

射频微电子学

译者序

近年来，由于无线通信的广泛需求和迅速发展，射频集成电路已成为 IC 设计中被普遍关注和大量研究的课题。尤其是因为硅 CMOS 工艺的不断改进，MOS 器件的截止频率已完全能满足多数现代无线通信应用要求（比如 0.13 微米工艺节点的 nMOS 截止频率可达 90GHz）。对一些要求不高的应用，如无线局域网（WLAN）射频前端收发器，CMOS 器件噪声性能等也能满足要求。因此了解和掌握射频集成电路的工作原理与分析方法对正确有效地设计射频、数字模拟混合信号集成电路是至关重要的。

Behzad Razavi 所著的《射频微电子学》（RF Microelectronics）一书就是针对上述需要的。它既可供高年级大学生或研究生作教科书用，又可满足集成电路设计工作者进一步提高自身知识和设计技能之目的。相对其他 RF CMOS 电路的专著，本书的特点是系统级的介绍较为详细。即它将无线通信电路系统的描述、器件特性及单元电路分析融合在一起，使读者能对射频电路的 IC 实现有一个完整的概念。

本书第 1 和第 2 章首先介绍了在射频电子学经常遇到的概念和术语，并给出了评价射频电路性能的主要指标，如表征非线性度的 1-dB 增益压缩点和三阶交调点、灵敏度和动态范围等。其中对随机信号和噪声的描述很有特色，例如定义随机过程为一组时间的函数或多次测量的集合，这对理解随机信号的本质很有帮助。

在通信系统中，模拟和数字信号的调制、解调是必须的过程。本书第 3 章先用简洁的数学公式定义和描述了模拟信号的调幅、调相和调频调制方法，然后自然地扩展到数字信号调制的幅移、相移和频移键控调制。本书对数字信号调制机理的讨论十分明了，是一个对数字信号系统不甚熟悉的读者很好的入门介绍。另外，在书中还对各种调制方法的功率效率进行了讨论。

本书第 4 章专门介绍 3 种在无线通信系统中通用的多路存取技术：FDMA，TDMA 和 CDMA，及几种常用的无线通信标准，如 GSM 和 Qualcomm 公司的 CDMA，这些对了解系统的工作方式很有帮助。

第 5 章是本书的重点，讨论了无线前端收发器的结构和集成电路的实现。它在介绍单元电路之前，先比较了各种接收器和发送器的结构，包括外差和零中频程式；然后给出了几种常见的无线收发器的设计实例。这些设计都已在产品中实现并被用于无线通信系统中，其中有 GSM 和 DECT 收发器的电路框图。这些高层次的讨论对理解后续的构成收发器的单元电路设计和工作原理打下了基础。

本书的第 6 章至第 9 章详细讨论了低噪声放大器和混频器、振荡器、频率综合器和功率放大器的电路原理和分析方法。电路的实现以 MOS 器件为主，但也有用双极型晶体管作实例的。电路的工作原理讨论较为简要，但对各个电路单元要达到的技术指标和评价标准给出了不少实际应用系统的具体数据，这对于读者建立相关电路设计的定量概念颇有益处。

本书的作者是美国 UCLA（加州大学洛杉矶分校）的教授，他曾在 HP 实验室工作过，有丰富的模拟电路设计经验。近年来更在射频 CMOS 的电路理论和设计研究方面大有建树，并成功地创办了用硅 CMOS 工艺实现的 WLAN 芯片设计公司。他写的另一本有关 CMOS 模拟集成电路的书 *Design of Analog CMOS Integrated Circuits*，作为电路教学的教科书，为美国多所大学采用，也在清华大学被用作本科生和研究生的模拟电路教材或参考书。

本书在讨论 RF CMOS 电路的工作原理和分析上略为简明。读者如果了解更多的电路分析知识和方法，[来](#)作为学习 RF 系统级设计、分析的补充，可进一步参考美国斯坦福（Stanford）大学 Thomas H. Lee 教授的 *The Design of CMOS Radio-Frequency Integrated Circuits* 一书（英国剑桥大学出版，中国电子出版社有第二版的影印本，2005）。

译者在清华大学讲授 CMOS 射频集成电路分析与设计研究生课程中，自 2000 年来同时采用 Thomas Lee 和本书（英文版）作为教材（电路部分以 Lee 书为主，系统部分参考 Razavi 书）。这本翻译稿就是和清华大学微电子学研究所的周润德教授以及该所 CAD 室修过这门课的博士研究生一起完成的。可能存在的错误或不准确的地方，则由本人负责。

射频 CMOS 电路是一门依然在发展中的活跃学科，希望通过普及这方面的基础知识，能起到介绍有兴趣的读者，尤其是年轻学人，投身于集成电路与系统的前沿研究。

余志平于清华大学

2005 年 12 月

目录

前言

1. 射频 (RF) 及无线技术简介
 - 1.1. 复杂性比较
 - 1.2. 设计瓶颈
 - 1.3. 应用
 - 1.4. 模拟与数字系统
 - 1.5. 工艺技术的选择

参考文献
2. RF 设计中的基本概念
 - 2.1. 非线性与时变性
 - 2.1.1. 非线性的影响
 - 2.1.2. 级联非线性级
 - 2.2. 符号间干扰
 - 2.3. 随机过程和噪声
 - 2.3.1. 随机过程
 - 2.3.2. 噪声
 - 2.4. 灵敏度与动态范围
 - 2.5. 无源阻抗变换

参考文献
3. 调制与解调
 - 3.1. 概述
 - 3.2. 模拟调制
 - 3.2.1. 幅度调制
 - 3.2.2. 相位调制与频率调制
 - 3.3. 数字调制
 - 3.3.1. 基本概念
 - 3.3.2. 二进制调制
 - 3.3.3. 正交调制

3.4. 调制技术的功率效率

3.4.1. 等幅包络与变幅包络信号

3.4.2. 频谱再增长

3.5. 非相干解调

参考文献

4. 多址访问技术与无线通信标准

4.1. 移动射频 (RF) 通信

4.2. 多址访问技术

4.2.1. 时分和频分复用

4.2.2. 频分多址

4.2.3. 时分多址

4.2.4. 码分多址

4.3. 无线通信标准

4.3.1. 先进移动电话服务 (AMPS)

4.3.2. 北美数字标准 (NADS)

4.3.3. 移动通信全球系统 (GSM)

4.3.4. 高通 (Qualcomm) CDMA

4.3.5. 欧洲数字无绳电话 (DECT)

