возможность варьировать при установке в очень широких пределах.

К настоящему времени спектроанализаторы СВЧ стали важнейшим измерительным и индицирующим средством характеристики высокочастотных и сверхвысокочастотных сигналов самой разнообразной структуры. Современный универсальный спектроанализатор ВЧ и СВЧ относится к четвертому поколению радиоизмерительных приборов; позволяет вести автоматическую регистрацию амплитудных (скалярный анализатор) и комплексных (векторный анализатор) спектров панорамно по частоте, вплоть до рабочей полосы 0–110 ГГц; в большой мере управляется и вычислительно обслуживается внутренним или внешним компьютером; использует как минимум тройное преобразование частоты, гетеродины которого построены на высокостабильных синтезаторах частоты; широко использует цифровые технологии, в частности, имеет полностью цифровой тракт, начиная с выхода последнего смесителя; автоматически представляет и сохраняет спектральные характеристики сигналов в различных форматах, масштабах и подробностях; может быть включен в систему метрологического трассирования вплоть до национальных стандартов.

Тем не менее, возможности увеличения эффективности векторных и скалярных спектроанализаторов далеко не исчерпаны. Косвенным свидетельством этому является ежегодное появление на рынке измерительных приборов СВЧ все новых, более совершенных моделей анализаторов ведущих мировых фирм. Можно предположить, что в связке с технологическими прорывами радиотехнического значения и рыночной конъюнктурой основными направлениями развития будут:

- 1) дальнейшая виртуализация системы, то есть передача выполняемых функций от радиотехнических устройств к компьютерам;

- 2) совершенствование системы измерительных калибровок;
- 3) дальнейшая автоматизация цикла измерений, вплоть до почти полного устранения ручных операций;
- 4) совершенствование системы интерпретации наблюдаемых данных на базе специально разработанной теории интерпретации спектрометрических данных;
- 5) дистанционное управление работой прибора;
- 6) расширение возможностей включения прибора в измерительный комплекс без доработки интерфейсов и т.п.

Теоретические предпосылки

Цель аппаратного спектрального анализа

Будем пока для простоты считать сигнал действительной конечной непрерывной функцией времени. Фактически при изложении на названном уровне обычно предполагается (явно или неявно), что сигнал детерминированный (в смысле – *не случайный*), то есть найденная аппаратно спектральная характеристика интерпретируется как характеристика данного сигнала, а не оценка характеристики некоторой совокупности реализаций (ансамбля) случайного процесса, полученная по одной реализации. Впрочем, предположение о детерминированности сигнала не сужает дальнейших технических возможностей, т.к. названная оценка легко получается из спектральной характеристики детерминированного сигнала.

Итак, в свете сделанных оговорок будем считать, что целью *скалярного аппаратного спектрального анализа* является получение *амплитудного спектра* (*амплитудной спектральной плотности*) $|S(\omega)|$ сигнала $f(t)$, а целью *векторного аппаратного спектрального анализа* является получение *комплексного спектра* (*комплексной спектральной плотности*) $S(\omega)$. Здесь комплексный спектр определен как *преобразование Фурье* (*интеграл Фурье*) от сигнала:

$$S(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt,$$

а амплитудный спектр – как модуль комплексного спектра.

В свою очередь, при определенных условиях (всегда выполняющихся при сделанных предположениях) сигнал как временная функция может быть получен из комплексного спектра (но не амплитудного!) *обратным преобразованием Фурье*:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{i\omega t} d\omega.$$

Говорят, что в виде $f(t)$ сигнал представлен *во временной области*, а в виде $S(\omega)$ – *в частотной области*. Амплитудный спектр суть неполное представление сигнала в частотной области, именно поэтому, зная только его, невозможно восстановить $f(t)$. Это обстоятельство – один из главных стимулов для создания гораздо более сложных векторных спектроанализаторов (иногда – на опциональной основе).

Используя экспоненциальную форму комплексного числа $z = |z| e^{i\varphi}$, запишем комплексный спектр в виде

$$S(\omega) = |S(\omega)| e^{i\varphi_s(\omega)}.$$

Как видим, комплексный спектр характеризуется двумя функциями: амплитудным спектром $|S(\omega)|$ и *фазовым спектром* $\varphi_s(\omega)$. Сделаем два замечания.

1) Как и при любом аппаратном измерении, функции $|S(\omega)|$ и $\varphi_s(\omega)$ находятся с ошибками (случайными и систематическими), меры которых сами являются функциями частоты ω ; то есть на самом деле мы получаем *статистические оценки* $|\hat{S}(\omega)|$, $\hat{\varphi}_s(\omega)$ искомым функций.

Частотные функции мер ошибок зависят как от специфических свойств спектральных оценок [2.1], так и структуры спектроанализатора и его обработки сигналов. Если первым зависимостям посвящены многочисленные работы (в основном, на общетеоретическом уровне), то исследование вторых представляет обширное поле для деятельности, частично освоенное на частном, утилитарном уровне.

2) Если представление сигнала во временной области получается обратным преобразованием Фурье из $\hat{S}(\omega)$ (а не непосредственной его записью), то и это представление есть оценка $\hat{f}(t)$ временной функции; временные функции мер ошибок этой оценки связаны непростыми зависимостями с частотными функциями мер ошибок спектральных оценок (плюс ошибки численного расчета обратного преобразования Фурье).

Параллельный и последовательный анализ

Для целей аппаратного спектрального анализа необходим *спектральный прибор*, реагирующий на частоту колебаний. Поиск рабочих физических эффектов для создания спектральных приборов шел с 17 века, когда было открыто и начало использоваться спектральное разложение света призмой. Тем не менее, и в середине двадцатого века можно было всего лишь констатировать, что «в основе действия спектральных приборов лежит одно из следующих явлений: интерференция, преломление при наличии дисперсии, резонанс» [2.2]. В радиотехнике, в том числе на СВЧ давно победили методы спектрального анализа с помощью резонанса, то есть с использованием резонаторов.

Существует две принципиальные возможности аппаратного спектрального анализа с помощью резонаторов. Первый способ состоит в применении набора резонаторов, настроенных на различные частоты и подвергающихся одновременно воздействию исследуемого колебания; этот способ называется *одновременным (параллельным) анализом*. Второй способ предполагает применение одного резонатора с переменной настройкой (модификацию этого способа мы рассмотрим ниже); этот способ называется *последовательным анализом*.

На первый взгляд, параллельный анализ имеет ряд преимуществ перед последовательным. Во-первых, кажется очевидным, что параллельный анализ более быстрый, т.к. скорость перестройки резонатора в последовательном анализе ограничена искажениями, создаваемыми переходными процессами. Во-вторых, последовательный анализ кажется малоприспособленным при быстро протекающих, быстро изменяющихся, нестационарных процессах. В-третьих, кажутся трудными и малоэффективными процессы калибровки считывания настройки резонатора в каждый момент времени.

Однако, логика технического прогресса в данной области привела к почти полной победе последовательного анализа, по крайней мере, в универсальных спектроанализаторах (правда, после некоторой модификации), оставившего параллельному анализу только исключительные, редкие ситуации: спонтанные, непрогнозируемые, неповторяющиеся СВЧ-«вспышки» и т.п. Решающими факторами этой «победы» оказались гибкость, автоматизируемость и точность анализа.

Упомянутая модификация последовательного анализа состоит в том, что вместо плавной управляемой перестройки резонатора используется управляемый сдвиг спектра самого сигнала во времени. Это стало возможным благодаря применению *супергетеродина*, в результате которого селективные функции отделены от функций перестройки и выполняются на постоянной *промежуточной* частоте, что обеспечивает возможность прецизионного регулирования параметров частотной селекции (полосы, крутизны среза и т.п.). В технике спектрального анализа супер-

гетеродинный принцип сыграл столь же революционную роль, как и вообще в технике перестраиваемого приема сигналов.

На самом деле, техническое совершенствование как параллельного, так и последовательно-анализа не является стратегической целью построения совершенного спектроанализатора будущего, именно потому, что эти методы опосредованы свойствами резонансной системы. Чтобы понять, к чему следует стремиться в аппаратном анализе сигнала, как ни парадоксально, необходимо отказаться от первого же утверждения, с которого мы начали данное изложение: что «... необходим спектральный прибор ...». Действительно, как будет показано ниже, спектральный прибор, будь то резонатор или другая система, принципиально вносит свои искажения в получаемый продукт, отличающийся от желаемого спектра сигнала; более того, эти искажения бывает трудно учесть.

Стратегическая цель развития аппаратного анализа спектра должна основываться на непосредственной записи сигнала и выполнении численного преобразования Фурье от него (здесь мы для простоты обсуждения предполагаем детерминированный сигнал; спектральный анализ случайного сигнала будет рассмотрен позже). На языке сегодняшней цифровой техники сказанное означает оцифровку сигнала уже на сверхвысоких частотах и выполнение *быстрого преобразования Фурье* (БПФ) от него. Достижение этой цели сдерживается лишь недостаточной скоростью АЦП. Например, приняв достаточным период отсчетов в 6 градусов фазы, потребуем 60 отсчетов на периоде синусоиды, что на частоте f_s сигнала требует скорости АЦП $60 f_s$ отсчетов в секунду (о/с). При частоте $f_s = 1$ ГГц это означает $6 \cdot 10^{10}$ о/с, при частоте 10 ГГц – $6 \cdot 10^{11}$ о/с и т.д. Ясно, что применительно к интересующему нас диапазону СВЧ такие скорости лежат далеко за пределами возможностей современных АЦП. Этот вывод не изменится, если мы ограничимся всего лишь 20 отсчетами на периоде синусоиды. Более того, достижение скоростей в $10^{11} - 10^{12}$ о/с и более, по-видимому, потребует изменений структурных элементов АЦП на технологическом уровне. В настоящее время оцифровка сигнала (с последующей цифровой обработкой) производится лишь в тракте последней промежуточной частоты, так что предыдущие аналоговые элементы фильтрации вносят свою лепту в искажения получаемой спектральной характеристики.

Текущий спектр

Пусть бесконечный по длительности сигнал воспринимается спектроанализатором по мере поступления и обрабатывается в реальном или почти реальном времени. В этих условиях вычисление прямого преобразования Фурье в бесконечных пределах невозможно по двум причинам: во-первых, сигнал $f(t)$ поступает на систему обработки лишь с некоторого конечного момента времени t_0 ; во-вторых, он известен (поступал) лишь до текущего момента t . Приняв момент t_0 за начало отсчета времени, мы констатируем, что в качестве спектральной характеристики сигнала вынуждены взять функцию частоты, задаваемую усеченным интегралом Фурье:

$$S_t(\omega; t) = \int_0^t f(\tau) e^{-i\omega\tau} d\tau.$$

Функция частоты $S_t(\omega; t)$, параметрически зависящая от времени t , называется *текущим спектром* сигнала $f(t)$.

Введение этого понятия учитывает один из важнейших аспектов реального измерения и уточняет цель аппаратного спектрального анализа сигнала: этой целью не может быть получение «истинного» спектра $S(\omega)$; максимум, на что можно рассчитывать, это текущий спектр $S_t(\omega; t)$. В связи с этим возникает ряд вопросов: при каких условиях и в каком смысле текущий спектр $S_t(\omega; t)$ сходится при $t \rightarrow \infty$ к «истинному» спектру $S(\omega)$? Какова скорость этой сходимости? Какова оценка ошибки текущего спектра $|S_t(\omega; t) - S(\omega)|$ как функции времени? Рассмотрение этих вопросов увело бы нас далеко в сторону от цели данной главы. Рассмотрим лишь простой пример.

Пусть сигнал имеет вид $f(t) = \cos \Omega t$, $-\infty < t < \infty$. «Истинный» спектр (см. 3.10):

$$\begin{aligned} S(\omega) &= \int_{-\infty}^{\infty} \cos \Omega t e^{-i\omega t} dt = \int_{-\infty}^{\infty} \cos \Omega t \cdot \cos \omega t \cdot dt = \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \{ \cos[(\omega - \Omega)t] + \cos[(\omega + \Omega)t] \} dt = \pi [\delta(\omega - \Omega) + \delta(\omega + \Omega)]. \end{aligned}$$

Для удобства перехода к пределу и избавления от мнимой части, текущий спектр вычислим, поместив начало отсчета времени в середине интервала наблюдения сигнала:

$$S_i(\omega, t) = \int_{-t/2}^{t/2} \cos \Omega \tau \cdot e^{-i\omega \tau} d\tau = \int_{-t/2}^{t/2} \cos \Omega \tau \cdot \cos \omega \tau d\tau = \frac{\sin [(\omega - \Omega)t/2]}{\omega - \Omega} + \frac{\sin [(\omega + \Omega)t/2]}{\omega + \Omega}.$$

Нетрудно найти, что при $t \rightarrow \infty$ $S_i(\omega, t) \rightarrow \pi [\delta(\omega - \Omega) + \delta(\omega + \Omega)] = S(\omega)$.

До сих пор предполагался общий случай существования сигнала на всей оси времени (возможно, с убыванием на $\pm\infty$). В частном случае сигнал *финитен*, то есть отличен от нуля на некотором конечном интервале времени $[t_1, t_2]$. Тогда, если спектроанализатор был готов принимать и обрабатывать сигнал до момента t_1 и делал это до момента $t \geq t_2$, текущий спектр $S_i(\omega; t)$ к моменту t_2 совпал с «истинным» спектром $S(\omega)$, а дальнейшее измерение его не изменило. Если сигнал финитен, но очень длительный, для выполнения этих условий потребуется отказаться от требования обработки в реальном или почти реальном времени.

Мгновенный спектр

Усеченный мгновенный спектр

Главным недостатком спектральной характеристики сигнала с помощью преобразования Фурье является отсутствие локализации по времени; предоставляемая при этом, в соответствии с соотношением неопределенности «время-частота», бесконечная локализация по частоте не используется по техническим причинам. Отсутствие локализации по времени не позволяет контролировать какие-либо временные изменения спектральной характеристики. Требуются такие спектральные характеристики, которые бы совмещали умеренные локализации по времени и по частоте, тем самым давая промежуточное представление сигнала между временной и частотной областями. Радикальное решение этой задачи – представление сигнала в вейвлетном базисе. Однако и в Фурье-базисе имеются некоторые возможности. Один из вариантов – введение спектральной характеристики, называемой *мгновенным спектром*.

Частная конструкция мгновенного спектра – функция частоты и времени, определяемая как [2.2]:

$$S_T(\omega, t) = \int_{t-T}^t f(\tau) e^{-i\omega \tau} d\tau,$$

то есть это спектр отрезка сигнала длительностью T , непосредственно предшествующего моменту наблюдения t . Назовем эту конструкцию *усеченным мгновенным спектром*. По мере течения времени мгновенный спектр получается скользящим интегрированием по перемещающемуся отрезку постоянной длины.

Чтобы лучше понять смысл новой характеристики, выразим ее через текущий спектр:

$$S_T(\omega, t) = \int_{-\infty}^t f(\tau) e^{-i\omega \tau} d\tau - \int_{-\infty}^{t-T} f(\tau) e^{-i\omega \tau} d\tau = S_i(\omega, t) - S_i(\omega, t-T),$$

то есть на каждой частоте ω мгновенный спектр равен приращению текущего спектра на интервале времени длиной T , предшествующем моменту наблюдения t . Предполагая, что функция текущего спектра $S_i(\omega, t)$ дифференцируема по аргументу t , при малых T можно это приращение выразить так:

$$S_T(\omega, t) \approx \frac{\partial S_i(\omega, t)}{\partial t} \cdot T.$$

Эту же приближенную формулу можно было получить сразу дифференцированием интеграла по параметру t в определении мгновенного спектра.

Кроме пояснения смысла мгновенного спектра, формула дает полезный рецепт для приближенной оценки верхней грани скорости изменения текущего спектра нестационарного сигнала на каждой частоте ω . Действительно, с одной стороны, из этого равенства следует:

$$\frac{\partial S_i}{\partial t} \approx \frac{S_T}{T};$$

с другой стороны, записав комплексный текущий спектр в виде

$$S_i(\omega, t) = |S_i(\omega, t)| e^{i\varphi_i(\omega, t)}$$

и дифференцируя по t , получаем:

$$\frac{\partial S_i(\omega, t)}{\partial t} = e^{i\varphi_i(\omega, t)} \left[\frac{\partial |S_i(\omega, t)|}{\partial t} + i |S_i(\omega, t)| \frac{\partial \varphi_i(\omega, t)}{\partial t} \right].$$

Приравнявая правые части равенств и переходя к модулям, имеем:

$$\frac{|S_T(\omega, t)|}{T} \approx \sqrt{\left(\frac{\partial |S_i(\omega, t)|}{\partial t} \right)^2 + |S_i(\omega, t)|^2 \left(\frac{\partial \varphi_i(\omega, t)}{\partial t} \right)^2} \geq \left| \frac{\partial |S_i(\omega, t)|}{\partial t} \right|,$$

и мы получили искомое неравенство:

$$\left| \frac{\partial |S_i(\omega, t)|}{\partial t} \right| \leq \frac{|S_T(\omega, t)|}{T}.$$

Взвешенный мгновенный спектр

Вернемся к определению мгновенного спектра. Данное выше определение было названо частной конструкцией. Более общее определение мгновенного спектра [2.2, 2.3] связано с конструкцией, в которой в интеграл Фурье вводится скользящая (связанная с текущим временем) весовая функция:

$$S_r(\omega, t) = \int_{-\infty}^{\infty} r(\tau - t) f(\tau) e^{-i\omega\tau} d\tau;$$

получилась свертка функций $r(t)$ и $f(t)e^{-i\omega t}$. Такая конструкция имеет два достоинства: она удобна как в аналитических преобразованиях, так и в учете искажающего спектральный анализ действия реальных цепей. Назовем такую конструкцию *взвешенным мгновенным спектром*.

То что $S_r(\omega, t)$ – более общая конструкция мгновенного спектра, чем $S_T(\omega, t)$, видно из того, что последняя есть частный случай первой при $r(t) = H(t+T) - H(t)$, где

$$H(t) = \begin{cases} 0, & \text{при } x < 0, \\ 1, & \text{при } x \geq 0, \end{cases}$$

– функция Хевисайда (единичный скачок).

Важный частный пример весовой функции [2.3]:

$$r(t) = e^{\alpha t} H(-t), \quad -\infty < t < \infty.$$

Эта функция учитывает прошлое процесса с весом, экспоненциально убывающим по мере удаления от момента наблюдения. Заметим, что $H(-t)$ – обращенный единичный скачок:

$$H(-t) = \begin{cases} 1, & \text{при } t \leq 0. \\ 0, & \text{при } t > 0. \end{cases}$$

Такая весовая функция отражает влияние на спектральный анализ реальных цепей с однорелаксационным затуханием, например, резонансных фильтров первого порядка; параметр α учитывает скорость затухания.

Плотности мощности

Было бы привлекательно построить такую спектральную характеристику, которая служила бы временной плотностью некоей интегральной характеристики, от которой она получалась бы дифференцированием по времени. Осуществлению этой идеи мешает комплексный характер «истинного» спектра и текущего спектра. Наиболее просто и удобно реализовать эту идею, взяв за основу не текущий спектр, и не амплитудный текущий спектр, а спектр мощности сигнала. Как показано в главе 6, спектральной характеристикой случайного процесса является *спектральная плотность мощности (энергетический спектр)*, по определению являющийся средней мощностью на единицу полосы процесса. Аппаратную оценку этой характеристики дает величина, пропорциональная квадрату амплитудного спектра, полученного достаточно долгим усреднением по реализации процесса, то есть квадрату модуля текущего спектра. Замечательно, что эта оценка сохраняет свой смысл спектральной плотности мощности (но уже без добавления слова «средней») и для детерминированного сигнала. Удобно, что эта функция частоты не только действительна, но и всюду неотрицательна, а также связана простым интегральным соотношением с мощностью сигнала.

Итак, пусть по-прежнему $S_i(\omega, t)$ – текущий (комплексный) спектр сигнала, $|S_i(\omega, t)|$ – текущий амплитудный спектр сигнала, тогда $F_i(\omega, t) = K |S_i(\omega, t)|^2$ – можно назвать *текущим энергетическим спектром* (K – коэффициент с размерностью сек⁻¹). Смысл $F_i(\omega, t)$ раскрывается равенством:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} F_i(\omega, t) d\omega = f^2(t) = P_i(t)$$

(в правой части – мгновенная мощность сигнала). Оно означает, что $F_i(\omega, t)$ это спектральная плотность мгновенной мощности – мощность на единицу (циклической) частоты.

Полезно и поучительно контролировать размерности рассматриваемых величин. Пусть сигнал и все связанные с ним величины приведены к сопротивлению в 1 Ом, в такой системе мощность равна квадрату напряжения. Тогда размерности: $[f(t)] = \text{В}$; $[S_i] = \text{В/Гц}$; $[F_i] = \text{В}^2/\text{Гц} = \text{Вт/Гц}$.

Следуя Пейджу [2.4], введем такую всюду неотрицательную функцию частоты ω и времени t $\rho(\omega, t)$ (с размерностью Вт), что

$$\int_{-\infty}^t \rho(\omega, \tau) d\tau = F_i(\omega, t).$$

Дифференцируя это равенство по t , получаем:

$$\rho(\omega, t) = \frac{\partial}{\partial t} F_i(\omega, t),$$

а интегрируя его по ω , имеем

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^t \rho(\omega, \tau) d\omega d\tau = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_i(\omega, t) d\omega = P_i(t),$$

откуда следует, что $\rho(\omega, t)$ суть *спектрально-временная плотность мощности* сигнала.

Можно пойти дальше в образовании плотностей мощности. Пусть сигнал снимается с апертуры W . Тогда можно ввести такую всюду неотрицательную функцию частоты ω , точки на апертуре \mathbf{r} и времени t $\phi(\omega, \mathbf{r}, t)$, что

$$\iint_W \phi(\omega, \mathbf{r}, t) d\mathbf{r} = \rho(\omega, t).$$

Эту функцию с размерностью $\frac{\text{Вт}}{\text{с} \cdot \text{Гц} \cdot \text{м}^2} = \frac{\text{Вт}}{\text{м}^2}$ назовем *спектрально-пространственно-временной плотностью мощности* сигнала.

Полезны также *пространственно-временная плотность* мощности $\chi(\mathbf{r}, t)$:

$$\chi(\mathbf{r}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(\omega, \mathbf{r}, t) d\omega$$

с размерностью $\frac{\text{Вт}}{\text{с} \cdot \text{м}^2}$ и *спектрально-пространственная плотность мощности* $\mu(\omega, \mathbf{r})$:

$$\mu(\omega, \mathbf{r}) = \int_{-\infty}^{\infty} \phi(\omega, \mathbf{r}, \tau) d\tau$$

с размерностью $\frac{\text{Вт}}{\text{Гц} \cdot \text{м}^2}$.

Какой спектр измеряет конкретный спектроанализатор?

Теперь, после рассмотрения различных спектральных характеристик, возникает вопрос: какую из них измеряет спектроанализатор? Конечно, точный и подробный ответ на этот вопрос возможен только после определения структуры СА, но принципиальную часть ответа возможно сформулировать уже сейчас.

Забегая вперед, будем иметь в виду СА последовательного типа (см. 2.1.2), с переносом спектра с помощью супергетеродина, гетеродин которого (первый, если их больше одного) суть синтезатор частоты (СЧ), последовательно генерирующий сетку частот с устанавливаемыми оператором или по умолчанию диапазоном и шагом перестройки.

Обозначим:

f_{\min}, f_{\max} – минимальная и максимальная частоты перестройки СЧ;

Δf – шаг перестройки СЧ;

$N = (f_{\max} - f_{\min}) / \Delta f$ – число частотных подынтервалов;

$f_k = f_{\min} + k \cdot \Delta f$, $k = 0, 1, 2, \dots, N$ – частотные точки;

$t_{\text{пер}} = t_{\text{тех}} + t_{\text{уст}}$ – время перестройки СЧ на следующую частотную точку;

$t_{\text{тех}}$ – технической время перестройки СЧ;

$t_{\text{уст}}$ – длительность переходных процессов при перестройке СЧ (время установления);

$t_{\text{ст}} \gg 1/\Delta f$ – время «стояния» СЧ на частотной точке;

$t_{\text{точ}} = t_{\text{ст}} + t_{\text{пер}}$ – время, затрачиваемое на одну частотную точку;

$t_k = t_0 + k \cdot t_{\text{точ}}$, $k = 0, 1, 2, \dots, N$ – момент времени, соответствующий частотной точке f_k ;

$t_{\text{разв}} = t_{\text{точ}} \cdot (N+1) + t_{\text{пер,раз}}$ – время, затрачиваемое на одну развертку частоты;

$t_{\text{пер,раз}}$ – время переключения на следующую развертку частоты.

СА с установленными предварительно параметрами f_{min} , f_{max} и Δf последовательно «становится» на частотные точки f_k , от первой $f_0 = f_{\text{min}}$ до последней $f_N = f_{\text{max}}$; в каждой из них СА находит оценку усеченного мгновенного спектра

$$S(\omega, t_k; t_{\text{ст}}) = |S(\omega, t_k; t_{\text{ст}})| e^{i\varphi(\omega, t_k; t_{\text{ст}})} = \int_{t_k}^{t_k+t_{\text{ст}}} f(\tau) e^{-i\omega\tau} d\tau, \quad 2\pi f_k \leq \omega < 2\pi f_{k+1}$$

участка спектра шириной Δf . Если это скалярный СА, то он выдает в качестве результата лишь оценку амплитудного спектра:

$$|S(\omega, t_k; t_{\text{ст}})| = \sqrt{A^2(\omega, t_k; t_{\text{ст}}) + B^2(\omega, t_k; t_{\text{ст}})},$$

где

$$A(\omega, t_k; t_{\text{ст}}) = \int_{t_k}^{t_k+t_{\text{ст}}} f(\tau) \cos(\omega\tau) d\tau; \quad B(\omega, t_k; t_{\text{ст}}) = \int_{t_k}^{t_k+t_{\text{ст}}} f(\tau) \sin(\omega\tau) d\tau.$$

Если же это векторный СА, то он выдает в качестве результата оценку комплексного спектра, скорее всего, в виде двух характеристик: амплитудного спектра $|S(\omega, t_k; t_{\text{ст}})|$ и фазового спектра

$$\varphi(\omega, t_k; t_{\text{ст}}) = \arctg [B(\omega, t_k; t_{\text{ст}})/A(\omega, t_k; t_{\text{ст}})].$$

Определим *скважность спектрального анализа* как отношение времени, затрачиваемого на одну развертку, к времени «стояния» на одной частотной точке:

$$Q_{\text{СА}} = \frac{t_{\text{разв}}}{t_{\text{ст}}} = (N+1) \left(1 + \frac{t_{\text{пер}}}{t_{\text{ст}}} \right) + \frac{t_{\text{пер,раз}}}{t_{\text{ст}}}.$$

Обычно $N \gg 1$ и составляет несколько десятков или сотен, поэтому последним слагаемым справа можно пренебречь; если второе слагаемое в скобках много меньше единицы, то $Q_{\text{СА}} \approx N+1$.

Итак, СА последовательного анализа измеряет мгновенный усеченный амплитудный спектр (или амплитудный и фазовый спектры) по графику, в котором:

- участку спектра шириной Δf соответствует время анализа $t_{\text{ст}} \gg 1/\Delta f$;
- каждый участок спектра шириной Δf анализируется на своем интервале времени;
- разные участки спектра разнесены по времени: минимально – на $t_{\text{точ}} \square t_{\text{ст}}$, максимально – на $Nt_{\text{точ}} \square Nt_{\text{ст}}$;
- анализ каждого участка спектра повторяется установленное оператором или по умолчанию число раз со скважностью $\square N+1$.

Искажающее действие реального фильтра

Спектральный прибор принципиально вносит свои искажения в оценку спектра сигнала. На этом этапе изложения нет смысла проводить глубокий анализ этих искажений; мы ограничимся

простой демонстрацией присутствия искажений. Пусть на вход фильтра ПЧ с импульсной характеристикой $h(t)$ поступает сигнал $f(t)$ на интервале $(t_k, t_k + t_{\text{CT}})$. Сигнал на выходе фильтра равен

$$x(t) = \int_{t_k}^{t_k + t_{\text{CT}}} f(\tau) h(t - \tau) d\tau.$$

Если $h(t) = \sin \omega_0 t$ (идеальный фильтр пропускания частоты ω_0), то

$$x(t) = \int_{t_k}^{t_k + t_{\text{CT}}} f(\tau) \sin[\omega_0(t - \tau)] d\tau = A(\omega_0, t_k; t_{\text{CT}}) \sin \omega_0 t - B(\omega_0, t_k; t_{\text{CT}}) \cos \omega_0 t,$$

и получается оценка истинного амплитудного спектра на частоте ω_0 :

$$|S(\omega_0, t_k; t_{\text{CT}})| = \sqrt{A^2(\omega_0, t_k; t_{\text{CT}}) + B^2(\omega_0, t_k; t_{\text{CT}})}.$$

Если же для реального полосового фильтра с потерями с центральной частотой ω_0 принять

$$h(t) = e^{-\alpha t} \sin \omega_0 t,$$

то получается

$$x(t) = \int_{t_k}^{t_k + t_{\text{CT}}} f(\tau) e^{\alpha(\tau - t_k)} \cos(\omega_0 \tau) d\tau = A_\alpha(\omega_0, t_k; t_{\text{CT}}) \sin \omega_0 t - B_\alpha(\omega_0, t_k; t_{\text{CT}}) \cos \omega_0 t,$$

где

$$A_\alpha(\omega_0, t_k; t_{\text{CT}}) = \int_{t_k}^{t_k + t_{\text{CT}}} f(\tau) e^{\alpha(\tau - t_k)} \cos(\omega_0 \tau) d\tau; \quad B_\alpha(\omega_0, t_k; t_{\text{CT}}) = \int_{t_k}^{t_k + t_{\text{CT}}} f(\tau) e^{\alpha(\tau - t_k)} \sin(\omega_0 \tau) d\tau.$$

Как видим, получается мгновенный спектр не функции $f(t)$, а взвешенной функции $f_\alpha(t) = e^{\alpha(t - t_k)} f(t)$. В более общем случае реального полосового фильтра с центральной частотой ω_0 [2.2] импульсная характеристика имеет вид:

$$h(t) = p(t) \sin \omega_0 t + q(t) \cos \omega_0 t = \text{Re} \left[s(t) e^{i \omega_0 t} \right],$$

где $s(t) = p(t) - iq(t)$, и на выходе фильтра:

$$\begin{aligned} x(t) &= \int_{t_k}^{t_k + t_{\text{CT}}} f(\tau) h(t_k - \tau) d\tau = \int_{t_k}^{t_k + t_{\text{CT}}} f(\tau) \text{Re} \left[s(t_k - \tau) e^{i \omega_0 (t_k - \tau)} \right] d\tau = \\ &= \text{Re} \left[e^{i \omega_0 t_k} \int_{t_k}^{t_k + t_{\text{CT}}} f(\tau) s(t_k - \tau) e^{-i \omega_0 \tau} d\tau \right] \end{aligned}$$

то есть получается мгновенный спектр взвешенной функции $f_s(t) = s(t_k - t) f(t)$.

Разрешающая способность последовательного анализа

На протяжении веков применения спектрального анализа, сначала в оптике, а затем и в радиотехнике, центральным вопросом качества анализа был вопрос разрешающей способности анализа. Выражаясь современным языком, можно сказать, что «... разрешающая способность – важнейшая метрологическая характеристика (спектрального) анализатора» [2.2]. На этом этапе изложения мы лишь наметим общие контуры проблемы. Классический подход к этому вопросу состоит в следующем. В самом общем смысле под разрешающей способностью спектроанализатора понимается его способность *разрешить* (то есть «видеть» по отдельности) две соседние

спектральные линии в сигнале. В общем случае эти линии могут представлять собой локальные максимумы спектра с достаточно глубокими провалами вокруг этих максимумов, то есть иметь разные спектральные амплитуды и разные спектральные ширины. Мы сразу сузим задачу, предположив, что линии имеют бесконечно малую ширину и одинаковые амплитуды. Первое предположение на практике означает, что линии много уже полос фильтров анализатора, оба предположения – что мы хотим оценить предельное разрешение.

Количественной мерой *разрешающей способности* является наименьший частотный интервал (*интервал разрешения*) между двумя спектральными линиями, при котором они еще разрешаются спектроанализатором. Чем меньше интервал разрешения, тем лучше разрешающая способность. Но что значит «разрешаются» в точном математическом смысле? Об этом условимся чуть ниже.

Будем, как прежде, иметь в виду последовательный спектральный анализ. Сначала рассмотрим гипотетический СА с плавной перестройкой спектра, в котором дискретизации по частоте нет: спектр входного сигнала плавно (и очень медленно, чтобы результат анализа прямо не зависел от скорости перестройки) сдвигается путем непрерывного изменения частоты гетеродина. Для рассмотрения данной задачи удобно считать, что сдвигается не спектр, а анализирующий фильтр плавно перестраивается, не меняя ни формы АЧХ, ни ее ширины.

Сначала предположим, что в спектре входного сигнала имеется одна спектральная линия с частотой f_1 . По мере перестройки фильтра амплитуда выходного сигнала фильтра описывает кривую в функции времени, совпадающую с АЧХ как функцией частоты. При совпадении резонансной частоты фильтра с f_1 кривая достигает максимума.