参考文献

5. 接收发送器结构

5.1. 概述

5.2. 接收器结构

5.2.1. 外差接收器

5.2.2. 零差接收器

5.2.3. 镜像抑制接收器

5.2.4. 数字中频接收器

5.2.5. 亚采样接收器

5.3. 发送器结构

5.3.1. 直接变换发送器

5.3.2. 两步发送器

5.4. 接收发送器的性能测试

5.5. 实例研究

5.5.1. Motorola 调频接收器

5.5.2. Philips 传呼机接收器

5.5.3. Philips DECT 接收发送器

5.5.4. Lucent (朗讯) GSM 接收发送器

5.5.5. Philips GSM 接收发送器

参考文献

6. 低噪声放大器与混频器

6.1. 低噪声放大器

6.1.1. 概述

6.1.2. 输入匹配

6.1.3. 双极型 LNA

6.1.4. CMOS LNA

6.2. 下变频混频器

6.2.1. 概述

6.2.2. 双极型混频器

6.2.3. CMOS 混频器

6.2.4. 混频器中的噪声

6.3. 级联级再讨论

参考文献

7. 振荡器

7.1. 概述

7.2. 基本的 LC 振荡器拓扑

7.3. 电压控制 (压控) 振荡器

7.4. 相位噪声

7.4.1. 相位噪声在射频通信中的影响

7.4.2. 振荡器的 Q 值

7.4.3. 相位噪声机制

7.4.4. 噪声—功率折衷关系

- 7.4.5. 分频与倍频对相位噪声的影响
 - 7.4.6. 振荡器的牵引与推动
 - 7.5. 双极型与 CMOS LC 振荡器
 - 7.5.1. 负跨导 (G_m) 振荡器
 - 7.5.2. 插值振荡器
 - 7.6. 单片集成电感
 - 7.7. 不带谐振腔的压控振荡器 (VCO)
 - 7.8. 正交信号产生
 - 7.8.1. RC-CR 网络
 - 7.8.2. Havens 方法
 - 7.8.3. 分频
 - 7.9. 单边带信号生成
 - 参考文献
- 8. 频率综合器
 - 8.1. 概述
 - 8.2. 锁相环 (PLL)
 - 8.2.1. 基本概念
 - 8.3. 基本的锁相环
 - 8.3.1. 电荷泵锁相环
 - 8.3.2. I型和II型锁相环
 - 8.3.3. 锁相环中的噪声
 - 8.3.4. 输入端的相位噪声
 - 8.3.5. 压控振荡器的相位噪声
 - 8.3.6. 倍频
 - 8.4. 射频频率综合器结构
 - 8.4.1. 整数-N结构
 - 8.4.2. 分数-N结构
 - 8.4.3. 双环路结构
 - 8.4.4. 直接数字频率综合
 - 8.5. 分频器

8.5.1. 二分频电路

8.5.2. 双模分频器

参考文献

9. 功率放大器

9.1. 概述

9.1.1. 线性与非线性功率放大器

9.2. 功率放大器的分类

9.2.1. A类和B类功率放大器

9.2.2. C类功率放大器

9.3. 高效率功率放大器

9.4. 大信号阻抗匹配

9.5. 线性化技术

9.5.1. 前馈

9.5.2. 反馈

9.5.3. 包络消除及恢复

9.5.4. 采用非线性元件的线性放大 (LINC)

9.6. 设计实例

参考文献

前言

全球的手机年销售量已超过 25 亿美元。有 4 亿 50 万客户的家庭卫星网络也成了 25 亿美元的工业。到了 2000 年，全球定位系统有望成为 50 亿美元的市场。在欧洲，移动通信设备的销售和服务将在 1998 年达到 300 亿美元。这些统计数字令人眩目。

射频 (RF) 和无线通信的市场突然一下子膨胀到不可想象的地步。诸如传呼机、手机及无绳电话、有线电视调制解调器以及射频标签的电讯产品正在迅速地渗入我们的生活。从奢华用品变成不可缺少的工具。大大小小的半导体和系统公司，不管是模拟或数字电路的，都看到了这些统计数字。他们都在化大力气推出各种各样的 RF 产品来占据各自的市场。

RF 设计与众不同，它需要涉及许多与集成电路 (IC) 无关的学科领域。有关 RF 的知识已经成长了几乎一个世纪，因而形成了对于新入门的人看起来似乎是无穷尽的文献资料。

这本书讨论 RF 集成电路及系统的分析与设计。通过用启蒙语言来提供一个对射频电子学的系统介绍，本书首先给出从微波到通信理论的必要的背景知识，然后引入 RF 接收发送器和电路的设计。书中从 VLSI 技术的单片实现出发，同时强调系统结构和电路级的课题。内容的处理主要集中在双极型和 CMOS 设计，但是大多数概念也可以用到其它工艺技术。我们假定读者对模拟集成电路设计和信号及系统的理论已有了一些基本了解。

全书由九章组成。第 1 章给出了一般性的介绍，提出了问题，并为后续各章提供了学习的动力。第 2 章描述了 RF 与微波设计的基本概念，强调了非线性和噪声的影响。

第 3 和第 4 章把读者带到通信系统的层次上，回顾了调制、解调、多址访问技术，以及无线标准。这些内容尽管初看起来没有必要，但事实上它们对 RF 电路与系统的并发设计是不可缺少的。

第 5 章研究接收发送器的系统结构，讨论各种不同的接收器和发送器的电路拓扑结构以及它们的优劣之处。这一章也包括一些范例研究来说明在实际的 RF 产品中所采用的途经。

第 6 到 9 章讨论 RF 基本构筑模块的设计：低噪声放大器和混频器、振荡器、频率综合器以及功率放大器。特别研究如何使片外的元件数保持到最少。这些章节的一个重要目的是演示为满足系统的要求如何去定义电路的参数，以及每个电路的性能又是如何影响整个接收发送器。

我在 UCLA (美国加利福尼亚大学洛杉矶分校) 的一门 4 学分的研究生课程中曾讲授过本书大约 80% 的材料。对于只有十周课的学季制，第 3, 4, 8, 和 9 章的内容不得被缩减，但对于学期制，它们可以被较仔细地讲授。

我的 RF 设计知识许多是来自和同事们的交流。贝尔实验室的 Helen Kim, Ting-Ping Liu, 及 Dan Avidor, 惠普 (Hewlett-Packard) 公司实验室的 David Su 和 Andrew Gzegorek 在许多方面对本书的材料作出了贡献。这本书也为许多专家审阅过：Stefan Heinen (西门子), Bart Jansen (惠普), Ting-Ping Liu (贝尔实验室), John Long (多伦多大学), Tadao Nakagawa (NTT), Gitty Nasserbakht

(德州仪器公司), Ted Rappaport (佛吉尼亚理工学院), Tirdad Sowlati (Gennum), Trudy Stetzler (贝尔实验室), David Su (惠普), 以及 Rick Wesel (UCLA)。此外, 许多 UCLA 的学生, 包括 Farbod Behbahani, Hooman Darabi, John Leete, 以及 Jacob Rael, “试用”了不同的章节, 并且提供了有用的反馈意见。我对所有上面提到的人表示感谢他们的热心帮助。我也要感谢 Prentice Hall 出版社的工作人员, 尤其是 Russ Hall, Maureen Diana, 及 Kerry Reardon, 感谢他们的支持。

Behzad Razavi
1997年7月

第 1 章 射频 (RF) 和无线技术简介

一位工程师说“这个手机用的是在 DECT-TDMA/TDD 模式下的 GFSK 方式调制以及零中频的 I/Q 检波”。他的经理问道，“我们如何能将它改成也能在 DCS1800 制式下工作？”工程师回答说，“我们需要加一个天线收发转换开关，至少一个带有 SSB 混频的本机振荡器，或许还要两个 SAW 滤波器，”等等。