Пусть теперь в спектре входного сигнала имеются две спектральные линии f_1 и f_2 одинаковой интенсивности. В резонаторе фильтра теперь одновременно существуют колебания двух частот и происходят биения с частотой $f_B = |f_1 - f_2|$. Амплитуда сложного выходного колебания теперь не постоянна, а периодична с периодом $T_B = 1/f_B$, повторяя с этим периодом свое наибольшее значение, равное сумме выходных амплитуд обеих спектральных составляющих. Поскольку перестройка резонатора происходит очень медленно, можно регистрировать эту сумму как амплитуду выходного сигнала. Каждая из выходных амплитуд отдельных линий повторяет однокоробую кривую, рассмотренную выше, максимумы кривых отстоят друг от друга на $f_B = |f_1 - f_2|$. При достаточно большом f_B суммарная кривая имеет провал (локальный минимум) на средней частоте $f_S = (f_1 + f_2)/2$ и «горбы» (локальные максимумы) на частотах f_1 и f_2 . Если теперь сближать частоты f_1 и f_2 , то выраженность провала уменьшается, затем исчезает вовсе, затем в точке f_S образуется максимум, а локальные максимумы в точках f_1 и f_2 исчезают.

Теперь мы готовы дать математическое определение термину «разрешается». Это можно сделать по-разному. Наиболее простое и удобное определение таково: линии разрешаются, если частотное расстояние между ними больше или равно частотному интервалу разрешения, который, в свою очередь, равен такой $f_B = |f_1 - f_2|$, при которой отношение амплитуды y_0 в провале к амплитуде y_{\max} максимумов равно заданному числу:

$$\frac{y_0}{y_{\max}} = q < 1.$$

Это уравнение можно, в принципе, решить относительно f_B , если задать АЧХ конкретного фильтра. Например, для простого резонансного контура с добротностью Q , относительная АЧХ которого:

$$y = \frac{1}{\sqrt{\left(\frac{f-f_0}{f_0}\right)^2 + \frac{1}{4Q^2}}},$$

решение уравнения имеет вид [2.2]:

$$f_B = |f_1 - f_2| = \frac{f_s}{Q} \sqrt{\frac{4}{q^2} - 1} = K_q \frac{f_s}{Q}.$$

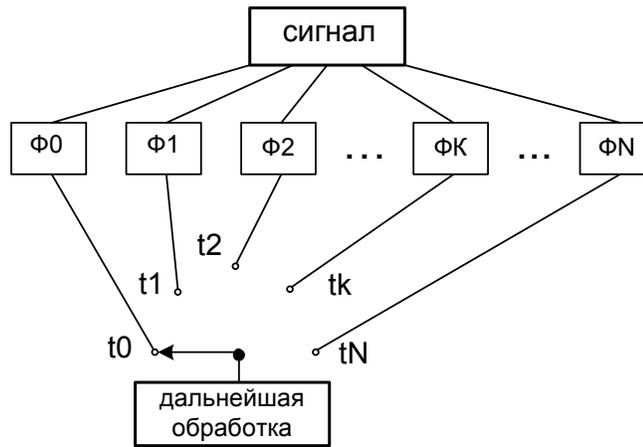
Как видим, полоса разрешения для простого резонансного контура равна произведению двух множителей. Множитель K_q определяется только произвольно принятым нами значением q , составляет порядок нескольких единиц, например, при $q=1/3, 1/2, 1/\sqrt{2}$ $K_q=5.92, 3.87, 2.65$ соответственно, и отражает надежность разрешения: чем глубже критический провал, тем больше K_q . Множитель f_s/Q есть ничто иное, как полоса того же фильтра при его настройке на f_s . При использовании фильтра более высокого порядка, имеющего более высокий коэффициент прямоугольности, множитель K_q уменьшается (при той же надежности разрешения).

Рассмотренная нами разрешающая способность называется *статической*, поскольку не учитывает временные эффекты, подобные влиянию момента начала сигнала, конечной скорости перестройки и т.п. Учет этих факторов довольно сложен и приводит к понятию *динамической разрешающей способности*, которую мы пока не рассматриваем.

Вспомним теперь, что в современном СА в качестве гетеродина используется синтезатор частот, благодаря чему спектр сигнала сдвигается на дискретные значения, кратные величине Δf . Опять, как и при рассмотрении последовательного анализа, удобно считать, что сдвигается не спектр, а анализирующий фильтр (на этот раз не плавно, а шагами по Δf) перестраивается, не меняя ни формы АЧХ, ни ее ширины. Резонатор анализирующего фильтра последовательно настраивается в резонанс на частотную точку $f_k = f_{\min} + k \cdot \Delta f$, $k = 0, 1, 2, \dots, N$ к моменту времени $t_k = t_0 + k \cdot t_{\text{точ}}$ и остается в этой настройке в течение времени $t_{\text{кр}} \gg 1/\Delta f$, затем к следующему моменту времени перестраивается на следующую частотную точку.

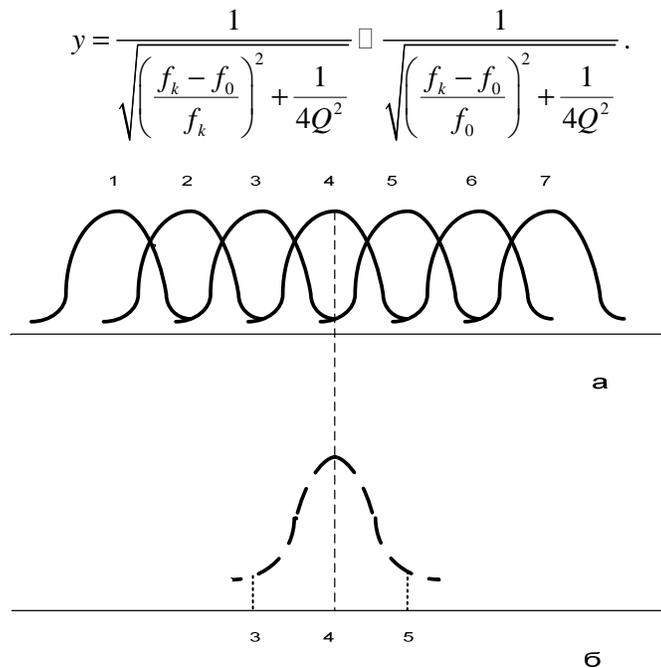
Заметим, что если бы фильтр на самом деле перестраивался, то для сохранения его полосы Δf_ϕ пришлось бы по мере перестройки менять его добротность в соответствии с формулой $\Delta f_\phi = f_k / Q_k$, но поскольку на самом деле анализирующий фильтр есть постоянный во времени фильтр ПЧ, постоянство его полосы выполняется автоматически. Далее, заметим, что полоса анализирующего фильтра (фильтра ПЧ) должна быть не больше и не меньше, а равна шагу частотной перестройки: $\Delta f_\phi = \Delta f$. Действительно, если будет $\Delta f_\phi < \Delta f$, то часть информации об анализируемом спектре сигнала будет утеряно; если будет $\Delta f_\phi > \Delta f$, то возникнет неоднозначность отсчетов.

Теперь мы готовы представить удобную виртуальную модель обработки сигнала в вышеописанных условиях, полезную не только при анализе разрешающей способности СА. Модель иллюстрируется рис. 2.1.1 и представляет схему параллельного спектрального анализа с последовательным опросом каналов (фильтров).



Виртуальная модель обработки сигнала при последовательном анализе

Пусть каждый фильтр модели (то есть фильтр ПЧ) построен на резонаторе низкого (например, первого) порядка, а входной сигнал содержит одну единственную спектральную линию на частоте f_0 , показанную на рис. 2.1.2а [2.2] вертикальной пунктирной линией. На этом рисунке показаны положения АЧХ фильтров модели, а пересечения пунктирной линии с кривыми показывают амплитуды отклика спектральной линии в разных фильтрах. Эти амплитуды можно привести к одной резонансной кривой, показанной на рис. 2.1.2б, поскольку их относительные значения (для фильтра первого порядка):



Таким образом, одна спектральная линия вызывает (разные по величине) отклики в нескольких фильтрах; идентифицировать линию можно по максимальному отклику. С увеличением N (то есть уменьшением $\Delta f = (f_{\max} - f_{\min})/N$) частотное расстояние между откликами уменьшается и при $N \rightarrow \infty$ спектральный отклик модели стремится к дискретному аналогу спектрального отклика рассмотренного выше последовательного анализа с непрерывной перестройкой фильтра. В реальных СА обычно N очень велико, например, $N = 400$. Поэтому с точностью до дискретизации откликов статическая разрешающая способность СА с дискретно-последовательным анали-

зом та же, что для рассмотренного выше СА с медленной непрерывной перестройкой, то есть задается полосой разрешения

$$|f_1 - f_2| = K_q \frac{f_s}{Q} = K_q \cdot \Delta f.$$

Заметим теперь, что для СА с дискретно-последовательным анализом использование в тракте ПЧ фильтра на резонаторе низкого порядка еще более неоптимально, чем для гипотетического СА с плавной перестройкой. Действительно, большая неравномерность вершины АЧХ придает разный вес спектральным откликам в пределах полосы фильтра, что требует трудной компенсации, а малая прямоугльность АЧХ приводит к перераспределению откликов на соседние полосы. Возникает императив: использование в фильтре ПЧ резонатора высокого порядка с высокой прямоугльностью. При этом множитель K_q уменьшается (при той же надежности разрешения). Если тракт последней ПЧ выполнен полностью цифровым, то есть и фильтрация в тракте осуществляется цифровым фильтром, то его прямоугльность может быть практически идеальной и множитель K_q близок к единице. В этом случае полоса разрешения практически равна шагу перестройки Δf .

Замечания:

1) До сих пор предполагалось, что шаг перестройки Δf равен полосе $\Delta f_{\text{посл ПЧ}}$ фильтра последней ПЧ. Обычно СА сам подбирает Δf в зависимости от установленной $\Delta f_{\text{посл ПЧ}}$. Если при текущей установке частотного диапазона анализа ($f_{\text{min}}, f_{\text{max}}$) и числа шагов N получающееся отношение $\Delta f_{\text{отн}} = (f_{\text{max}} - f_{\text{min}}) / N$ больше $\Delta f_{\text{посл ПЧ}}$, то СА устанавливает $\Delta f = \Delta f_{\text{посл ПЧ}}$, а компенсация производится при дисплейном детектировании с автоматически устанавливаемым эквивалентным видео-сглаживанием в $f_{\text{отн}} / \Delta f_{\text{посл ПЧ}}$ раз. Если же получается $\Delta f_{\text{отн}} < \Delta f_{\text{посл ПЧ}}$, то опять СА устанавливает $\Delta f = \Delta f_{\text{посл ПЧ}}$, а реализуемое число частотных точек становится больше установленного в $\Delta f_{\text{посл ПЧ}} / \Delta f_{\text{отн}}$ раз.

2) При анализе разрешающей способности СА мы не учитывали влияние шумов и помех, в частности, фазовых шумов гетеродинов.

Классификация анализируемых сигналов

Классификация сигналов как упорядоченный перечень потенциальных для анализа объектов суть классификация подходящих функций времени во временной и частотной областях. Такая классификация не может быть замкнутой и должна ограничиваться субъективным отбором признаков, влияющих на конструктивные решения при разработке СА, на параметрические установки – при эксплуатации СА, на эффективность проводимого анализа, на интерпретацию результатов анализа.

На этапе разработки СА классификация сигналов двусторонне связана с шириной претензий на универсальность прибора: с одной стороны, перечень видов сигналов помогает уточнить параметрические и характеристические границы желаемого, с другой – технические и стоимостные ограничения диктуют достижимые границы в пространстве сигналов.

Особая задача – интерпретация результатов измерения, в частности, математическая аппроксимация наблюдаемых результатов, например, амплитудных спектров; эта задача решается тем успешнее, чем четче априорное разделение сигналов на классы.

Нужно иметь в виду, что универсальный СА СВЧ должен удовлетворять потребностям анализа информативных сигналов не только в области радиоэлектроники, но в любых областях науки и техники, где используются СВЧ-сигналы для получения и передачи информации. Анализируемые сигналы могут поступать на СА как непосредственно с генератора, так и с выхода при-

емной антенны, с датчика экспериментальной аппаратуры, с выхода макета конструируемой аппаратуры и т.д.

1) Анализируемые колебания могут представлять детерминированный процесс, случайный процесс, квазидетерминированный процесс, смешанный детерминированно-случайный процесс.

а) *Детерминированный процесс* – математическая идеализация, поскольку любой сигнал всегда содержит долю случайных составляющих и не вполне прогнозируем на будущее. Например, сигнал любого генератора подвержен фазовым и амплитудным флуктуациям; любая цепь, через которую проходит сигнал, вносит стохастические элементы и шумы и т.д. Но выделение в определенных случаях и на приемлемых условиях и приближениях класса детерминированных сигналов совершенно необходимо для решения задач, где предполагается (хотя и приближенно) полное описание и полная прогнозируемость сигналов; именно эти задачи составляют большую часть радиотехники. Следует особо подчеркнуть, что один и тот же процесс, в зависимости от необходимой точности описания и прогнозирования и других требований, может быть отнесен или не отнесен к классу детерминированных процессов.

б) *Случайные процессы* – традиционная область путаницы, неточных представлений, нестрогих определений и других феноменов обыденного сознания в инженерной среде. Например, часто под случайными сигналами подразумевают неожиданные, спорадические, или редкие, или даже непериодические и т.д. сигналы. Иногда полагают, что для случайных процессов характерно отсутствие четкого описания, строгих характеристик, какой-либо возможности прогноза, экстраполяции. Степень неадекватности представлений о случайных процессах зависит от «просвещенности» того или иного работника.

в) Важность этого класса процессов для аппаратного анализа сигналов возрастает год от года и определяется колоссальным количеством видов флуктуаций и помех, сопровождающих генерацию и формирование сигналов, распространение их по каналам различной физической природы (линии передачи, атмосфера и т.д.), их прием, обработку, индикацию. С другой стороны, технологическое, системное, информационное развитие радиоэлектроники приводит к необходимости рассмотрения и учета все более слабых флуктуаций. Так, до определенного периода фазовые флуктуации высокочастотных генераторов СВЧ были просто интересным явлением, мало влияющим на эффективность радиотехнических систем, сейчас же качественные показатели ряда этих систем критически зависят от уровня фазовых флуктуаций применяемых генераторов. Надо признать, что радиоэлектроника и ее приложения являются «чемпионом» по использованию статистических подходов и методов. Ни в какой области технического знания так широко и глубоко не представлены статистические методы анализа, синтеза и оптимизации сигналов и систем, как в радиоэлектронике и ее разделах: беспроводной связи, радиолокации, оптическом и радио-зондировании сред (активном и пассивном), радионавигации и многих других.

г) Случайный процесс может быть описан на одном из двух уровней: на *уровне распределений* до некоторого порядка и некоторой точности, или на уровне конечного числа числовых характеристик, в частности, *уровне моментов*. Уровень распределений более информативен, но более сложен в математической реализации и интерпретации результатов математического анализа и экспериментов. С одной стороны, теоретический вывод распределений вероятностей процесса (особенно второго и более высоких порядков) часто бывает неоправданно сложным или даже недоступным в конкретных физических и иных задачах; к тому же он может потребовать ряда допущений, которые затруднительно обосновать. С другой стороны, экспериментальное определение распределений случайных процессов предъявляет серьезные требования к измерительным приборам и системам обработки данных. Более того, существует огромное число теоретических задач, которые практически неразрешимы на уровне уже двумерных, а иногда даже одномерных распределений, по крайней мере, в аналитическом виде, пока не будут сделаны некоторые радикальные предположения, например, о гауссовости случайного процесса или поля. Достаточно привести примеры задач, которыми изобилует статистическая радиотехника: о линейном инерционном преобразовании, нелинейном безынерционном, тем более – нелинейном инерционном преобразовании случайного процесса. Зачастую такие радикальные предположения

делаются в интересах «решения» задачи даже тогда, когда они заведомо не согласуются с реальностью задачи, а иногда даже со здравым смыслом. Насколько радикально, например, предположение о гауссовости процесса на входе нелинейного преобразования, в ситуации, когда оно неверно, видно уже из того, что для стационарного гауссовского процесса вся информация, содержащаяся в многомерном распределении любого порядка и любой точности, содержится уже в корреляционной функции процесса, другими словами, описание на уровне распределений автоматически воспроизводится по корреляционной функции, т. е. по описанию на уровне первых и вторых моментов. Иногда в описании на уровне распределений нет особой необходимости и можно ограничиться описанием на уровне моментов. При использовании конечного числа моментов это описание заведомо неполное (если класс распределений не задан), но может удовлетворять потребности конкретного приложения. В других случаях переход на этот уровень может быть вынужденной мерой. Кроме того, иногда именно временная зависимость некоторых конкретных моментов представляет главный интерес. Описания на уровне моментов отличаются друг от друга наивысшим порядком участвующих в описании моментов. Среди этих описаний исключительное практическое значение имеет то, которое использует моменты до второго порядка включительно. Оно включает одноточечный момент первого порядка (математическое ожидание) и двухточечный момент второго порядка (корреляционную функцию), и может быть названо *корреляционным описанием* (рассмотрения на этом уровне описания часто называют «корреляционными теориями»). Используя СА, мы имеем дело, в первую очередь, именно с корреляционным описанием, доставляемое измерением *спектральной плотности мощности (энергетического спектра)* (кроме того, опционально, можно по получаемым данным рассчитывать моменты более высокого порядка, гистограммы распределений, сглаживать их и т.д.). Если есть основания считать случайный процесс стационарным по крайней мере в широком смысле (см. ниже и главу 6), то обратное преобразование Фурье от энергетического спектра есть *ковариационная функция* – второй двухточечный центральный момент как функция разности аргументов процесса. Важность и даже уникальность ковариационной функции как характеристики случайного процесса обусловлена двумя обстоятельствами. Во-первых, она полностью описывает структуру *линейной* статистической зависимости (корреляции) процесса. Во-вторых, ее Фурье-образ есть энергетический спектр процесса, то есть полностью описывает распределение средней мощности процесса по частотам. Вместе пара ЭС – КФ описывает статистическую динамику случайного процесса как во временной области (КФ), так и в частотной области (ЭС). В подборе математических аппроксимаций пар ЭС-КФ может помочь работа [2.5], где собран большой набор пар. Таким образом, еще один классификационный признак для случайных процессов – форма энергетического спектра (ковариационной функции).

д) Существует обширный и практически важный для радиоэлектроники класс случайных процессов, занимающих промежуточное положение между *чисто-случайными* и детерминированными процессами и называемых *квазидетерминированными процессами*. Временные реализации этих процессов детерминированы, но *имеют случайные параметры*. Например, это может быть синусоида со случайной амплитудой, или со случайными амплитудой и частотой; или это «пила» со случайным наклоном и/или периодом и/или эпохой; или это последовательность импульсов с одинаковой случайной шириной и/или постоянной случайной скважностью и т.д. и т.п. Имеется в виду, что случайные параметры реализуются в начале временной реализации процесса, а затем остаются постоянными. О квазидетерминированном процессе можно правильно судить, наблюдая не менее двух его реализаций: наблюдая только одну реализацию, можно решить, что это детерминированный процесс, но вторая реализация покажет, что это не так. Для того же, чтобы получить представление о свойствах случайных параметров (например, дисперсиях или распределениях), нужно наблюдать достаточно много реализаций.

е) Также промежуточное положение, но в другом смысле, занимают смешанные *детерминированно-случайные* процессы. Реализация такого процесса суть некоторая комбинация реализаций детерминированного и случайного процессов. К таким процессам относятся, в частности, детерминированные узкополосные сигналы, модулированные по амплитуде и/или по фазе (частоте) случайным процессом, и, наоборот, случайные процессы, модулированные по амплитуде и/или фазе детерминированным процессом. К этому классу примыкают формально детер-

минированные сигналы со сложными М-арными модуляциями (последние в некоторых случаях можно считать шумоподобными).

2) Сигнал может быть периодическим или аperiodическим.

а) Детерминированный *периодический процесс* – функция времени, удовлетворяющая условию $f(t+T) = f(t)$ для любого $0 < t < \infty$, где T называется периодом. Поскольку целое число величин T – также период, иногда периодом называется наименьшая из величин, для которых удовлетворяется вышенаписанное условие. На практике периодическим часто считается сигнал, существующий на конечном интервале времени, если этот интервал много больше «периода». Как известно, преобразование Фурье от строго периодического процесса есть ряд Фурье, дающий амплитудный спектр в виде эквидистантной последовательности дельта-функций. Однако на практике спектральные линии имеют конечную ширину из-за конечности длительности анализируемой реализации и из-за неизбежных флуктуаций, делающих процесс приближенно периодическим, например, из-за фазовых шумов.

б) Случайный процесс общего вида не может быть периодическим в строгом смысле слова. Однако некоторые подклассы случайных процессов могут быть периодическими в условном смысле. Для квазидетерминированных процессов условная периодичность означает периодичность его реализаций (может быть, с разными периодами). Для детерминированно-случайных процессов вида «случайный процесс с детерминированной амплитудной модуляцией» условная периодичность означает периодичность модулирующей функции.

3) Случайный процесс может быть стационарным в каком-либо смысле, или нестационарным. Стационарность случайного процесса это свойство инвариантности тех или иных его характеристик относительно сдвига во времени. Поскольку число характеристик процесса бесконечно, то и видов стационарности бесконечное число. Для спектрального анализа случайных процессов особое значение имеют *стационарность в широком смысле* и *стационарность в узком смысле (строгая стационарность)*. Часто при обсуждении вопросов аппаратного спектрального анализа изначально предполагают, что анализируемый процесс стационарен по крайней мере в широком смысле. На самом деле, спектроанализаторы рассматриваемого нами типа имеют определенные возможности спектрального анализа и некоторых типов нестационарных процессов. Оставляя пока рассмотрение этого вопроса, заметим только, что важный для приложений класс нестационарных процессов – периодически нестационарные случайные процессы, характеристики которых – периодичные функции времени.

4) По спектральному составу сигналы можно разделить на узкополосные, широкополосные и общего вида.

а) Сигнал называется *узкополосным*, если почти вся его мощность сконцентрирована в полосе $\Delta f \ll f_0$ вокруг характерной частоты (например, частоты максимальной спектральной амплитуды) f_0 . Например, узкополосным является сигнал генератора с нагруженной добротностью $Q \gg 1$. Многие современные радиотехнические цепи и системы предъявляют к сигналам требования по спектральной концентрации, для которых данное определение слишком неконкретно. Например, может потребоваться, чтобы 99,9% мощности сигнала концентрировалась в полосе $\Delta f \leq 0,001 \cdot f_0$. В диапазоне СВЧ класс узкополосных сигналов – основной; именно на такие сигналы рассчитаны линии передачи и многие другие устройства.

б) *Широкополосным* обычно называют сигнал, в спектре которого отношение максимальной и минимальной частот много больше единицы.

5) В математике функцию, отличную от нуля на конечном носителе, называют *финитной*. С этой точки зрения все сигналы – финитные функции времени, так как существуют лишь на конечном интервале времени. Более актуально сравнивать время существования сигнала с каким-либо интервалом, характерным для анализирующего прибора, например, временем развертки частоты, или временем сеанса измерения и т.п. Если сигнал финитен в такого рода смысле, его спектральные свойства имеют особые черты. Аналогично, если сигнал имеет финитный спектр,

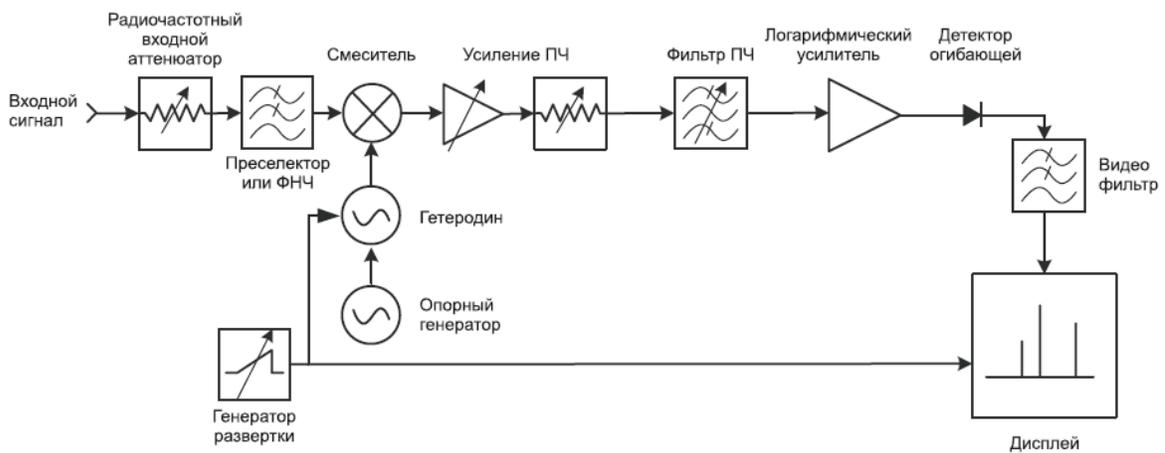
то он дает специфические проявления во временной области. Это частные случаи общего принципа: ограничения на аргументы в частотной области порождают специфичное поведение во временной области, и наоборот. Например, квазимонохроматический детерминированный сигнал – частный вид узкополосных сигналов – допускает специфичные представления (через огибающую и фазу, через квадратуры) во временной области. Аналогично, узкополосный случайный процесс допускает похожие специфичные представления своей ковариационной функции.

б) Специфические классы сигналов составляют детерминированные и случайные процессы со скважностью во временной области или в частотной области. Сигналы со скважностью во временной области являются обобщением импульсной сигнальной последовательности, сигналы со скважностью в частотной области – обобщением периодических сигналов. В соответствии с высказанным принципом, первая группа сигналов имеет специфические характеристики в частотной области, вторая – во временной области.

Структура спектроанализатора

Современные спектроанализаторы (СА) в разной и все возрастающей степени используют цифровую обработку сигналов. Но описание структуры СА удобно и полезно начать с классической полностью аналоговой схемы.

Структурная схема аналогового СА



Упрощенная структурная схема аналогового СА СВЧ

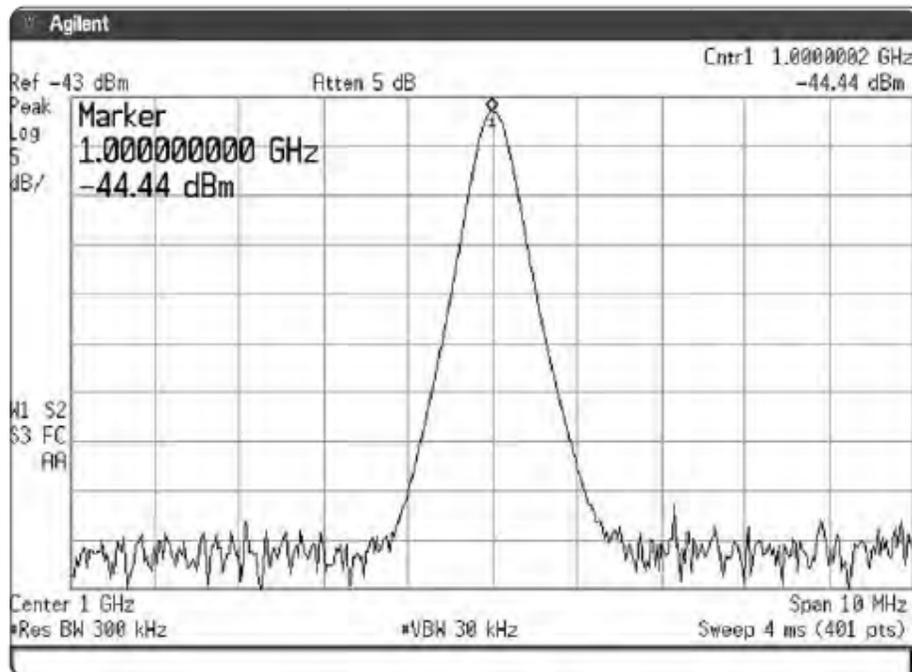
Входной сигнал СВЧ проходит регулируемый (автоматически и, в необходимых случаях, вручную) радиочастотный аттенюатор, управляемый фильтр нижних частот, или фильтр, называемый *преселектором*, и подвергается преобразованию частоты. На схеме показано однократное преобразование, на самом деле в СА используется от двух до четырех каскадов преобразования, что решает известные проблемы устранения зеркального канала, выбора оптимальной промежуточной частоты (ПЧ) и др. Гетеродин первого каскада преобразования (первый гетеродин) выполнен как синтезатор, дискретно перестраиваемый по частоте, давая эквидистантную сетку частот. Управляющая этим гетеродином цепь условно показана как «генератор развертки». После смесителя на схеме показан тракт последней промежуточной частоты (тракты остальных ПЧ стереотипны). Принципиальные элементы в нем: усилитель с регулируемым коэффициентом усиления, управляемый аттенюатор и управляемый полосовой фильтр с дискретным набором полос. Показанный на схеме условно детектор огибающей символизирует как последний, так и набор переключаемых оператором т.н. *дисплейных детекторов* разного вида. Принципиальный элемент видео-тракта – ФНЧ (видео-фильтр), символизирующий набор фильтров с разной полосой, позволяющий устанавливать разную степень сглаживания отображаемого спектра. Визуальное отображение результатов анализа осуществляется калибруемым по частоте и уровню дисплеем, синхронизованным со схемой управления первого гетеродина.

Первый гетеродин

Первый гетеродин последовательно генерирует сетку частот, обеспечивая соответствующую сетку входных частот СА $f_k = f_{\min} + k \cdot \Delta f$, $k = 0, 1, 2, \dots, N$, $N = (f_{\max} - f_{\min}) / \Delta f$. Минимальная и максимальная частоты f_{\min}, f_{\max} анализируемого диапазона устанавливаются оператором до и во время сеанса, шаг перестройки Δf обуславливает разрешающую способность. Из двух теоретических возможностей соотношения частот сигнала f_c и гетеродина f_r : $f_r < f_c$ или $f_r > f_c$, в данном случае возможна только вторая, т.к. f_{\min} может быть очень малой, например, 10 МГц. Первый гетеродин может отвечать предъявляемым к нему требованиям только если выполнен как *синтезатор частот*. Важнейшие из этих требований таковы:

- максимальный устанавливаемый диапазон перестройки должен соответствовать максимальному частотному диапазону СА: от $f_{r_{\min}} = f_{\min} + f_{\text{ПЧ1}}$ до $f_{r_{\max}} = f_{\max} + f_{\text{ПЧ1}}$, где, в данном случае, f_{\min}, f_{\max} – минимально и максимально возможные устанавливаемые частоты диапазона сигналов; максимальный частотный диапазон как правило очень широк, например, от 10 МГц до 4, 10, 20 или даже 50 ГГц;
- шаг перестройки Δf должен варьироваться в очень широких пределах, а минимальный шаг быть очень малым, например, $\Delta f = 10^k$ Гц, $k = 0, 1, \dots, 5$; заметим, что, например, при шаге перестройки 1 Гц и диапазоне перестройки 0.01–20 ГГц 1-ый гетеродин за один цикл перестройки должен настроиться на $\square 2 \cdot 10^{10}$ частотных точек (!), причем на каждую – с ошибкой $\ll 1$ Гц (!);
- время $t_{\text{пер}}$ переключения на следующую частотную точку и время $t_{\text{пер.раз}}$ переключения на следующую развертку частоты должны быть достаточно малыми, например, $t_{\text{пер}}$ порядка 1–5 мкс (в зависимости от Δf), $t_{\text{пер.раз}}$ порядка 1 мс;
- уровень фазового шума гетеродина должен быть достаточно низким.

Дисплей



Вид обычного дисплея СА

В разных моделях СА функции дисплея выполняет или ноутбук, осуществляющий одновременно управление прибором, обработку данных и вычисления, или собственный индикатор, оформленный как отдельный блок или входящий в общий блок прибора. И в том, и другом случае применяется цифровой формат сигнала, выводимого на экран. Такая индикация называется цифровой, а сам дисплей – «цифровым дисплеем». Применение цифровой техники позволяет легко выводить на экран обширную информацию о действующих установках, калибровках, маркерах и т.д., располагая ее на периферии экрана.

До появления и массового применения цифровых дисплеев индикаторы СА прошли долгий путь развития и совершенствования: от чисто аналогового дисплея (как и всего СА) до середины 70-х годов, главной проблемой которого была необходимость применения долгой развертки при узких полосах разрешения, затем – накопительная ЭЛТ с переменным свечением и регулировкой скорости угасания изображения, требующая перенастройки того и другого для каждого нового сеанса, наконец, цифровая индикация, дающая ряд преимуществ. Теперь, как только кривая представлялась в цифровом виде и записывалась в память, она становилась доступной для отображения на дисплее; стало просто обновлять изображение, минуя мерцания, размытость и блеклость. Данные в памяти обновляются со скоростью развертки, а содержимое памяти отображаются на экране без мерцания, поэтому можно следить за этими обновлениями одновременно с тем, как анализатор проходил по установленной полосе обзора частот.