电话已经变得比先前复杂多了。射频电路也是如此。普通一个用手机往家打电话的人多半不会知道成百上千的科学家和工程师几乎努力了整整一个世纪才使得无线技术变得用得起。他也不会知道在手机里计算能力比某些早期的个人计算机还要强。

在 1901 年当古格列尔莫·马可尼 (Marconi) 成功地将无线电信号发送过大西洋时，无线技术就出现了。这个演示产生的影响和展示的前景是巨大的：用无线电波在“以太”中的传递来取代电报和电话描绘了一个令人振奋的未来。然而尽管双向无线通信在军事应用中实现了，日常生活中的无线传播仍然局限于由大型、昂贵的发射站单向发送无线电广播和电视。普通的双向电话通话几十年来一直通过电话线进行。

晶体管的发明，仙农 (Shannon) 信息论的发展，以及蜂窝系统概念的提出 — 所有这一切都发生在贝尔实验室 — 这为提供廉价的移动通信铺平了道路。就像最初在汽车电话中实现，最终又在便携式的手机中实现。

但是为什么会在无线通信电子上出现突然的上升？市场调查表明在美国每天有超过 28,000 人新加入使用手机系统。这就激励了互相竞争的制造厂商提供更具高性能和低成本的电话机。事实上，当前的目标是使手机的功耗和成本每年减少 30%，尽管不清楚这种速率能维持多久。然而，当双向无线通信被引入我们生活的其它方面时 — 如家庭电话，计算机，传真机和电视，一个更为光彩的前景呈现了。虽然无线通信工业界的近期目标是将无绳电话和手机结合起来形成几乎任何地方都可使用的“无缝”通信，长期的目标是实现一个无所不能的无线终端，它能处理声音，数据，图象，以及提供计算能力。其它一些更为“奢华”的应用，诸如全球定位系统 (GPS)，也会在今后的某一时候通过这个终端变成现实。个人通信服务 (PCS) 几乎就在眼前。

1.1 复杂性比较

为了得到一个现代 RF 设计的初步认识，首先让我们来考虑一下图 1.1 所示的一个简单的频率调制 (FM) 接收发送器系统。在图 1.1(a)中， Q_1 既作为一个振荡器又作为一个频率调制器；亦即是话筒产生的音频信号改变了变容二极管 D_1 两端的偏置电压，因此调制了振荡频率。在接收路径中， Q_1 同时用作振荡器和频率解调器，在节点 X 产生音频信号。这个音频信号然后被放大并加到扬声器上去。

图 1.1 (a) 调频发送器，(b) 调频接收器

现在考虑图 1.2 的电路，即手机的 RF 部分[1]。这个电路比图 1.1 所示的 FM 电路要复杂好几个数量级，我们将其分析推迟到第 5 章进行。为什么 RF 设计者从图 1.1 的电路进展到图 1.2 的电路？什么是这个演变的思维过程？这些复杂性真的必需吗？当我们在第 2 到第 5 章中逐渐了解 RF 通信系统、体系结构和电路设计时，这些问题将得到回答。

图 1.2 手机的射频 (RF) 部分[1]

1.2 设计瓶颈

今天的袖珍电话（即手机）包含了一百多万晶体管，其中只有一小部分工作在射频范围，其余都是进行低频的“基带”模拟和数字信号处理（图 1.3）。换句话说，就器件的数目而言，基带部分要比图 1.2 的电路复杂几个数量级。RF 和基带的定义随后会变得清楚，这里我们只是指出 RF 部分仍是整个系统设计的瓶颈。这是因为如下三个原因。

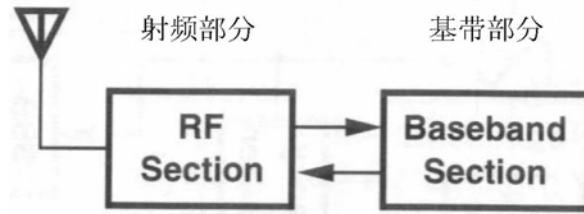


图 1.3 接收发送器的射频和基带处理

多学科领域 和其它类型的模拟与混合信号电路不同，RF 系统要求了解并不直接与集成电路（IC）相关的许多领域。图 1.4 中所示，大多数这些领域已经被深入地研究了半个多世纪，这使得一个集成电路设计者要在较短的时间内掌握必要的知识相当困难。即使是在当前，有关 RF 设计的文献就出现在 30 多种杂志与会议上。

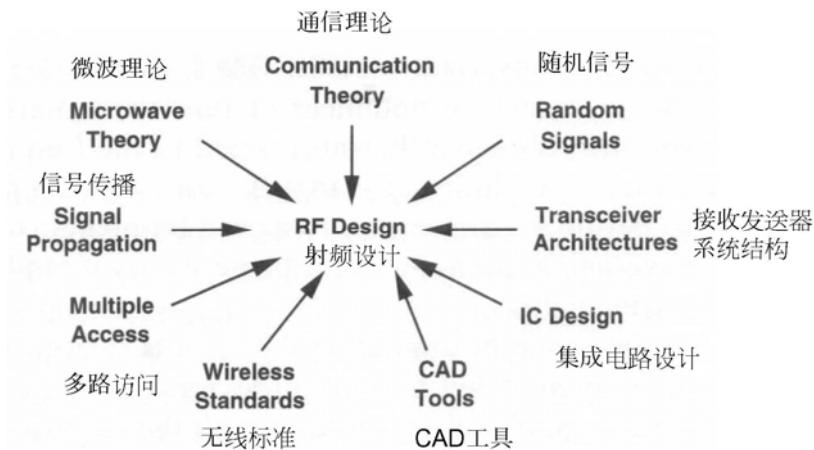


图 1.4 射频设计要求的学科

由于这个原因，传统的无线系统设计是在某种程度上互相不连接的抽象级别上进行的：通信理论工作者创造了调制方案和基带信号处理；射频系统专家规划接收发送器的系统结构；集成电路（IC）设计者开发了每一个基本模块；制造厂则将 IC 及其它外部元件“粘”在一起。事实上，系统结构往往是按照可以购得的现成（off-the-shelf）产品元件来规划的，而 IC 则是设计成能适用于尽可能多的体系结构。这样就造成了在系统和电路级大量的重复性。随着工业界向更高的集成度和低成本发展，射频和无线的设计要求日渐增加的“并发工程”，因而要求 IC 设计者具有在图 1.4 所示的所有领域里的足够知识。

射频设计六边形 RF 电路必须在很宽的动态范围及高频下处理模拟信号。有意思的是即使调制信号是数字信号，或者振幅不载有任何信息，信号也必须按模拟信号处理。（这一点会在以

后的章节中澄清。)在这种电路设计中涉及的折衷考虑可以总结为如图 1.5 所示的“射频设计六边形”。在这个图中,六个参数中几乎任两个都可以在某种程度上进行折衷调节。我们将要在第 6 到 9 章中看到这种折衷关系。但我们现在要看到的是,尽管数字电路直接从 IC 技术的进步得到好处,RF 电路却并没有得到同样多的帮助。这个问题由于 RF 电路经常要用到外部的元件,比如说电感器,而变得更为严重。这些外部元件即使在现代的集成电路工艺中也是很难在片上实现的。