Графическая часть экрана обычно разграфлена масштабной сеткой из горизонтальных и вертикальных линий (обычно по 10 тех и других). По горизонтали откладывается частота в линейном масштабе, с указанием опорного уровня, по вертикали – спектральная плотность мощности (СПМ) в линейном или (чаще) логарифмическом масштабе, с указанием опорного уровня, уровня верхней горизонтальной линии, цены деления, что позволяет непосредственно или, более точно, с помощью маркера определять абсолютную величину СПМ, например, в дБм/Гц. Вообще, с помощью одного маркера обычно можно определять частоту и мощностной уровень заданной точки кривой, а с помощью двух маркеров – частотную и мощностную разность (в децибельной мере) двух точек.

На рис. 2.3.2 показаны отображения на экране типичного дисплея СА. Предварительно сделаны установки ширины частотной развертки (*Span*), числа частотных дискретов N , полосы фильтра ПЧ (*Res BW*), полосы видео-фильтра (*VBW*), опорного уровня мощности (верхняя горизонтальная линия) и т.д. Кривая обнаруженной спектральной линии поставлена оператором в удобное положение по горизонтали в центре экрана (с помощью т.н. блока управления центральной частотой) и выбран удобный масштаб по вертикали, а верхняя точка кривой путем ручной установки опорного уровня (верхней горизонтальной прямой) помещена близко к этому уровню. Вместо интерпретации наблюдаемой кривой в терминах СПМ, можно эквивалентно считать, что кривая показывает некоторую безразмерную функцию (например, в дБ) $\varphi(f)$, $\varphi(0) = 0$ в сочетании с указанной мощностью сигнала (например, в дБм).

Радиочастотный аттенюатор

Первый по ходу сигнала существенный элемент схемы на рис. 2.3.1 – радиочастотный управляемый аттенюатор. Его назначение – обеспечить приемлемый (или даже оптимальный) уровень мощности сигнала на входе первого смесителя, избегая перегрузки, сжатия усиления и искажений. Обычно ослабление этого аттенюатора регулируется автоматически при ручной установке опорного уровня мощности во время действия сигнала. Кроме того, возможна ручная шаговая установка ослабления в широких пределах, например, в диапазоне 70 дБ с шагом 1, или 2, или 5 дБ [2.6].

Из не показанных на схеме элементов отметим блокирующий конденсатор на входе сигнала, предотвращающий повреждение СА сигналом (компонентой) ПТ или смещением постоянной составляющей информативного сигнала. Образуя с сопротивлением нагрузки RC -цепочку, он создает фильтр ВЧ, частотнозависимо ослабляющий низкие частоты, что заставляет выбирать нижнюю частоту диапазона анализа f_{\min} достаточно большой.

Другой не показанный элемент (возможно, опциональный) – специальный прецизионный опорный генератор с фиксированной частотой и амплитудой, подключаемый ко входу СА для более точной автокалибровки. При подключении этого генератора блокирующий конденсатор отключается.

Преселектор

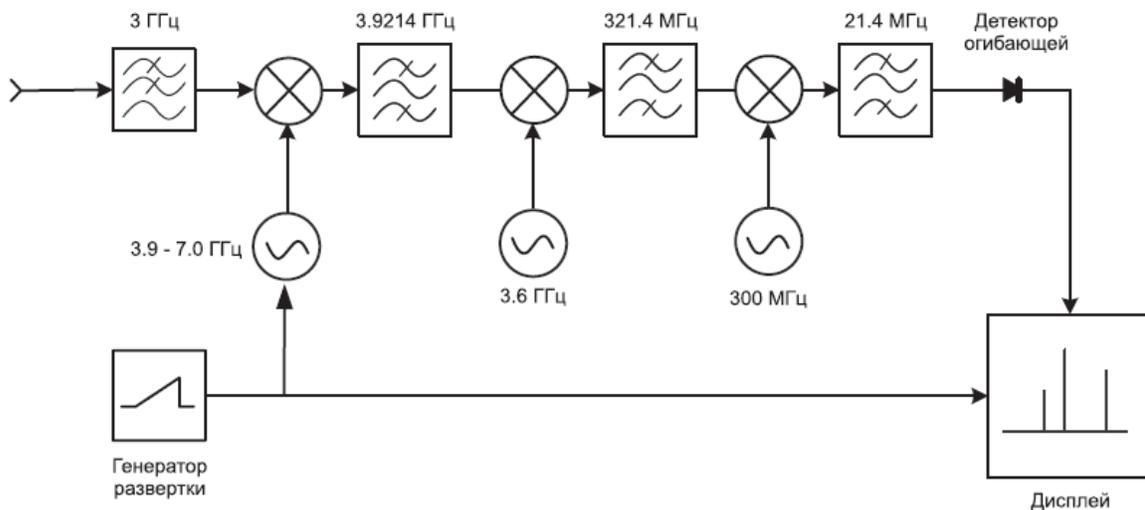
В диапазоне СВЧ недостаточно селективности входных сигналов с помощью ФНЧ, обычного для СА более низких частот и препятствующего попаданию на первый смеситель внеполосных сигналов, преобразующихся в нежелательные отклики на ПЧ; по ряду причин приходится применять перестраиваемый фильтр, называемый *преселектором*.

Частотный план

Как всегда при разработке приемника с несколькими преобразованиями частоты, выбор промежуточных частот, частот второго и последующих гетеродинов и диапазона перестройки первого гетеродина осуществляется с помощью построения частотного плана, позволяющего избежать ряда «опасностей». В результате выбора, частота сигнала $f_{\text{сиг}}$ при данной настройке $f_{\text{гет1}}$ удовлетворяет уравнению

$$f_{\text{сиг}} = f_{\text{гет1}} - f_{\text{ПЧ1}} = f_{\text{гет1}} - (f_{\text{гет2}} + f_{\text{гет3}} + \dots + f_{\text{посл ПЧ}}).$$

Приведем лишь пример результата выбора [2.6] (рис. 2.3.3) применительно к тройному преобразованию и входному диапазону частот сигналов 0–3 ГГц.



Пример частотного плана при тройном преобразовании частоты и входном диапазоне частот сигналов 0–3 ГГц

Усилитель тракта ПЧ с регулируемым усилением

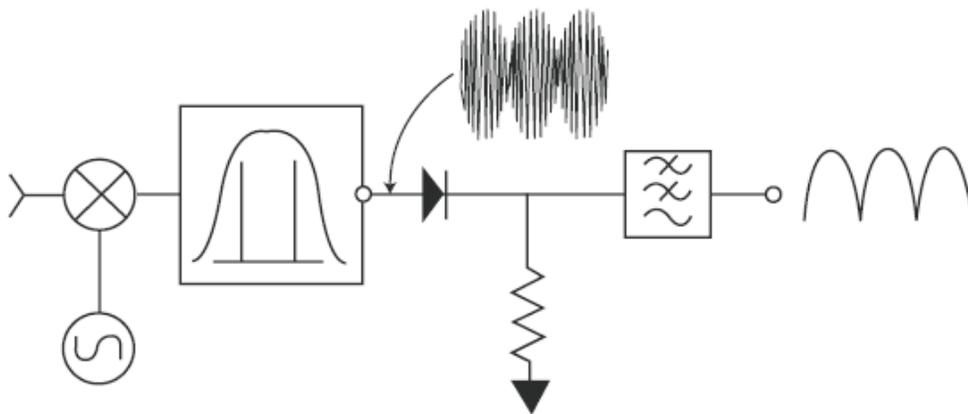
Следующий существенный элемент, показанный на схеме рис. 2.3.1, – усилитель тракта последней ПЧ с регулируемым коэффициентом усиления. Он используется для подстройки вертикального положения кривой спектра сигнала на дисплее без воздействия на уровень сигнала на первом смесителе. Соответственно этой подстройке меняется опорный уровень мощности (мощность, соответствующая верхней горизонтальной линии), так что отображение СПМ сигнала не нарушается. Чтобы опорный уровень мощности не изменялся при ручной регулировке ослабления входного аттенюатора, последний сопряжен со схемой регулировки усиления обсуждаемого усилителя так, чтобы происходила полная компенсация изменения мощности на выходе тракта последней ПЧ и кривая спектра на дисплее оставалась в том же положении.

Фильтр последней ПЧ

Поскольку фильтр последней ПЧ имеет намного меньшую полосу, чем фильтры остальных ПЧ, именно он является тем спектральным прибором, о необходимости которого для аппаратного спектрального анализа говорилось в 2.1.2. Российский ГОСТ предписывает называть такого рода фильтры *селективными анализирующими*. Именно полоса и прямоугольность этого фильтра определяют разрешающую способность СА. Если этот фильтр аналоговый, то для повышения разрешающей способности СА, как указывалось выше (см. 2.1.7), необходим фильтр высокого порядка, обладающий, при надлежащем конструировании, большой прямоугольностью АЧХ. Прямоугольность АЧХ фильтра обычно оценивается *полосовой избирательностью (избирательностью, коэффициентом прямоугольности)*, определяемой как отношение полос на низком и высоком уровнях АЧХ, например, на уровнях -60 дБ и -3 дБ. Чем меньше это отношение, тем выше избирательность (прямоугольность). Например, аналоговые фильтры, применявшиеся когда-то в СА фирмы Аджилент [2.6], были четвертого порядка (четыреполосные), выполнены по синхронно настроенной схеме, имели форму АЧХ, близкую к гауссовой, с избирательностью 12.7:1 по уровням -3, -60 дБ. Еще раньше та же фирма применяла для самых узких устанавливаемых полос фильтры пятого порядка с избирательностью 10:1 по тем же уровням. Применение цифровых фильтров (для самых узких или для всех устанавливаемых полос) позволяет достичь еще более лучших избирательностей.

Детектор огибающей

Детектор огибающей – стандартный элемент тех приемников, которые извлекают информацию из временного поведения амплитуды сигнала; поэтому он не требует какого-либо обстоятельного рассмотрения. Кратко прокомментируем лишь выбор инерционности аналогового детектора в обстоятельствах спектрального анализа. В своей простейшей форме схема детектора включает диод и RC- цепочку в качестве нагрузки (с учетом паразитных емкостей). Если в полосе $\Delta f_{\text{посл ПЧ}}$ фильтра последней ПЧ в данный момент находится одна спектральная компонента стационарного (в смысле неизменности спектра за время частотной развертки) сигнала, то амплитуда сигнала на входе детектора постоянна. Но если эта полоса вмещает две (или более) неразрешенных спектральных компонент, то между ними происходят биения, картина которых (для случая двух компонент) показана на рис. 2.3.4.



Картина биений в тракте ПЧ и выходной сигнал детектора

Время между соседними пучностями (узлами) равно $T_n = 1/2f_b$, где f_b – частота биений, в свою очередь равная частотному расстоянию между центрами спектральных составляющих f_1, f_2 . Поскольку максимально возможное расстояние при условии неразрешения равно $\Delta f_{\text{посл ПЧ}}$, то инерционность детектора должна удовлетворять условию $T_n \leq 1/2 \cdot \Delta f_{\text{посл ПЧ}}$, зависящему от полосы

фильтра последнего ПЧ. Поэтому хорошим решением является переключать инерционность при переключении полосы.

Обработка сигналов в СА

До сих пор рассматривалась структура полностью аналогового СА и предполагалась полностью аналоговая обработка сигналов во всех блоках СА. Общая тенденция развития СА – постепенный переход к цифровой схемотехнике и обработке «от конца к началу» СА: сначала в видео тракте и управлении дисплеем, затем также в части тракта последней ПЧ (например, в отношении самых узкополосных, а затем всех фильтров ПЧ), затем также в отношении всего тракта последней ПЧ и т.д. Как указывалось выше, этот процесс «цифровизации» сдерживается лишь недостаточной скоростью АЦП. В настоящее время большинство моделей СА выполнены в архитектуре с полностью цифровой последней ПЧ. Конечно, цифровая и аналоговая обработки сигналов различаются в ряде отношений, как и характеристики их эффективности. На этом этапе мы лишь кратко отметим эти различия для двух случаев: СА с цифровой обработкой в видео-тракте и в управлении дисплеем, и СА с цифровой обработкой начиная с тракта последней ПЧ. Рассматривать обработку сигнала в полностью аналоговом СА нет смысла: в настоящее время такой СА вряд-ли будет разрабатываться и производиться.

Последовательность преобразований сигнала

а) СА с цифровой обработкой только в видео-тракте и в управлении дисплеем. Глядя на структурную схему рис. 2.3.1, имеем следующую последовательность преобразований:

– подаваемый на входной разъем СА сигнал испытывает ослабление в управляемом входном радиочастотном аттенуаторе, регулируемом автоматически и мануально (в частном случае аттенуатор может находиться в положении нулевого ослабления);

– сигнал фильтруется перестраиваемым фильтром СВЧ, называемым преселектором;

– сигнал проходит несколько преобразований частоты, причем 1-ый гетеродин дискретно по частоте и по времени перестраивается, реализуя эквидистантные сетки времени и частоты; в результате на выходе 1-го смесителя в каждый момент времени присутствует сдвинутая по спектру копия входного сигнала; именно на этапе первого смешения к собственным флуктуациям сигнала прибавляются существенные (в сравнении с шумами входного аттенуатора и преселектора) флуктуации за счет фазовых и амплитудных шумов 1-го гетеродина и схемы управления им, а также шумов 1-го смесителя; для уменьшения первых обычно используется фазовая привязка к высокостабильному опорному генератору (показана на схеме, обычно кварцевый генератор) и схема оптимизации фазовых шумов (см. ниже), для уменьшения вторых необходимо тщательно выбирать режим смесителя; кроме того, «медленные» флуктуации мощности 1-го гетеродина, в зависимости от его конструкции, могут достигать величин, требующих применения автоматической регулировки его мощности (АРМ); каждый тракт от первого до предпоследнего ПЧ как минимум содержит фильтр постоянной настройки;

– после последнего смесителя сдвинутая по спектру (в соответствии с моментом времени) и фильтрованная в предыдущих трактах ПЧ копия сигнала попадает в тракт последней ПЧ, где подвергается регулируемому (в процессе подстройки вертикального положения кривой спектра на дисплее и в процессе ручной регулировки ослабления входного аттенуатора) усилению; ослаблению (возможно, нулевому) в регулируемом аттенуаторе последней ПЧ; фильтрации в фильтре последней ПЧ с устанавливаемой, в широком диапазоне значений, полосой $\Delta f_{\text{посл ПЧ}}$, причем время $t_{\text{ст}}$ «стояния» на очередной точке сетки частот находится в обратной зависимости от $\Delta f_{\text{посл ПЧ}}$; возможно, пропусканию через логарифмический усилитель (для расширения динамического диапазона СА);

– сигнал с выхода тракта последней ПЧ детектируется амплитудным детектором огибающей;

– с выхода детектора огибающей начинается полностью цифровой видеотракт: сигнал оцифровывается и дальнейшие операции выполняются с помощью цифровой техники; как минимум видеотракт содержит цифровой видео-фильтр с устанавливаемой полосой; этот фильтр чрезвычайно важен, т.к. он устраняет неинформативные колебания после детектора и осуществляет сглаживание получаемого спектра; эти две функции могут быть разделены, а функция сглаживания выполняется с помощью скользящего взвешенного интегрирования (суммирования) с регулируемой постоянной времени;

– оцифровка это дискретизация по времени и квантование по уровню; квантование должно обеспечивать достаточно большое число двоичных разрядов для получения нужной точности, дискретизация должна быть достаточно частой для отслеживания информативных изменений амплитуды; частота $f_{\text{дискр}}$ дискретизации почти всегда больше частоты $f_{\text{экр}}$ нанесения точек кривой на экране дисплея, особенно если $\Delta f_{\text{отн}} = (f_{\text{max}} - f_{\text{min}}) / N$ больше $\Delta f_{\text{посл ПЧ}}$; поэтому каждая нанесенная точка есть результат обработки набора цифровых значений (каждый такой набор назовем *блоком отсчетов*, а соответствующую обработку – *комбинированием отсчетов* (иногда [2.6] эту обработку называют дисплейным детектированием); обычно в СА предусмотрено на выбор несколько видов комбинирования отсчетов; в разных СА реализованы разные варианты сочетания видео-фильтра и блока комбинирования отсчетов; так или иначе после видео-фильтра сигнал проходит блок комбинирования отсчетов, дающий одно значение на блок отсчетов; это значение как очередное отображается на экране дисплея; в необходимых случаях осуществляется интерполяция для представления последовательности точек на экране в виде непрерывной кривой.

б) СА с цифровым трактом последней ПЧ.

– обработка сигналов от входа СА до последнего смесителя совпадает со случаем а);

– после аналогового последнего смесителя и до оцифровки сигнал проходит через (аналоговый) полосовой фильтр с центром на последней ПЧ; этот фильтр, при необходимости, то есть когда частота $f_{\text{дискр}}$ отсчетов не достигает двух $f_{\text{посл ПЧ}}$ и критерий Найквиста не выполняется, выполняет функцию защиты от наложения спектров (элиэзинга), заодно он отфильтровывает вредные и неинформативные составляющие на выходе смесителя;

– далее сигнал оцифровывается АЦП с частотой $f_{\text{дискр}}$ отсчетов; все дальнейшие преобразования осуществляются в цифровом формате сигналов;

– следующее преобразование – прохождение сигнала через цифровой фильтр последней ПЧ с устанавливаемой, в широком диапазоне значений, полосой $\Delta f_{\text{посл ПЧ}}$, причем время $t_{\text{ст}}$ «стояния» на очередной точке сетки частот находится в обратной зависимости от $\Delta f_{\text{посл ПЧ}}$;

– далее выполняется «цифровое детектирование» оцифрованного сигнала: вычисляются «медленно меняющиеся» квадратуры I и Q и каждый отсчет находится как $\sqrt{I^2 + Q^2}$;

– остальные преобразования – те же, что в случае а).

Комбинирование цифровых отсчетов амплитуды

Задача комбинирования цифровых отсчетов амплитуды ставится так: на одну отображаемую на экране точку (в дальнейшем интерполированной) кривой приходится $M \gg 1$ цифровых отсчетов амплитуды (блок отсчетов); нужно указать математическое отображение M величин x_1, x_2, \dots, x_M в одну величину y , которая и будет индексироваться на экране, то есть указать подходящую функцию $y = \varphi(x_1, x_2, \dots, x_M)$ M переменных.

Но что значит «подходящая»? На языке математической статистики, а точнее, той ее части, которая называется теорией оценок, это задача построения «хорошей» или даже «оптимальной» непараметрической *оценки* для y по выборочным данным $\{x_k\}$, $k = 1, 2, \dots, M$. Решение этой задачи сильно зависит от того, что мы знаем (априори, или из предыдущих наблюдений) о структуре

выборки $\{x_k\}$, в частности, к какому классу относится входной сигнал. Например, если входной сигнал суть стационарный случайный процесс с порождающим выборку $\{x_k\}$ белым (или почти белым) спектром мощности, то выборка $\{x_k\}$ – *случайна и однородна*; оптимальное построение таких выборок хорошо проработано в классической математической статистике. Если входной сигнал относится к тому же классу, но с небелым и приближенно наблюдаемым спектром, то задачу можно свести к предыдущему варианту путем введения весовой функции неоднородности выборки.

Довольно общий случай – входной сигнал в виде смеси стационарного или с известной нестационарностью случайного процесса (например, как проявление шумов разного рода) и детерминированного процесса с резко неоднородным спектром. Сформулировав подходящий критерий оптимальности, можно построить оптимальную оценку, максимизирующую значение этого критерия. Другими словами, теоретическое построение хорошей или оптимальной оценки требует как минимум частотно-временной структурной модели входного сигнала. Практическое применение такого подхода сводится к следующему:

а) Используя опыт и теоретические предпосылки, заранее подбираются несколько спектрально-временных моделей сигналов, могущих появиться на входе СА.

б) Для каждой модели разрабатывается оптимальное правило комбинирования отсчетов и каждое правило закладывается в математическое обеспечение СА.

в) Далее возможны два способа действий:

в1) вырабатывается система индикаторов $U_1(\{x_k\}), U_2(\{x_k\}), \dots$ (функций выборки $\{x_k\}, k = 1, 2, \dots, M$) таких, что в зависимости от областей значений, в которые они попадают при наблюдаемых $\{x_k\}$, автоматически принимается та или иная модель и реализуется соответствующее правило комбинирования;

в2) оператор предварительно обучается так, чтобы по наблюдаемым на экране (некомбинированным) отсчетам он мог, более или менее уверенно, назначать подходящую модель и, тем самым, алгоритм комбинирования.

Изложенный подход можно рассматривать как план на будущее; представим теперь подход, реализованный во многих моделях СА. В математическое обеспечение заложено несколько алгоритмов комбинирования отсчетов каждого блока; выбор того или иного алгоритма поручается интуиции тренированного оператора, предварительно наблюдающему последовательность отсчетов на некотором участке спектра.

Прокомментируем некоторые популярные алгоритмы комбинирования.

а) Метод центрального отсчета. Это простейший, интуитивно понятный метод комбинирования: в качестве представителя блока отсчетов берется отсчет, занимающий в блоке центральное по времени положение (если в блоке четное число отсчетов, можно взять или любое из двух центральных отсчетов, или среднее арифметическое от них). Точки, представляющие отдельные блоки отсчетов, можно соединить отрезками прямых или, более сложным образом, с помощью сплайнов. Часто считают [2.6], что этот метод хорош для модели (однородного) случайного процесса, но неудобен для анализа детерминированного узкополосного сигнала.

б) Метод максимального значения. В качестве представителя блока отсчетов берется наибольший из них. Метод дает относительную гарантию того, что отдельные локальные спектральные линии сигнала не останутся незамеченными при индикации; поэтому метод считается одним из самых важных и установлен по умолчанию в ряде СА; однако, считается [2.6], что этот метод недостаточно хорош для анализа случайного процесса.

в) Метод минимального значения. В качестве представителя блока отсчетов берется наименьший из них. Этот метод предусмотрен в большинстве анализаторов, хотя используется гораздо реже

метода пикового значения и метода центрального значения. Считается [2.6], что этот метод хорош в некоторых процедурах идентификации сигналов, а также для отделения гармонических сигналов от импульсных, что, например, необходимо производить в тестах на электромагнитную совместимость.

г) Метод нормального комбинирования. Чтобы объединить положительные свойства методов центрального отсчета и пикового значения, во многих СА предлагается метод нормального комбинирования, привлекающий информацию о росте или падении отсчетов в пределах блока. Если отсчеты в рамках одного блока и падал, и рос, и это повторялось в ряде других блоков, то оценки чередуются: для первого блока берется оценка максимального значения, для второго – минимального значения, для третьего – снова максимального значения и т.д. Если же на протяжении нескольких блоков отсчеты будут только расти или только падать, то метод нормального детектирования будет работать как метод максимального значения. Идея метода в том, что если встретится спектральная линия шириной в несколько блоков отсчетов, то она не будет пропущена, если же на входе СА действует случайный процесс, то его специфика будет хорошо отображена. Применение этого метода может привести к сдвигу максимума спектральной кривой на экране вправо на одну точку; в некоторых моделях СА этот эффект компенсируется, в других – игнорируется.

д) Метод среднего. Для тех, кто имеет опыт использования терминологии, методов и рекомендаций математической статистики, наиболее естественной оценкой отображаемого значения $y = \varphi(x_1, x_2, \dots, x_M)$ по выборке $\{x_k\}$ является среднее арифметическое выборочных значений:

$$y = \frac{1}{M} \cdot \sum_{k=1}^M x_k .$$

Если выборка $\{x_k\}$ чисто случайна (выборочные значения – реализации независимых случайных величин) и стационарна, то такая оценка эффективна в смысле Фишера, то есть состоятельна, несмещенна и имеет минимальную дисперсию. В то же время, например, оценки центрального отсчета, максимального отсчета и минимального отсчета имеют большую дисперсию, чем оценка среднего, что объясняется просто: первая вообще не учитывает остальные отсчеты, вторая и третья – учитывают их слабо.

Приведенная оценка усредняет значения отсчетов спектральных амплитуд. Логически к ней примыкает оценка, усредняющая квадраты амплитуд, то есть значения отсчетов спектральной плотности мощности:

$$y = \sqrt{\frac{1}{M} \sum_{k=1}^M x_k^2} ,$$

и оценка, усредняющая логарифмы амплитуд:

$$y = 10^W, \quad W = \frac{1}{M} \sum_{k=1}^M \lg x_k ,$$

которые могут показаться более естественными при интерпретации результатов анализа в терминах спектральной плотности мощности или децибелов. При аналоговой обработке методу среднего соответствует пропускание видеосигнала через фильтр нижних частот или интегратор.

Фильтрация в ПЧ и сглаживание индицируемого спектра

Даже в целом детерминированный входной сигнал имеет некоторую (аддитивную) случайную составляющую, относительный уровень которой можно оценить параметром, называемым *отношением сигнал-шум*. По мере прохождения сигнала по тракту к его случайной составляющей прибавляются собственные (внутренние) шумы аппаратуры и отношение сигнал-шум уменьшается; это уменьшение оценивается параметром, называемым *коэффициентом шума*. В результате, на выходе последнего смесителя СА мощность сигнала есть сумма мощностей де-

терминированной части и случайной части. Каждая из этих частей имеет свой спектр мощности. Обычно спектр детерминированной части локализован в некоторой конечной полосе, в то время как спектр случайной части размыт в очень широкой полосе, ограниченной только фильтрующим действием преселектора и предыдущих каскадов преобразования частоты. Поскольку мы рассматриваем СА последовательного анализа с дискретной частотной перестройкой, то при установленной полосе $\Delta f_{\text{посл ПЧ}}$ фильтра последнего ПЧ спектр детерминированной части не искажается, а лишь просматривается с некоторым разрешением, в то время как мощность случайной части уменьшается в тем большей степени, чем уже полоса $\Delta f_{\text{посл ПЧ}}$. В результате, отношение сигнал-шум на выходе фильтра много больше отношения на выходе последнего смесителя. Этот эффект есть ничто иное, как *усреднение*, он обязан уменьшению относительной дисперсии при сложении некоррелированных случайных значений и может быть строго математически описан.

Из сказанного не следует, что нужно во всех случаях устанавливать как можно меньшее значение $\Delta f_{\text{посл ПЧ}}$, получая выигрыш в отношении сигнал-шум и в разрешающей способности.

Дело в том, что с уменьшением $\Delta f_{\text{посл ПЧ}}$ резко увеличивается время частотной развертки, а это не только снижает оперативность измерений, но и надолго разносит по времени наблюдение разных участков спектра сигнала; последнее важно, если сигнал имеет конечное время существования или/и нестационарный спектр, меняющийся с некоторой скоростью. Действительно, пусть выбраны анализируемый интервал частот (установкой значений $f_{\text{min}}, f_{\text{max}}$) и величина $\Delta f_{\text{посл ПЧ}}$. Для обеспечения нужной точности спектрального анализа необходимо, чтобы время «стояния» последнего гетеродина на одной частотной точке $t_{\text{ст}} = K / \Delta f_{\text{посл ПЧ}}$, $K \gg 1$ было достаточно большим, а число частотных точек в развертке будет $(f_{\text{max}} - f_{\text{min}}) / \Delta f_{\text{посл ПЧ}} + 1$; в результате время частотной развертки обратно пропорционально квадрату $\Delta f_{\text{посл ПЧ}}$ и может оказаться недопустимо большим.

Рассмотрим пример. Пусть частотный диапазон анализа СА составляет от 10 МГц до 10 ГГц (\square 10 ГГц); выберем для анализа всего 1% диапазона: $f_{\text{max}} - f_{\text{min}} = 100$ МГц. Установим $\Delta f_{\text{посл ПЧ}} = 10$ Гц, примем $K = 10$, тогда $t_{\text{ст}} = 1$ с. Время развертки составит 10^8 с, или 2780 часов (!), что абсолютно неприемлемо. Если же при тех же параметрах установить $\Delta f_{\text{посл ПЧ}} = 10$ кГц, то время развертки составит 10 с.

Обратимся теперь к манипуляциям с цифровыми отсчетами амплитуды, получаемыми после детектирования (аналогового или цифрового). Приведенное выше среднее

$$y_s = \frac{1}{M} \cdot \sum_{k=1}^M x_{sk}$$

от отсчетов $\{x_{sk}\}$ s -го блока есть вид *сглаживания* видеоданных; если же среднее захватывает несколько (S) блоков, то получается более сильное сглаживание, чем при комбинировании отсчетов блока:

$$Y = \frac{1}{S} \sum_{s=1}^S y_s = \frac{1}{M \cdot S} \sum_{s=1}^S \sum_{k=1}^M x_{sk}.$$

Природа сглаживания отличается от природы фильтрации (усреднения) в тракте ПЧ и состоит в следующем. Спектр как функция частоты, который отображался бы при отсутствии сглаживания (при условии соединения точек на экране отрезками прямых) имеет колебания разных частотных масштабов; если частотный аргумент заменить на временной, то этот исходный спектр станет функцией времени, от которой можно взять преобразование Фурье. В результате получится функция частоты, которую можно назвать «спектром спектра» или «вторым спектром». Осуществляя сглаживание, мы усредняем (убираем) наиболее высокочастотные колебания второго спектра. Чем больше параметр SM сглаживания, тем левее (размытая) граница уда-

ленных частот второго спектра; в результате отображаемая кривая (первого) спектра становится все глаже, откуда и название операции. Подытоживая, можно сказать, что усиливая фильтрацию в тракте ПЧ (обуждая фильтр ПЧ) мы отслеживаем спектр сигнала все с большим разрешением (расплачиваясь, правда, за это увеличением длительности развертки), а усиливая сглаживание (увеличивая параметр SM), мы делаем кривую индицируемого спектра более гладкой, расплачиваясь обеднением «второго спектра» за счет его высоких частот.

Процесс сглаживания дает повод ввести еще одно понятие спектрального разрешения (разрешающей способности), совершенно отличное от ранее рассмотренного классического. Если последнее определялось полосой разрешения, то есть спектральным расстоянием между двумя спектральными линиями, при котором они разрешены, то новое разрешение определяется числом отсчетов АЦП, идентифицируемое как одна точка на экране, или соответствующая этому числу полоса. Этот вид разрешения можно назвать *разрешением индикации*, а соответствующую полосу – *полосой разрешения индикации*. Чем больше эта полоса, тем хуже разрешение индикации, но более гладкая, «красивая» отображаемая кривая спектра.

Межкадровое усреднение

Цифровое управление дисплеем (цифровой дисплей) дает возможность осуществления другого вида сглаживания, называемого *межкадровым усреднением*. Здесь усреднение проводится поточно, за две или более частотные развертки. Удобно находить межкадровое среднее рекуррентно, по мере прохождения все новых разверток, по схеме:

$$\tilde{x}_n = \frac{n-1}{n} \cdot \tilde{x}_{n-1} + \frac{1}{n} x_n,$$

где $\tilde{x}_{n-1}, \tilde{x}_n$ – средние после $n-1$ -ой и после n -ой разверток соответственно;

x_n – отсчет на n -ой развертки.

Такой алгоритм придает процессу наглядность: картина на дисплее в каждой частотной точке постепенно сходится к среднему за установленное число проходов. Эта установка задает степень усреднения, так же как при обычном видео-усреднении, которое теперь можно назвать *трассовым*, степень усреднения задается параметром SM .

Межкадровое усреднение, также как и трассовое, сглаживает индицируемую кривую спектра, расплачиваясь примерно таким же временем. Однако результаты межкадрового и трассового усреднений могут быть разными, например, в модели сигнала с нестационарным (меняющимся во времени) спектром. Поэтому важно, что эти два вида усреднения можно совместить в одном эксперименте. Установим число M отсчетов в блоке, число S блоков в трассовом усреднении, число Q разверток в межкадровом усреднении. Тогда индицируемое значение в одной частотной точке равно

$$Y = \frac{1}{W S M} \sum_{n=1}^Q \sum_{s=1}^S \sum_{k=1}^M x_{nsk}.$$

Оптимальный выбор величин M, S, Q позволяет адекватно учесть особенности модели сигнала.

Временное стробирование

Как указывалось выше, недостатком представления сигналов в частотной области на базе преобразования Фурье является отсутствие временной локализации в последнем. В то же время, сам метод последовательного спектрального анализа приводит к тому, что различные участки индицируемого спектра относятся к различным временным интервалам существования сигнала, но это никак не проявляется при наблюдении кривой спектра на экране. Кроме того, в практике радиотехнических систем СВЧ, особенно в беспроводных системах связи, часто используются сигналы, занимающие перекрывающиеся, или даже совпадающие области частотного спектра, но раз-

деленные во временной области. Перечисленные факторы подталкивают к идее осуществлять, помимо спектрального анализа, еще и управление наблюдением сигналов по времени. Одним из простейших механизмов такого управления является *временное стробирование*. Оно использует внешний управляющий сигнал, активизирующий те или иные функции СА на определенных интервалах времени (стробах), и закрывающий эти функции на других на других интервалах времени (антистробах).

Временное стробирование позволяет раздельно наблюдать спектры сигналов, разделенных по времени. Эта функция полезна, например, при наблюдении неоднородной радио-импульсной последовательности, сигналов с временным мультиплексированием, сигналов в системе много-станционного доступа с временным разделением каналов (TDMA), перемежающихся или уплотненных импульсных сигналов, модулируемых всплеском сигналов.

Выполняемые при этом операции можно условно разделить на следующие:

- измерение спектральных характеристик разделенных по времени сигналов;
- измерение спектра сигнала в одной временной ячейке системы TDMA;
- отделение и наблюдение спектра неинформативных временных частей сигналов, например, переходных процессов периодической последовательности импульсов.