图 1.5 射频设计的六边形

设计工具 用于 RF IC 的计算机辅助分析和综合的工具仍然处于初始阶段,这迫使设计者不得不依赖于经验,直觉或不那么有效的模拟技术来预测电路的性能。比如说,RF 电路中的非线性性,时变性和噪声通常需要研究信号的频谱。但是,SPICE 具有的标准交流分析仅用到线性和时不变模型。这样,电路只能在时域里模拟以便包括非线性和时变效应,所得到的波形随后再变换到频域来得到频率响应。困难在于时域里的分析必须要足够长以能分辨频率相近的成分。此外,假若在时域分析中用到随机噪声,则可能需要用到频谱平均技术。模拟 RF 电路时的另外一个问题是一些外部元件无法用 SPICE 中已有的典型器件来建立模型。比如,在接收和发送路径中都要用到表面声波(SAW)滤波器,其输入输出特性只能用散射(S)参数来表征——实质上是一个数值表格。用 RLC 网络来对这样的电路建模提供了一阶近似,但是,这种方法可能预测不到非稳定性及阻抗不匹配的影响。

1.3 应用

除了诸如传呼机和手机这些为人们熟悉的无线通信产品以外,RF 技术已经创造了许多其他市场。这些市场展示了快速成长的巨大潜力,每一个都对 RF 设计者造成挑战。

无线局域网(WLAN) 在一个拥挤的场所,人们或设备之间的通信可以通过无线局域网(WLAN)来实现。采用在 900 MHz 和 2.4 GHz 附近的频带,WLAN 接收发送器能在办公室,医院,工厂等地提供移动通信连接,这样就免除了使用笨拙的有线网络的需要。便携性与重构性是 WLAN 的显著特征。

全球定位系统(GPS) 随着 GPS 接收器的成本和功耗下降,用它来确定一个目标的位置及寻得方向变得对消费市场十分有吸引力。这样的系统在 1.5 GHz 下工作,不仅为汽车制造厂家所考虑,而且在不久的将来可以作为低成本的手持产品。

射频标签(RF ID) 射频标识系统,简称“RF ID”是小的,低成本的标签。它们可以附加到物品或个人来跟踪其位置。应用范围包括飞机场的行李到军事行动的部队。因为标签的寿命可能为单个小电池的寿命所决定,所以低功耗的要求尤其重要。工作在 900 MHz 和 2.4 GHz 频率范围的 RF ID 产品最近已出现在市场上。

家庭卫星网络(Home Satellite Network) 卫星电视所提供的节目与服务已经使众多的用户为家庭卫星网络所吸引。这些网络工作在 10 GHz 频段,需要附加碟形天线以及连到电视机的接收器。它们直接与有线电视竞争。

1.4 模拟与数字系统

为了得到一个无线接收发送器的一般的和初步的了解,首先考虑图 1.6 所示的简单的“模拟”体系结构。在发送路径中,话筒产生的信号对高频载波进行调制,调制后的信号被放大并经过“缓冲”来驱动天线。在接收路径中,信号经低噪声放大器(LNA)放大,频谱由一个下变频器(通常是一个混频器)变换到较低的频率以有利于嗣后的解调操作。解调后的输出被放大来驱动扬声器。

图 1.6 一般化了的模拟射频系统框图: (a) 发送器, (b) 接收器

现在,考虑图 1.7 所示的“数字”接收发送器。这里,声音信号首先被模数转换器(ADC)数字化,并经压缩以减少位率(bit rate),因而也就是减少所需要的带宽(图 1.7(a))。下一步,数据经过“编码”和“间插”[2]。这两个操作对数据进行了格式化,这样接收器就能够通过反向操作来解调并将误差减小到最低程度。因为矩形脉冲对调制而言并不是最佳,因此数据在送到调制器及功率放大器(PA)之前先进行“整形”。在接收路径中(图 1.7(b)),信号经过放大、下变频及数字化。然后在数字域里进行解调、平衡化、解码,去间插及去压缩。这样得到的数据然后被数模转换器(DAC)转换成模拟信号,再经放大并送到扬声器去。

图 1.7 一般化了的数字射频系统框图: (a) 发送器, (b) 接收器

在这个最简单的例子里,这些接收发送器可以用它们在提供满意的接收效果条件下能工作的最大距离来定量地评价其性能。这个距离是由送到天线上去的功率和接收器的灵敏度,尤其是 LNA¹的噪声共同决定的。然而,在实际的环境下,许多不同的接收发送器同时在工作,而且往往彼此间距离很近。这样引起了互相干扰。更进一步,通信媒介是十分复杂的:在建筑物和其它物体上发生的反射会在给定的地点产生破坏性的干涉,使得接收到的信号强度被抑制到可检测到的水准之下。在这样一个环境下,数字接收发送器的信号处理可以取得比模拟系统更高的性能。

研究一下图 1.6 及 1.7 的两个接收发送器,读者会问哪些部分是和射频(RF)电子学有关的?我们将在其它章节看到,RF 的确切定义取决于具体的系统,但是我们暂时将频谱中心不在零频率的模拟信号视作是在 RF 区域。这样在图 1.6 中除了音频放大器之外的所有电路模块都是 RF 电路。类似地,在图 1.7 中,调制器、功放(PA)、低噪声放大器(LNA)及下变频器都工作在 RF 范围。然而,现代的 RF 电子学包括比以上这些电路单元多得多的内容。比如说对每种情况都要仔细地研究系统级的折衷考虑 [3, 4]。

这本书主要讨论 RF 集成电路的设计。但是,依照图 1.4 所示的领域和课题,本书也会提供相关的背景材料。我们在全书中将会看到,RF IC、RF 系统结构以及无线系统相互间是紧密联系着的,因此读者经常要用到这些背景知识。

1.5 工艺技术的选择

用于 RF 电路的关键 IC 工艺技术不断地在变化。性能、成本以及进入市场的时间是三个在竞争激烈的 RF 工业界中影响相关工艺技术选择的主要因素。此外,诸如集成度,体积面积

¹ 调制的类型也是一个重要因素,但是我们目前先不考虑。

因子 (form factor), 以及过去 (成功) 的经验都在设计者作决定时起重要作用。目前, 砷化镓 (GaAs) 及硅双极型和 BiCMOS (双极 CMOS) 工艺技术构成了 RF 市场的主要部分。GaAs 场效应及异质结器件通常被看作是低成品率、高功耗和高成本的选择。但是它们仍然在 RF 产品中占重要位置, 特别是在功率放大器及前端开关中。尽管 GaAs 工艺具有诸如更高的击穿电压—截止频率的乘积值、半绝缘的衬底、高品质因子的电感和电容等有用的特点, 但 VLSI 技术中的硅器件有潜力提供更高的集成度和更低的总成本。这一点在诸如频率综合器一类的复杂电路中已得到演证。事实上, 许多制造厂提供用硅双极型技术实现的用于典型接收发送器中的所有构筑模块。第三个竞争者是 CMOS 技术。受数字电路市场的巨大动量的支持, CMOS 器件已经取得了很高的截止频率, 比如用 0.35 μm CMOS 技术可以达到几十 GHz。正像我们在这本书中将要看到的, “RF CMOS” 已突然变成活跃的研究课题。然而, CMOS 技术必须解决一些具体的问题: 幅度相差 100dB 的信号通过衬底的耦合, 参数因温度和工艺过程的变化, 以及针对 RF 工作情况建立器件模型。

参考文献

- [1] C. Marshall *et al.*, “2.7 V GSM Transceiver ICs with On-Chip Filtering,” *ISSCC Dig. Tech. Papers*, pp. 148-149, February 1995.
- [2] B. Sklar, *Digital Communications*, Englewood Cliffs, NJ: Prentice Hall, 1988.
- [3] K. Feher, *Wireless Digital Communications*, Englewood Cliffs, NJ: Prentice Hall, 1995.
- [4] R. Steele, ed. *Mobile Radio Communications*, Piscataway, NJ: IEEE Press, 1992
- [5] R. Schneiderman, “GaAs Continues to Gain in Wireless Applications,” *Wireless Systems Design*, pp. 14-16, March 1997.

第 2 章 RF 设计中的基本概念

RF 设计者用到很多来自信号与系统理论的概念。在本章中，我们将介绍这些概念并且定义一些用在射频电子学（RF electronics）中的术语，以便读者学习随后章节的材料。

从非线性系统开始，我们介绍诸如谐波失真（harmonic distortion）、增益压缩（gain compression）、互调（cross modulation）和交调（intermodulation）等效应。然后我们将简要地研究符号间的干扰（intersymbol interference）和奈奎斯特采样（Nyquist signaling），回顾随机过程和噪声的知识，并且将介绍在电路中表示噪声的方法。最后我们将描述用无源元件实现阻抗变换。

2.1 非线性与时变性

如果一个系统的输出可以表示成每个输入分别对应的输出的线性叠加，那么这个系统就是一个线性系统。更准确地说，如果对于输入 $x_1(t)$ 和 $x_2(t)$ ，我们有

$$x_1(t) \rightarrow y_1(t), x_2(t) \rightarrow y_2(t) \quad (2.1)$$

其中箭头表示线性系统对信号的操作，那么对应于常数 a 和 b 的所有值，我们有

$$ax_1(t) + bx_2(t) \rightarrow ay_1(t) + by_2(t) \quad (2.2)$$

任何不满足这一条件的系统都是非线性的。注意到根据这个定义，如果一个系统具有非零的初始条件或者有限的偏移量，我们就认为是非线性的系统。

如果一个系统的输入的时间平移将导致输出有相同的时间平移，那么这个系统就是时不变的。也就是说如果 $x(t) \rightarrow y(t)$ ，那么时不变系统对于所有的 τ 值都有 $x(t - \tau) \rightarrow y(t - \tau)$ 。如果一个系统不满足这个条件，就叫做时变系统。

非线性和时变性都是直观上很显然的概念，但是它们在某些情况下可能相互混淆。例如考虑如图 2.1(a) 中的开关电路，开关的控制端由 $v_{in1}(t) = A_1 \cos \omega_1 t$ 驱动，输入端由 $v_{in2}(t) = A_2 \cos \omega_2 t$ 驱动。我们假设当 $v_{in1} > 0$ 时开关接通，反之断开。那么这个系统是非线性的还是时变的？如果如图 2.1(b) 所示我们考虑的通路是从 v_{in1} 到 v_{out} （此时 v_{in2} 是系统的一部分仍然等于 $A_2 \cos \omega_2 t$ ），那么这个系统是非线性的，因为这时控制只与 v_{in1} 的极性有关；同时又是时变的，因为 v_{out} 还与 v_{in2} 有关。另一方面，如图 2.1(c) 所示，如果所考虑的路径是从 v_{in2} 到 v_{out} （此时 v_{in1} 是系统的一部分并仍然等于 $A_1 \cos \omega_1 t$ ），那么系统是线性的[式(2.2)]和时

变的。所以诸如“开关是非线性的”等笼统的说法是很含糊的。我们将在第六章中看到这些区别在混频器的设计中是很重要的。

图 2.1 (a) 简单的开关电路, (b) 非线性时变系统, (c) 线性时变系统

上述观察的另一个有趣的结果是一个线性系统可以产生输入信号中并不存在的频率成分。这在系统是时变的情况下是可能的, 例如图 2.1(c)所示的电路就是如此。因为在这样的电路中 v_{out} 可以看作是 v_{in2} 和一个在 0、1 之间翻转的方波相乘的积, 其输出频谱是

$$V_{out}(f) = V_{in2}(f) * \sum_{n=-\infty}^{+\infty} \frac{\sin(n\pi/2)}{n\pi} \delta(f - \frac{n}{T_1}) \quad (2.3)$$

$$= \sum_{n=-\infty}^{+\infty} \frac{\sin(n\pi/2)}{n\pi} V_{in2}(f - \frac{n}{T_1}) \quad (2.4)$$

其中 * 代表卷积, $\delta(\bullet)$ 是 Dirac δ 函数, $T_1 = 2\pi/\omega_1$ 。所以输出是由 $v_{in2}(f)$ 在时间上平移 n/T_1 在幅度上乘以不同倍数后叠加而成。

如果一个系统的输出与过去的输入信号无关, 那么这个系统就是无记忆的 (memoryless) 系统。对于一个无记忆的线性系统

$$y(t) = \alpha x(t) \quad (2.5)$$

如果系统是时变的, 那么其中的 α 是时间的函数[如图 2.1(c)]。对于一个无记忆的非线性系统, 输入—输出的关系可以近似的用多项式表示为

$$y(t) = \alpha_0 + \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t) + \dots \quad (2.6)$$

如果系统是时变的, 其中的 α_j 一般是时间的函数。图 2.2(a)描述了一个例子, 其中输入信号作用于 Q_1 的基极, 而 Q_2 和 Q_3 则由方波周期性地导通和截止。对于理想的双极型晶体管, 这个电路可以看作如图 2.2(b)的形式, 并且有

$$v_{out}(t) = (I_{S1} \exp \frac{v_{in}}{V_T}) s(t) \cdot R \quad (2.7)$$

其中 I_{S1} 表示 Q_1 的饱和电流, $V_T = kT/q$, $s(t)$ 是一个在 -1、1 之间翻转的方波。

图 2.2 (a) 尾电流源为信号所驱动的开关差分对, (b) (a) 的等效电路

如果由式 (2.6) 描述的系统对于相反的输入 $x(t)$ 和 $-x(t)$ 有相反的输出, 那么它具有奇对称性。这一情况发生在偶数 j 所对应的 $\alpha_j = 0$ 的时候。一个有奇对