Из всех возможных объектов стробирования СА наиболее часто используются два:

- выход видеодетектора; сигнал отключается в антистробах от дальнейшего тракта, применяется пиковый режим комбинирования, время развертки устанавливается так, чтобы строб действовал хотя бы один раз на каждую отображаемую точку дисплея;
- последний гетеродин; во время антистроба гетеродин не переключается на новые частоты и дисплей отображает одну точку спектра.

Преимущества цифровой обработки

Начиная с 80-х годов прошлого века СА испытали значительные изменения, в основном за счет быстрого наступления цифровой обработки сигналов. Это стало возможным благодаря прогрессу в высокопроизводительных АЦП. Цифровые технологии обработки сигналов позволили улучшить все показатели эффективности СА: скорость, точность, разрешение и т.д. Наиболее сильные улучшения стали результатом освоения полностью цифровой секции ПЧ. Цифровая секция последней ПЧ обеспечивает целый букет преимуществ для пользователя. Ключевым преимуществом, доставляемым цифровыми разрешающими фильтрами последней ПЧ, является резко возросшая разрешающая способность СА, обязанная недостижимой для аналоговых фильтров полосовой избирательности цифровых фильтров, например, 4:1 на уровнях -60 дБ, -3 дБ. Наилучшие измерения возможны при комбинации БПФ-анализа (при отсутствии частотной развертки) – для узких полос обзора, и последовательного анализа с частотной разверткой – для широких полос обзора. Дальнейшие преимущества высокой надежности, стереотипности исполнения, удешевления реализации и др. принесли разработка, производство и применение двух интегральных схем: ИС преобразований от вычислений квадратур I, Q до преддисплейной комбинаторики, и ИС RISC-процессора («интегральная микросхема с сокращенным набором»), осуществляющая дальнейшие преобразования до дисплея.

В современных СА с высоким весом цифровой обработки имеется возможность повысить точность и разрешение установки частоты, и без того очень высокой. Для этого в тракте последнего ПЧ предусмотрен счетчик частоты. Он подсчитывает количество пересечений нуля сигналом в этом тракте за конечный интервал времени, например, за 0.1 с, и пересчитывает результат в текущее отклонение от номинальной частоты. Само это отклонение обусловлено двумя составляющими: флуктуациями частоты гетеродинов за счет фазовых шумов и флуктуациями мгновенной частоты сигнала за счет его зашумленности. Поскольку гетеродины СА выполнены как высоко частотно-стабильные синтезаторы (особенно первый гетеродин, стабилизированный кварцевым генератором), первую составляющую частотного отклонения можно сделать достаточно

малой, и на первый план выходит составляющая, обусловленная зашумленностью анализируемого сигнала. Таким образом, появляется дополнительная функция СА – измерение частоты. Благодаря фазовым ухищрениям, дополняющим счет пересечений нуля, можно достичь разрешения частоты в десятки миллигерц на периоде 0.1 с.

Подытожим достижения СА, ставшие возможными благодаря высокой степени «цифровизации»:

- возможно усреднение с регулируемой постоянной времени как напряжения, так и мощности и логарифма мощности;
- возможно трассовое, межкадровое, а также смешанное сглаживание;
- возможен счет частоты с высоким разрешением;
- хранящиеся в памяти данные можно извлекать как в линейном, так и в логарифмическом масштабе;
- превосходный коэффициент прямоугольности цифровых фильтров обеспечивает уникальное частотное разрешение анализа, почти идеальное для последовательного анализа;
- коэффициент K , связывающий необходимое время «стояния» $t_{ст}$ на одной частоте последнего гетеродина с полосой разрешения $\Delta f_{\text{полс ПЧ}}$ ($t_{ст} = K / \Delta f_{\text{полс ПЧ}}$) в случае цифрового исполнения разрешающего фильтра может быть значительно меньше, чем при аналоговом исполнении;
- возможность легкого исполнения комбинирования по методу среднего;
- цифровое исполнение всех разрешающих фильтров последней ПЧ позволяет создать уникальный набор устанавливаемых полос разрешения (например, 160 различных полос), удовлетворяющий потребности любого пользователя;
- возможно включить, в зависимости от полосы обзора, режим последовательного анализа или режим БПФ;
- присутствующий на схеме рис. 2.3.1 логарифмический усилитель до сих пор никак не комментировался; при полностью цифровом тракте последней ПЧ и этот усилитель – также цифровой; его точность (точность выполнения им функции логарифмирования) значительно выше, чем при аналоговом исполнении;
- точность установки каждой полосы разрешения намного выше, чем при аналоговых разрешающих фильтрах; за счет этого, в частности, выше точность измерения спектральной амплитуды, спектральной плотности мощности и мощности в канале;
- разброс пропусков в максимуме разрешающих цифровых фильтров меньше, чем при аналоговом исполнении; это в частности, облегчает калибровку прибора;
- при отсутствии в тракте последнего ПЧ аналоговых звеньев усиления, зависящих от опорного уровня, не существует неопределенности «усиления тракта последней ПЧ»;
- возможно изменять настройки СА без существенного влияния на точность измерения;
- итоговое влияние отмеченных улучшений таково, что СА с полностью цифровым трактом последней ПЧ имеет значительно лучшую точность спектрального анализа, чем СА с аналоговым трактом.

Эффективность спектроанализатора

Как и для анализаторов цепей, для СА пока не сформулирован комплексный критерий эффективности. На этом этапе изложения мы не будем углубляться в предложенные частные критерии эффективности, а перечислим и поясним некоторые количественные факторы, отражающие свойства АЦ и входящие в комплексный и некоторые частные критерии.

1. Точностные характеристики СА. Выше уже было дано общее понятие о комплексе параметров и характеристик, группирующихся вокруг понятия «точность измерения» и характеризующих степень приближения результатов измерений к истинным значениям измеряемых величин: погрешность измерения (погрешность прибора) и факторы, ее обуславливающие; разновидности числового представления погрешности: абсолютные, относительные, приведенные; систематические, случайные погрешности и промахи, и факторы, их обуславливающие, а также понятия измерительной и производственной калибровке. Там же указано на роль теории оценок как раздела математической статистики. Также как при анализе цепей, при анализе сигналов в общем случае объектами измерения являются комплекснозначные функции частоты и времени, например, (комплексный) спектр сигнала или (действительная) частота сигнала в тракте последнего ПЧ. Было показано, почему абсолютная и относительная комплексные ошибки не годятся для практического использования. Поэтому и в случае спектрального анализа сигналов мы приходим к двум системам измеряемых величин и их ошибок; первая система: амплитудные спектры и спектры мощности, то есть действительные неотрицательные функции частоты, а также они же в децибельной мере, и абсолютные и относительные ошибки этих функций; вторая система: фазовые спектры и их абсолютные ошибки. Существуют два популярных подхода к интерпретации случайных ошибок амплитудного и фазового спектров: в виде доверительного интервала при фиксированной доверительной вероятности и в виде среднеквадратичной ошибки; и два популярных подхода к интерпретации систематических ошибок амплитудного и фазового спектров: в виде интервала неопределенности между возможными минимальным и максимальным значениями и в рамках квазистатистического подхода.

Не вдаваясь в собственно анализ точностных характеристик тех или иных измеряемых величин СА, назовем факторы неопределенности, порождающие парциальные неопределенности, то есть вклады в общую неопределенность измерения. Основные факторы:

- рассогласования СВЧ-секций: выхода источника сигнала и входа СА; выхода радиочастотного тракта и входа смесителя и т.д.;
- неопределенность, вносимая радиочастотным аттенуатором;
- неопределенность, вносимая частотной неравномерностью входного фильтра и преселектора;
- неопределенность, вносимая частотной неравномерностью смесителей и гетеродинов;
- неопределенность опорного уровня за счет неопределенности усиления трактов;
- неопределенность, вносимая разрешающими фильтрами;
- неопределенность масштаба отображения экрана;
- остаточная неопределенность после измерительной калибровки.

Неопределенность, вносимая рассогласованиями, возникает из-за переотражений от стыков разных импедансов; адекватным аппаратом для учета возникающих при этом парциальных волн и, соответственно, вычисления возникающей ошибки, являются *потокосигнальные графы*.

Неопределенность, вносимая радиочастотным аттенуатором, состоит из двух частей: неопределенности ослаблений при различных положениях аттенуатора на разных частотах и вклада в общую частотную характеристику СА, особенно в ее высокочастотной части.

Неопределенность, вносимая частотной неравномерностью входного фильтра и преселектора различна на разных частотах. В нижнем поддиапазоне СА используется фиксированный ФНЧ, в верхнем – перестраиваемый полосовой фильтр, преселектор. У ФНЧ неравномерность АЧХ мала и близка к расчетной, поэтому вносимая им неопределенность также мала. Преселектор, обычно выполненный на управляемом ЖИГ-резонаторе, имеет значительные неравномерности АЧХ и потому вносит значительную неопределенность. Эта неопределенность может быть существенно снижена с помощью внешних достаточно точных генератора монохроматического сиг-

нала и измерителя мощности: Последовательно, на нескольких интересующих нас частотах сигнал подается на СА и измеритель мощности и их показания сравниваются.

Неопределенность, вносимая частотной неравномерностью смесителей и гетеродинов – объект сложного анализа, из-за использования *гармонического смешения*, следующего отсюда разбиения диапазона СА на *гармонические полосы* и т.д.

Неопределенность опорного уровня за счет неопределенности усиления трактов обязана, в основном, тракту последней ПЧ. Если этот тракт выполнен как аналоговый, то его усилитель и аттенюатор настроены так, чтобы скомпенсировать изменения в установках радиочастотного аттенюатора и потери преобразования последнего смесителя, так что амплитуды входного сигнала соотносятся с верхней отметкой координатной сетки дисплея, то есть опорным уровнем. Усилитель и аттенюатор последней ПЧ работают на одной частоте, поэтому не вносят вклада в неопределенность частотной характеристики, однако дают вклад в амплитудную неопределенность, обусловленный неточностью итоговой настройки, который и называется *неопределенностью опорного уровня*. Однако, при полностью цифровой последней ПЧ аналогового усилителя в этом тракте нет, поэтому неопределенность опорного уровня отсутствует.

– Неопределенность, вносимая разрешающими фильтрами, обусловлена тем, что у аналоговых разрешающих фильтров с разной полосой вносимые потери различны, особенно если это фильтры разного типа, например, LC-фильтры, применяемые при широких полосах разрешения, и кварцевые фильтры, применяемые при узких полосах разрешения. В результате образуется *неопределенность переключения полос разрешения*. Однако, в полностью цифровом тракте последней ПЧ условия возникновения названной неопределенности другие. Во-первых, разрешающие фильтры для всех полос разрешения исполнены как цифровые, часто в рамках специализированной интегральной схемы. Хотя эти фильтры не идеальны, они отличаются высокой повторяемостью, вдобавок, применяется некоторый алгоритм компенсации, уменьшающий ошибку повторяемости. Поэтому вклад цифровых фильтров в неопределенность переключения полос разрешения достаточно мал. Во-вторых, кроме аналогового фильтра, препятствующего наложению спектров, иногда [2.6] перед оцифровкой в тракте последней ПЧ ставят аналоговый *предварительный фильтр* с переключаемой полосой, в 2.5 раз шире соответствующей полосы цифрового фильтра разрешения. У этого фильтра при каждой включенной его полосе имеются неопределенности центральной частоты, полосы пропускания и коэффициента ослабления, которые и определяют вклад в неопределенность переключения полос пропускания.

– Неопределенность масштаба отображения экрана в аналоговом тракте последней ПЧ возникает из-за того, что усиление усилителя тракта в логарифмическом масштабе амплитуды не идеально аппроксимирует логарифмическую кривую, а в линейном масштабе – не идеально линейна. В обоих режимах любое отклонение от идеальной характеристики дает вклад в неопределенность амплитуды (мощности). Этот тип неопределенности называется *достоверностью масштаба отображения*. Цифровой тракт последней ПЧ не включает аналоговый усилитель; функция логарифмирования (как и линейная функция) осуществляется математически почти идеально (то есть с пренебрежимой ошибкой); поэтому в традиционном смысле неопределенности масштаба отображения в этом случае нет. Однако, сама эта неопределенность имеет место по другим причинам. Во-первых, в отсутствие усилителя последней ПЧ на поверхность выходит маленькая неопределенность амплитуды за счет сжатия усиления в радиочастотной секции, заметная уже при мощности входного сигнала СА выше -20 дБм. Во-вторых, еще одним источником неопределенности амплитуды является неточность реализованного расположения диапазона усиления АЦП по отношению к модельному расположению. В-третьих, вклад в неопределенность масштаба отображения дает специфическая для цифровой техники ошибка квантования, обусловленная тем, что оцифровка – это процесс дискретизации аналогового сигнала по времени и квантования по уровню.

Заметим, что в данном случае нас интересует вызываемая шумом квантования нелинейность средней характеристики АЦП. Поэтому вклад в рассматриваемую неопределенность можно уменьшить классическим приемом, представляющим собой вид *рандомизации* данных: добав-

лением шумовой (некоррелированной) цифровой последовательности (*подмешивающий сигнал* или дифер, dither). Это подмешивание изменяет структуру шума квантования: сглаживает среднюю характеристику АЦП, делая ее более линейной, но, естественно, увеличивает средний уровень шума квантования. Если до подмешивания отношение (сигнал/шум квантования) достаточно велико (например, больше 10 дБ), то можно считать, что выигрыш в линейности средней характеристики АЦП более важен, чем проигрыш в точности единичного отсчета сигнала (то есть до усреднения); если же это отношение недостаточно велико, то названный проигрыш не окупается выигрышем. Поэтому производители СА с подмешиванием оставляют пользователю возможность выключить его при значительной зашумленности оцифрованного сигнала до усреднения.

– Остаточная неопределенность после измерительной калибровки это многокомпонентная неопределенность, оставшаяся после измерительной калибровки из-за неидеальности последней. Все СА используют встроенный или внешний прецизионный генератор, обеспечивающий монохроматический сигнал заданной частоты, используемый как опорный сигнал для абсолютной калибровки СА. В процессе последней СА переносит абсолютную калибровку опорного сигнала на другие частоты и амплитуды. Из-за неточностей амплитуды и частоты опорного сигнала, в частности, обусловленных частотной (фазовой) нестабильностью его генератора, образуется неопределенность измерения этих величин СА, обычно малая на частоте опорного сигнала и увеличивающаяся при отклонении от этой частоты. Например, во многих СА фирмы Аджилент (США) используется встроенный калибратор на частоте 50 МГц и на этой частоте абсолютная ошибка измерения амплитуды спектроанализатором, заявленная в спецификациях для PSA-серии СА: ± 0.24 дБ, но значительно больше на частотах в несколько ГГц. Если у СА есть встроенный калибратор, то ничто не мешает при случае использовать более точный или более близкий к интересующим нас частотам внешний калибратор.

– Точности измерения спектральной амплитуды (спектральной плотности мощности) и мощности сигнала. Представленные абсолютными и относительными, систематическими (в форме максимальных значений или квазислучайных параметров) и случайными ошибками, эти точности зависят, в общем случае, от установок СА (разрешающая полоса, степень усреднения, степень сглаживания и т.д.), от истинных значений измеряемых величин, формы их представления (линейная, логарифмическая), от степени зашумленности сигнала. Представляющие ошибки аккумулируют вклады всех неопределенностей (перечисленных выше и других). Отдельная точностная характеристика – амплитудная (мощностная) ошибка маркера, устанавливаемого на кривую спектра для получения информации об абсолютных значениях спектральной амплитуды и частоты. Эта точность, кроме прочего, зависит от точности наведения маркера на кривую, в свою очередь, наибольшей при наведении на экстремум кривой.

– Точность измерения частоты. Формально, судя по названию прибора – СА, измерение частоты – побочная (хотя и важная) функция прибора. С помощью СА можно проводить два вида измерений частоты: абсолютное и относительное. Абсолютные измерения проводятся с целью определить или проверить частоту конкретного сигнала, а также для анализа нежелательных сигналов. Например, для проверки, что радиостанция работает на предписанной ей частоте, или при исследовании некоторой паразитной помехи. Относительные измерения частоты полезны для выяснения частотного расстояния между двумя конкретными сигналами, для определения частоты модуляции и т.д.

Переход гетеродинов СА на высокостабильные синтезаторы частоты (в основном – косвенного синтеза, то есть с использованием опорных кварцевых генераторов), произошедший еще в 80-е годы 20-го века, резко повысил точность ее измерения. Дальнейший существенный прогресс в точности был связан с применением усовершенствованных счетчиков пересечений уровня. Большое значение для точности измерения частоты имеет короткопериодная (фазовый шум) и длиннопериодная (дрейф и температурные уходы) стабильности частоты гетеродинов, наличие или отсутствие встроенного опорного генератора частоты, период между частотными калибровками и подстройками, наличие или отсутствие трассируемого лабораторного стандарта частоты, время непрерывной работы и т.д. Точность относительных измерений частоты зависит, в частно-

сти, от *точности полосы обзора*, определяемой неопределенностью отображения на дисплее частотных расстояний между спектральными компонентами.