称性的电路叫做差分电路或者平衡（balanced）电路。例如图 2.3 的双极型差分对显示了下列输入—输出特性

$$v_{out} = RI_{EE} \tanh \frac{v_{in}}{2V_T} \quad (2.8)$$

上式是一个奇函数。

图 2.3 双极型差分对以及其输入—输出特性

如果一个系统的输出依赖于输入或者输出以前的值，那么这个系统叫做动态（dynamic）系统。例如对于一个线性时不变动态系统

$$y(t) = h(t) * x(t) \quad (2.9)$$

其中 $h(t)$ 表示脉冲响应。如果一个动态系统是线性但时变的，它的脉冲响应将和时间原点有关；如果 $\delta(t) \rightarrow h(t)$ ，那么 $\delta(t-\tau) \rightarrow h(t, \tau)$ 。所以

$$y(t) = h(t, \tau) * x(t) \quad (2.10)$$

最后，如果一个系统是非线性和动态的，那么它的脉冲响应可以近似的用 Volterra 序列表示[1, 2]，但这一内容超出了本书的范围。

2.1.1 非线性的影响

虽则很多模拟电路和 RF 电路可以用一个线性模型近似表示来得到它们对小信号的响应，但是非线性经常导致一些有趣而且重要的现象。为简单起见，我们把我们的分析局限于无记忆的时变系统并假设

$$y(t) \approx \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t) \quad (2.11)$$

但是读者必须注意好好的检查储能元件和高阶非线性项以保证 (2.11) 的合理性。

谐波 (Harmonics) 如果一个正弦信号作用于一个非线性系统，输出一般将包含输入信号频率的整数倍频。在式 (2.11) 中，如果 $x(t) = A \cos \omega t$ 那么

$$y(t) = \alpha_1 A \cos \omega t + \alpha_2 A^2 \cos^2 \omega t + \alpha_3 A^3 \cos^3 \omega t \quad (2.12)$$

$$= \alpha_1 A \cos \omega t + \frac{\alpha_2 A^2}{2} (1 + \cos 2\omega t) + \frac{\alpha_3 A^3}{4} (3 \cos \omega t + \cos 3\omega t) \quad (2.13)$$

$$= \frac{\alpha_2 A^2}{2} + \left(\alpha_1 A + \frac{3\alpha_3 A^3}{4} \right) \cos \omega t + \frac{\alpha_2 A^2}{2} \cos 2\omega t + \frac{\alpha_3 A^3}{4} \cos 3\omega t \quad (2.14)$$

在式 (2.14) 中，含输入频率的项叫做“基频”，高阶项叫做“谐波” (harmonics)。

从上面的展开式中可以发现两点。第一，由偶数 j 对应的 α_j 产生偶次谐波，并且若系统是奇对称时偶次谐波将消失，也即该系统是完全差分的。然而实际中的不匹配将破坏对称性，并产生一定的偶次谐波。第二，在 (2.14) 中 n 次谐波的幅度由正比于 A^n 的项及其它正比于 A 的更高次幂的项组成。若对于较小的 A 忽略后者，我们可以认为 n 次谐波的幅度近似正比于 A^n 。

增益压缩 (Gain Compression) 一个电路的小信号增益一般是在忽略谐波的假设下得到的。例如在 (2.14) 中若 $\alpha_1 A$ 远大于所有其他任何含 A 的系数，那么小信号增益就等于 α_1 。可以看到在图 2.3 我们所熟悉的差分对中这一量等于

$$\frac{v_{out}}{v_{in}} = \frac{I_{EE} R}{2V_T} \quad (2.15)$$

但是，当信号幅度增加时，增益将有所变化。实际上，非线性可以看作是小信号增益随着输入信号的大小而变化。这一点可以从 (2.14) 中 $3\alpha_3 A^3/4$ 加上 $\alpha_1 A$ 构成的项中看出；同样也可以在图 2.3 所示的输入—输出特性中看出。

在大多数我们感兴趣的电路中，输出是输入的一个压缩或者饱和函数；也就是说对于足够高的输入，增益将接近于 0。在 (2.14) 中如果 $\alpha_3 < 0$ 将发生这种情况，当写成 $\alpha_1 + 3\alpha_3 A^2/4$ 时可以看到增益是 A 的减函数。在 RF 电路中这一影响由 1dB 压缩点 (1-dB compression point) 量化, 定义为使小信号增益下降 1dB 时输入信号的水平。在对数坐标中绘制出输出信号作为输入信号的函数时，输出信号幅度将在 1dB 压缩点处比它的理想值下降 1dB (图 2.4)。

图 2.4 1-dB 压缩点的定义

要计算 1dB 压缩点，我们可以由 (2.14) 得到：

$$20 \log \left| \alpha_1 + \frac{3}{4} \alpha_3 A_{1-dB}^2 \right| = 20 \log |\alpha_1| - 1 \text{dB} \quad (2.16)$$

$$A_{1-dB} = \sqrt{0.145 \frac{|\alpha_1|}{|\alpha_3|}} \quad (2.17)$$

作为电路最大输入范围的一种度量，典型 RF 前端放大器的 1dB 压缩点大约发生在 -20 到 -25dBm (对于 50 Ω 系统为 63.2 到 35.6mVpp)。

减敏和阻塞 (Desensitization and Blocking) 当具有压缩特性的电路处理一个微弱的有用信号时，若还存在一个比较强的干扰信号，则会出现一个有趣的现象。因为一个大信号会降低电路的“平均”增益，这个小信号将会具有一个

近乎零的小增益。这一现象叫做“减敏现象”（desensitization），我们可以通过假设 $x(t) = A_1 \cos \omega_1 t + A_2 \cos \omega_2 t$ 针对式 (2.11) 的特性来分析这一影响。此时输出为：

$$y(t) = (\alpha_1 A_1 + \frac{3}{4} \alpha_3 A_1^3 + \frac{3}{2} \alpha_3 A_1 A_2^2) \cos \omega_1 t + \dots \quad (2.18)$$

当 $A_1 \ll A_2$ 时，上式可简化为

$$y(t) = (\alpha_1 + \frac{3}{2} \alpha_3 A_2^2) A_1 \cos \omega_1 t + \dots \quad (2.19)$$

所以有用信号的增益为 $\alpha_1 + 3\alpha_3 A_2^2 / 4$ ，如果 $\alpha_3 < 0$ 该增益是 A_2 的减函数。对于足够大的 A_2 ，增益将减为 0，这时我们说该信号被“阻塞”（blocked）了。在 RF 设计中，“阻塞信号”（blocking signal）一般是指减敏一个电路的干扰信号，即使这时增益并没有减为 0。许多 RF 接受器必须能够承受比有用信号大 60 到 70dB 的阻塞信号。

互调（Cross Modulation） 当一个弱的信号和一个强的干扰信号同时经过一个非线性系统时发生的另一个现象就是对干扰信号的幅度调制会影响有用信号的幅度。这一现象叫做“互调”（cross modulation），式 (2.19) 可以清楚地说明这一现象，其中 A_2 的变化将会影响到在频率 ω_1 处输出分量幅度的大小。例如：

如果干扰信号的幅度受正弦波 $A_2(1 + m \cos \omega_m t) \cos \omega_2 t$ 调制，其中 m 是调制指数并且小于 1，那么式 (2.19) 就具有以下形式：

$$y(t) = [\alpha_1 A_1 + \frac{3}{2} \alpha_3 A_1 A_2^2 (1 + \frac{m^2}{2} + \frac{m^2}{2} \cos 2\omega_m t + 2m \cos \omega_m t)] \cos \omega_1 t + \dots \quad (2.20)$$

于是，输出的有用信号就受到频率 ω_m 和 $2\omega_m$ 的幅度调节。

互调的一种典型情形出现在那些必须同时处理多个独立的信号通道的放大器中，比如有线电视发射器（cable TV transmitter）中。由于相邻信道幅度的变化，放大器的非线性将会破坏每个信号。