Частотная точность маркера зависит от частотной калибровки дисплея, выбранного количества дисплейных точек, усреднений и сглаживаний, точности наведения маркера на кривую. Последняя, как и амплитудная точность, максимальна при наведении на экстремум. Чтобы добиться хорошего результата, полезно сузить полосу обзора и полосу разрешения и тем самым визуально сузить отображаемый пик кривой, на который наводится маркер.

У многих СА есть специальный *режим маркера*, в котором действуют схемы внутренних счетчиков, предназначенных для устранения влияния полосы обзора и полосы разрешения на частотную точность маркера. Важно, что в этом случае счетчик считает не пересечения входного сигнала СА, а пересечения сигнала в тракте последней ПЧ. Кроме того, чтобы отклонения частот гетеродинов от номинальных не увеличивали ошибку измерения частоты, счетчики в режиме маркера считают и пересечения гетеродинных сигналов. Процессор обрабатывает эту информацию и вычисляет частоту входного сигнала. Чтобы в режиме маркера устранить шум как фактор ошибки, стараются, по крайней мере в этом режиме, добиться достаточно большого отношения сигнал-шум, или хотя бы наводить маркер на точки, достаточно далекие от индицируемого уровня шума. При счете пересечений простым счетчиком неустранимо влияние на точность измерения частоты параметра, называемого *разрешением счетчика* и задаваемого наименьшим значимым разрядом в представлении результата счета. Влияние разрешения счета удваивается при пользовании *дельта-маркерами*, то есть парами маркеров для визуализации частотных разностей.

2. Статические диапазоны измерений. Зададимся вопросом: в каком диапазоне может быть измерена величина, определяемая СА, в «медленном» режиме, то есть при последовательном изменении установок, например, разрешающих полос фильтра последней ПЧ, положений входного аттенюатора и т.п. Такой диапазон можно назвать *статическим* (иногда [2.6] его называют *диапазоном измерения*). Ответы на этот вопрос не определены, если не принять, что диапазон измерений величины это область значений измеряемой величины, в пределах которой нормированы погрешности средства измерений. То есть для каждой измеряемой величины мы задаемся некоторой максимальной допустимой ошибкой и находим диапазон измерения этой величины, в пределах которого ошибка измерения не больше заданной.

С помощью скалярного СА можно измерять спектральную амплитуду (спектральную плотность мощности) как функцию частоты, мощность всего сигнала (среднюю мощность случайного процесса как сигнала), мощность сигнала в спектральном канале, «мгновенную» частоту сигнала (здесь мы, как и прежде, не касаемся таких дополнительных функций СА, как измерения коэффициента шума цепи, фазовых шумов, быстрых и медленных частотных (фазовых) нестабильностей генераторов и т.п.).

а) Статические диапазоны измерения мощности, средней мощности, спектральной амплитуды (спектральной плотности мощности) при заданных предельных погрешностях этих величин ограничены сверху сжатием первого смесителя и сжатием усиления усилителей при минимальном ослаблении управляемых аттенюаторов, а снизу – уровнем шума при минимальной доступной в данном измерении полосе разрешающего фильтра ПЧ (не всегда равной полосе самого узкого фильтра ПЧ).

б) Статический диапазон измерения частоты сигнала сильно зависит от того, применяется ли техника счета пересечений нуля, какова допустимая ошибка измерения частоты (в «досчетчиковую» эру допустимая ошибка в гигагерцовом диапазоне составляла несколько МГц!), каков временной интервал измерения и т.д. В итоге, статический диапазон измерения частоты не обязательно равен частотному диапазону СА (см. ниже).

Нужно отличать статический диапазон измерения величины от динамического диапазона.

3. Динамический диапазон амплитудных измерений. Несмотря на лексическую схожесть терминов, динамический диапазон принципиально отличается от статического по смыслу и по

назначению. Если статический диапазон есть просто диапазон измеряемых величин с использованием возможностей перестроек параметров СА, то динамический диапазон понимается как возможность одновременно измерять два или более сигнала разной величины. Точнее, величина динамического диапазона это отношение, выраженное в дБ, наибольшего и наименьшего сигналов, которые могут одновременно измеряться СА с заданной степенью (относительной) точности (упоминание точности почти всегда относится к меньшему сигналу, т.к. больший сигнал измеряется с большей относительной точностью).

Ситуация с одновременным измерением нескольких сигналов разной величины возникает во многих случаях. Например, при наблюдении двух или более тестовых спектральных компонент, при измерении гармонических искажений или интермодуляций и т.д. Динамический диапазон, в отличие от статического, относится к измерениям при фиксированном наборе установок СА, без их изменения. Динамический диапазон не единствен: он зависит почти от каждой из установок и поэтому различен при разных их наборах. В соответствии с общим принципом, чтобы сделать понятие диапазона определенным, необходимо задать максимальную допустимую ошибку измерения соответствующей величины. Поскольку динамический диапазон современных СА очень велик (например, 100 или более дБ) допустимо для удобства анализировать его величину как разность (в дБ) максимально достижимой верхней границы и минимально достижимой нижней границы (хотя разработаны графические методы одновременного анализа всего диапазона).

Во всех источниках и руководствах по СА указано, что снизу динамический диапазон амплитудных измерений ограничен а) шумами, б) нелинейными искажениями, вносимыми смесителем (в основном гармоническими искажениями второго и третьего порядков и интермодуляциями третьего порядка). Но математическая трактовка этого тезиса в части а) зачастую утрируется, размыта неопределенными понятиями, осложнена ненужными и даже неверными предположениями и т.д. Например, предлагается в качестве нижней границы диапазона брать индицируемый в области предполагаемого отсутствия сигнала уровень шума (который, кстати, обозначают неопределенными терминами типа «средний шум» и т.п.); или предлагается считать нижней границей уровень, отстоящий на определенное число дБ от этого «уровня шума», называя это число «отношением сигнал-шум», не связывая выбор этого отношения с предположениями о статистической структуре шума и классе сигналов; или, наоборот, делая мало обоснованные предположения о распределении вероятностей «уровня шума», связывают «отношение сигнал-шум» с относительной ошибкой измерения и считают, что нижняя граница диапазона обусловлена максимальной допустимой ошибкой измерения и т.д. и т.п.

Конечно, эти нечеткости, неопределенности, нестрогости трактовки не имеют большого значения, если нас удовлетворит знание динамического диапазона с точностью до нескольких дБ, например, нам подходит ответ: «динамический диапазон данного СА при таких-то установках – порядка 100-106 дБ». Но если динамический диапазон – критический параметр для данного применения СА, если от этого вопроса зависит приобретение данного СА, если «идет борьба за каждый дБ», то необходима более определенная, более четкая, более строгая трактовка этого вопроса на базе теории случайных процессов. Здесь мы ограничимся тезисами качественного характера о части а) нижней границы динамического диапазона:

– из внешних соображений должна быть задана максимальная допустимая ошибка измерения на нижней границе диапазона;

– должны быть четко ограничены классы потенциальных сигналов и помех (флуктуаций, шумов, в том числе фазовых шумов и т.д.), то есть конкретизирована математическая модель сигналов и помех;

– должно быть сформулировано строгое количественное понятие отношения сигнал-шум, пригодное в рамках данной модели сигналов и помех;

– это отношение, с одной стороны, должно быть связано с ошибкой измерения, с другой – с устанавливаемыми параметрами СА;

– в результате, нижняя граница диапазона в части а) должна быть представлена как функция установок СА и структур сигналов и помех.

Что касается части б) тезиса, то традиционно ставится и решается задача определения оптимального режима первого смесителя, в частности, оптимально уровня входного сигнала смесителя; в качестве условия оптимальности берется равенство неопределенности за счет шума и ошибки за счет нелинейных искажений, или другое условие.

Верхняя граница динамического диапазона ограничена допустимыми искажениями всякого рода (сжатие усилений, сжатие смесителей и т.д.), перегрузками и т.п.

Существует еще одно практически полезное диапазонное понятие, которое часто путают с динамическим диапазоном: *дисплейный диапазон*. Это калиброванный амплитудный диапазон дисплея СА. Например, дисплей с десятью делениями с калибровкой 10 дБ на деление, очевидно, имеет дисплейный диапазон 100 дБ. Сказанное в полной мере относится к СА с цифровой секцией последней ПЧ, или с аналоговой секцией – при использовании узких разрешающих фильтров, которые делаются цифровыми. Однако дисплеи СА с полностью аналоговой секцией последней ПЧ обычно калибруются лишь на 85-90 дБ вниз от опорного уровня. В этом случае нижняя часть дисплея обозначает область диапазона от -85 или -90 дБ до $-\infty$ дБ относительно опорного уровня. Возникают вопросы: а) можно ли в полной мере использовать дисплейный диапазон; б) может ли динамический диапазон превзойти дисплейный диапазон. На первый вопрос ответ положительный всегда, на второй – положительный для некоторых моделей СА. В последнем случае путем увеличения усиления тракта последней ПЧ малые сигналы переводятся на калиброванную область дисплея, а большие – выше опорного уровня. Хорошо, если при этом точность отображения сигналов не ухудшается, как это имеет место для СА серии PSA фирмы Аджилент.

4. Чувствительность СА. В технике вообще (не только радиотехнической) термин «чувствительность» используется в разных смыслах. Мы можем различать три понятия, связанных с этим термином.

а) Преобразовательная чувствительность. Так предлагается назвать параметр, который обычно называется просто *чувствительностью*, то есть способность реагировать на изменения входного сигнала; она оценивается отношением изменения выходного сигнала к вызвавшему его изменению входного сигнала. Например, чувствительность отклонения луча осциллографа измеряется в миллиметрах на вольт.

б) Обнаружительная чувствительность СА. Это способность обнаруживать с достаточной уверенностью (например, с достаточной вероятностью при заданной вероятности «ложной тревоги») малые сигналы или малые изменения спектра зафиксированного сигнала. Возможна постановка и решение этой задачи на базе статистической теории обнаружения как частной задачи испытания статистических гипотез.

в) Пороговая чувствительность СА: оценивается минимальной абсолютной мощностью входного сигнала СА, спектр которого, при заданном спектральном разрешении, может быть измерен СА с заданной степенью точности. По-существу, эту чувствительность можно понимать как пересчитанную на входную абсолютную мощность нижнюю границу динамического диапазона. Пороговая чувствительность зависит не только от установок СА, но и характера (класса) сигнала и структуры внутреннего шума СА.

5. Частотный диапазон СА может быть представлен в двух вариантах:

а) как общий диапазон, внутри которого могут размещаться поддиапазоны спектрального анализа до какой-то ширины каждый; например, некий СА имеет общий диапазон от \square 0 до 50 ГГц, внутри которого может быть произвольно размещен поддиапазон анализа шириной до 1 ГГц; такой частотный диапазон можно назвать *статическим*;

б) как максимальный поддиапазон спектрального анализа; такой частотный диапазон можно назвать *максимальной глубиной частотной панорамы*.

Нужно иметь в виду, что хотя подавляющее большинство СА сконструированы как универсальные приборы, с весьма широким классом анализируемых сигналов и неограниченным спектром приложений, в настоящее время может быть вполне оправданным создание более дешевого, быстродействующего, удобного, мобильного и т.д. специализированного СА с узким классом сигналов и/или узким спектром приложений; например, для изучения (поверки) структуры сигналов беспроводной связи, или для обнаружения газовых выделений по их спектрам излучения. Поэтому вариант а) вполне возможен.

6. Диапазон и дробность разрешающих способностей СА. Как показано выше, разрешающая способность СА задается полосой разрешения, равной

$$|f_1 - f_2| = K_q \cdot \Delta f ,$$

где K_q – коэффициент, зависящий от того, аналоговый или цифровой разрешающий фильтр, и если аналоговый, то от его порядка; для аналогового фильтра K_q – порядка нескольких единиц, для цифрового – несколько больше единицы; Δf – полоса разрешающего фильтра.

Таким образом, в конечном счете, разрешающая способность СА задается полосой разрешающего фильтра Δf . Когда-то разрешающие фильтры были аналоговыми, затем узкополосные из них стали выполняться цифровыми, а остальные – аналоговыми; наконец, в настоящее время все разрешающие фильтры принято выполнять цифровыми. Это привело к тому, что как фактор эффективности разрешающая способность представлена диапазоном полос разрешения и дробностью (количеством) их значений. Фактически и то, и другое ограничено лишь здравым смыслом. Минимальное значение Δf_{\min} разрешающей полосы ограничено тем, что для получения одного отсчета требуется время $K / \Delta f_{\min}$, где K – порядка единиц. Обычно берут Δf_{\min} порядка 1 Гц, при меньших полосах сеанс измерения будет длиться недопустимо долго. Максимальное значение Δf_{\max} ограничено тем, что должно выполняться условие $\Delta f_{\max} \ll f_{\text{посл ПЧ}}$. В результате, при значении последней ПЧ порядка нескольких десятков МГц отношение $\Delta f_{\max} / \Delta f_{\min}$ составляет не более 7 декад. При 10 разрешающих фильтрах на декаду получается всего порядка 70 фильтров, при 20 фильтров на декаду – порядка 140 фильтров. Эти фантастические возможности выбора разрешения – одно из преимуществ использования цифровых технологий.

7. Оперативность (быстродействие). Это важный фактор эффективности СА; от оперативности зависит сама применимость данной модели СА в ряде приложений, например, при настройке и тестировании радиотехнических устройств на поточной производственной линии, при анализе сигналов ограниченного времени действия, при выполнении функций обнаружения. Кроме того, оперативность часто является фактором *производительности* СА. Для измерительного прибора *быстродействие* характеризуется интервалом времени, необходимого для производства единичного измерения, в качестве которого в данном случае нужно понимать получение спектра сигнала в заданном частотном диапазоне с необходимым качеством, возможно, в течение нескольких проходов частотного диапазона. Принципиальная часть этого времени – *время измерения*, включающее время переходов с одной частотной точки на другую, время «стояния» на частотных точках, время переключения кадров (предполагается, что процессы записи, расчетов и индикации не удлиняют этого времени). Полезно различать быстродействие собственно измерений, быстродействие с учетом калибровки, быстродействие с учетом калибровки и подготовительных процедур. Например, чтобы СА вышел на стандартный режим теплового тренда, может понадобиться несколько десятков минут [2.6]; это часть времени подготовительных процедур; альтернатива – постоянно включенное питание.

8. Стабильность отражает постоянство во времени метрологических показателей средства измерений. Часто эта характеристика представляется обратной величиной – *нестабильностью* показателей во времени. Общая нестабильность СА складывается из нестабильности (в частности, частотной) гетеродинов, переключателей, разъемов и т.д.

9. Помехозащищенность и электромагнитная совместимость. *Электромагнитная совместимость* средства измерения определяется гарантией достаточно низкого уровня его влияния на окружающие электромагнитные устройства, обеспечивающей сохранение свойств последних. Поскольку СА – пассивное устройство, а единственными его активными элементами являются гетеродины, требования нормативных документов по электромагнитной совместимости легко выполнить.

Помехозащищенность – способность электронного средства измерений сохранять в процессе измерений свои характеристики при наличии внешних радиопомех. Применительно к СА требования к помехозащищенности определяются нормативными документами; выполнимость этих требований зависит от состава, расположения и характеристик окружающих устройств; с другой стороны, влияние внешних радиопомех выражается характеристиками точности, обнаружения, динамического диапазона измерения полезных сигналов под влиянием этих помех.

10. Надежность. Общие комментарии к этому фактору эффективности даны выше.

11. Вес и габариты. Эти факторы не требуют комментариев.

12. Рабочие условия (условия измерений). Общие комментарии к этому фактору эффективности даны выше.

13. Поколения СА. Общие соображения о поколениях радиотехнических средств измерения изложены выше. Что касается СА, то большинство современных моделей этого прибора можно отнести к четвертому поколению: они имеют автоматическое управление от встроенного или внешнего компьютера; используют микросхемы и микросборки большой степени интеграции, многослойные печатные платы; используют цифровые технологии обработки сигналов после детектора огибающей или даже после последнего смесителя.

Можно ожидать, что для пятого поколения СА будет характерна полная готовность средств измерения к объединению в измерительные системы, с межкомпьютерной трансляцией данных и результатов, автоматическим формированием, выдачей и распечаткой отчетов с широким выбором номенклатуры и форматов, хранение установленных параметров, данных и результатов измерений в выбираемых номенклатурах и форматах и т. д.; а также переход к цифровой обработке не позднее тракта второй ПЧ; векторный характер спектрального анализа; возможность анализа в отличных от Фурье базисах, в частности, вейвлет-анализа.