交调（Intermodulation） 虽然谐波失真经常被用来描述模拟电路的非线性，但在某些情况下还需要其他的方式来度量系统的非线性行为。比如设想要评价一个有源低通滤波器的非线性。如图 2.5 所示，如果输入正弦波的频率选择在使它的谐波频率在通带之外，那么即使这个滤波器的输入级引入了很大的非线性，输出失真也会非常小。所以这里需要另一种检测方式，通常采用的是在“双声测试”（two-tone test）中的“交调失真”（intermodulation distortion）

图 2.5 低通滤波器的谐波失真

当两个不同频率的信号通过一个非线性系统时，输出一般会含有一些不属于输入频率谐波的成分。这种现象来源于两个信号的混频（multiplication），叫做交调（intermodulation），此时这两个信号的和出现了高次幂。我们通过假设 $x(t) = A_1 \cos \omega_1 t + A_2 \cos \omega_2 t$ 来理解式（2.11）是如何产生交调的。由此，

$$y(t) = \alpha_1(A_1 \cos \omega_1 t + A_2 \cos \omega_2 t) + \alpha_2(A_1 \cos \omega_1 t + A_2 \cos \omega_2 t)^2 + \alpha_3(A_1 \cos \omega_1 t + A_2 \cos \omega_2 t)^3 \quad (2.21)$$

将上式右边展开并略去直流项和谐波项，我们得到了如下的交调项：

$$\omega = \omega_1 \pm \omega_2 : \alpha_2 A_1 A_2 \cos(\omega_1 + \omega_2)t + \alpha_2 A_1 A_2 \cos(\omega_1 - \omega_2)t \quad (2.22)$$

$$= 2\omega_1 \pm \omega_2 : \frac{3\alpha_3 A_1^2 A_2}{4} \cos(2\omega_1 + \omega_2)t + \frac{3\alpha_3 A_1^2 A_2}{4} \cos(2\omega_1 - \omega_2)t \quad (2.23)$$

$$= 2\omega_2 \pm \omega_1 : \frac{3\alpha_3 A_2^2 A_1}{4} \cos(2\omega_2 + \omega_1)t + \frac{3\alpha_3 A_2^2 A_1}{4} \cos(2\omega_2 - \omega_1)t \quad (2.24)$$

以及如下的基波项：

$$\omega = \omega_1, \omega_2 : (\alpha_1 A_1 + \frac{3}{4} \alpha_3 A_1^3 + \frac{3}{2} \alpha_3 A_1 A_2^2) \cos \omega_1 t + (\alpha_1 A_2 + \frac{3}{4} \alpha_3 A_2^3 + \frac{3}{2} \alpha_3 A_2 A_1^2) \cos \omega_2 t \quad (2.25)$$

我们特别感兴趣的是在 $2\omega_1 - \omega_2$ 和 $2\omega_2 - \omega_1$ 处的三阶 IM 项，如图 2.6 所示。此处的关键点是如果 ω_1 和 ω_2 之间的差距很小，那么 $2\omega_1 - \omega_2$ 和 $2\omega_2 - \omega_1$ 将出现在 ω_1 和 ω_2 附近，这样即使是图 2.5 所示的 LPF（低通滤波器）的情况也可以表现出非线性。在典型的双声测试中， $A_1 = A_2 = A$ ，输出的三阶项的幅度与 $\alpha_1 A$ 的比值定义为交调失真。例如如果 $\alpha_1 A = 1V_{PP}$ ， $3\alpha_3 A^3 / 4 = 10mV_{PP}$ ，那么我们说交调成分为 -40dBc，其中“c”代表“以载波为基准”（with respect to the carrier）。

图 2.6 非线性系统的交叉调制

交调是 RF 系统中一个很让人讨厌的现象。如图 2.7 所示，如果一个弱的信号和两个较强的干扰信号一起经过三阶非线性调制，那么将有一个交调项落入我们感兴趣的频带内，它将破坏有用的成分。当对信号的幅度进行操作时，即使是改变信号相位，交调产生的结果也会降低电路的性能（因为零交叉点仍然受影响）。注意这一现象并不能直接用谐波失真来定量描述。

图 2.7 两个干扰信号之间的交调造成的有用信号的损坏

由于两个邻近的干扰产生的三阶交调对信号的破坏是很普遍和很严重的，所以我们定义了一个性能指标来表征这一现象。这个参数叫做“三阶交调点”(third intercept point -- IP_3)，该参数的测量是通过在一个双声测试中，选择幅度 A 足够小，这样高阶的非线性项可以被忽略，因而增益为常数等于 α_1 。从式 (2.23)、(2.24) 和 (2.25) 我们看到随着 A 的增加，基波与 A 成比例的增加，而三阶交调项与 A^3 成比例的增加[如图 2.8(a)]。画在对数坐标中[如图 2.8(b)]时，IM 项的幅度将以三倍于基波幅度增长的速度增长。而三阶交调点就定义为这两条线的交点。这个点的横坐标叫做输入 IP_3 (IIP_3)，纵坐标叫做输出 IP_3 (OIP_3)。

图 2.8 交调测试中的输出分量的增长

十分重要的是我们可以看到 IP_3 要比简单的 IM 度量好。如果 IM 项的幅度(归一至基波幅度)被用作线性的度量，那么我们必须同时指定测试时的输入幅度。而对于三阶交调点，它本身就可以作为一个线性的度量来比较不同的电路。

从式 (2.11) 的输入—输出特性中我们可以得到一个关于 IP_3 的简单表达式。

假设 $x(t) = A \cos \omega_1 t + A \cos \omega_2 t$ ，那么：

$$y(t) = \left(\alpha_1 + \frac{9}{4}\alpha_3 A^2\right)A \cos \omega_1 t + \left(\alpha_1 + \frac{9}{4}\alpha_3 A^2\right)A \cos \omega_2 t + \frac{3}{4}\alpha_3 A^3 \cos(2\omega_1 - \omega_2)t + \frac{3}{4}\alpha_3 A^3 \cos(2\omega_2 - \omega_1)t + \dots \quad (2.26)$$

如果有 $\alpha_1 \gg 9\alpha_3 A^2 / 4$ ，那么使在 ω_1 、 ω_2 处和 $2\omega_1 - \omega_2$ 、 $2\omega_2 - \omega_1$ 处幅度相等的输入幅度可由下式得到：

$$|\alpha_1| A_{IP_3} = \frac{3}{4} |\alpha_3| A_{IP_3}^3 \quad (2.27)$$

所以输入 IP_3 为：

$$A_{IP_3} = \sqrt{\frac{4}{3} \frac{|\alpha_1|}{|\alpha_3|}} \quad (2.28)$$

输出 IP_3 就等于 $\alpha_1 A_{IP_3}$ 。

参数 IP_3 仅仅描述了三阶非线性的特性。实际中，如果输入幅度增大到达到交调点，那么 $\alpha_1 \gg 9\alpha_3 A^2 / 4$ 的假设将不再成立，增益将下降，高阶的交调项也变得显著。事实上，很多电路的 IP_3 都超出了允许的输入范围，甚至有时超过了电源电压。所以实际测量 IP_3 的方法是测量小输入幅度下的图 2.8(b)的特性，然后在对数坐标下线性外推找到交调点。

下面是一个快速测量 IP_3 的方法：我们用 A_m 表示每个频率上的输入水平，用 A_{ω_1, ω_2} 表示在频率 ω_1 、 ω_2 处的输出幅度，而 IM_3 的幅度用 A_{IM_3} 表示。由式 (2.26) 我们可以得到

$$\frac{A_{\omega_1, \omega_2}}{A_{IM_3}} \approx \frac{|\alpha_1| A_m}{3|\alpha_3| A_m^3 / 4} \quad (2.29)$$

$$= \frac{4|\alpha_1|}{3|\alpha_3| A_m^2} \quad (2.30)$$

结合式 (2.28) 上式可以简化为

$$\frac{A_{\omega_1, \omega_2}}{A_{IM_3}} = \frac{A_{IP_3}^2}{A_m^2} \quad (2.31)$$

最后有：

$$20 \log A_{\omega_1, \omega_2} - 20 \log A_{IM_3} = 20 \log A_{IP_3}^2 - 20 \log A_m^2 \quad (2.32)$$

及

$$20 \log A_{IP_3} = \frac{1}{2} (20 \log A_{\omega_1, \omega_2} - 20 \log A_{IM_3}) + 20 \log A_m \quad (2.33)$$

因此如果所有的信号幅度用 dBm 表示，那么输入三阶交调点就等于基频输出和 IM_3 项输出幅度之差的一半加上相应的输入幅度[图 2.9(a)]。这里的关键是 IP_3 可以只用一个输入幅度来度量，避免了使用外推的方法。

图 2.9 (a) 不需要外推法的 IP_3 计算，(b) (a) 的图示解释

图 2.9(b)所示为上述关系的几何图示。因为直线 L_1 的斜率为 1，而直线 L_2 的斜率为 3，因此输入增加 $\Delta P/2$ 将在 L_1 上产生相等的增量，而在 L_2 上产生 $3\Delta P/2$ 的增量，这使这两条线之间的差距减为 0。

上面的步骤为在设计或表征初期估计 IP_3 提供了方法，但是准确的 IP_3 值仍然必须经过准确地外推得到，以保证所有的非线性效应和与频率相关的效应都被包括在内。

在文献中遇到的另一种测量方法是使用单一声调 (single tone)，绘制三阶谐波幅度关于输入幅度的函数图线，并利用外推得到交调点。从图 2.5 的例子中我们可知这个方法不能得到 IP_3 的准确值。

为了较好的体会实际 RF 系统中所需的线性度，我们来计算一个 $1-\mu V_{rms}$ 信号在一个 IIP_3 是 $70mV_{rms}$ (约等于 -10dBm) 的放大器中被两个 $1-mV_{rms}$ 信号干扰的情形 (如图 2.10)。忽略减敏效应和互调效应，我们可以得到：

$$\frac{A_{sig, out}}{A_{sig, in}} \approx \frac{A_{int, out}}{A_{int, in}} \quad (2.34)$$

其中 A_{sig} 表示信号幅度， A_{int} 是干扰信号的幅度。由式 (2.31) 有：

$$\frac{A_{sig,out}}{A_{IM3,out}} = \frac{A_{sig,in} \cdot A_{IP3}^2}{A_{int,in}^3} \quad (2.35)$$

其中 $A_{sig,in} = 1\mu V_{rms}$ ， $A_{IP3} = 70mV_{rms}$ ， $A_{int,in} = 1mV_{rms}$ 。所以这一比值等于 4.9 (13.8dB)。

图 2.10 在大的干扰信号存在情况下的可得到的 SNR 实例

求出 1dB 压缩点和三阶非线性输入 IP_3 的关系很有启发性。从式 (2.17) 和 (2.28) 我们可得到这两者间的关系为：

$$\frac{A_{1-dB}}{A_{IP3}} = \frac{\sqrt{0.145}}{\sqrt{4/3}} \quad (2.36)$$

$$\approx -9.6dB \quad (2.37)$$

2.1.2 级连非线性级

由于在 RF 系统中信号都是由级联的各级来处理的，所以知道各级的非线性是如何关系到级联的输入是很重要的。特别是，我们希望利用各级的 IP_3 和增益来计算总的输入三阶交调点。

考虑两个非线性系统级联（如图 2.11）。如果这两级的输入—输出特性分别如下式表示：

$$y_1(t) = \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t) \quad (2.38)$$

$$y_2(t) = \beta_1 y_1(t) + \beta_2 y_1^2(t) + \beta_3 y_1^3(t) \quad (2.39)$$

那么

$$\begin{aligned} y_2(t) = & \beta_1 [\alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t)] \\ & + \beta_2 [\alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t)]^2 \\ & + \beta_3 [\alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t)]^3 \end{aligned} \quad (2.40)$$

只考虑一阶和三阶项，我们有

$$y_2(t) = \alpha_1 \beta_1 x(t) + (\alpha_3 \beta_1 + 2\alpha_1 \alpha_2 \beta_2 + \alpha_1^3 \beta_3) x^3(t) + \dots \quad (2.41)$$

所以，根据式 (2.28)

$$A_{IP3} = \sqrt{\frac{4}{3} \left| \frac{\alpha_1 \beta_1}{\alpha_3 \beta_1 + 2\alpha_1 \alpha_2 \beta_2 + \alpha_1^3 \beta_3} \right|} \quad (2.42)$$

有趣的是，合适的选择分母中各项的值和符号可以得到任意高的 IP_3 。但是在实际中其他考虑诸如噪声、增益和有源器件的特性等可能不允许这样的选择。作为最坏情况的估计，我们在分母中给这三项加上了绝对值符号。

图 2.11 级联的非线性电路级

式 (2.42) 两边取倒数平方后可化简为:

$$\frac{1}{A_{IP3}^2} = \frac{3}{4} \frac{|\alpha_3\beta_1| + |2\alpha_1\alpha_2\beta_2| + |\alpha_1^3\beta_3|}{|\alpha_1\beta_1|} \quad (2.43)$$

$$= \frac{1}{A_{IP3,1}^2} + \frac{3\alpha_2\beta_2}{2\beta_1} + \frac{\alpha_1^2}{A_{IP3,2}^2} \quad (2.44)$$

其中 $A_{IP3,1}$ 和 $A_{IP3,2}$ 分别表示第一、第二级的输入 IP_3 点, 注意到 A_{IP3} 、 $A_{IP3,1}$ 和 $A_{IP3,2}$ 是电压量而不是功率量。

从上面的结果可知随着 α_1 的增加, 总的 IP_3 减小。这是因为如果第一级有更高的增益, 第二级就会接收到更高的输入幅度的信号, 所以产生了大得多的 IM_3 项。(回忆一下 IM_3 项随着输入幅度的三次方而增长。)

为了更深刻地理解, 我们假设 $x(t) = A\cos\omega_1t + A\cos\omega_2t$, 并由此来确定各个 IM 项。根据图 2.12, 我们观察到如下结果¹: (1) 输入信号的基波成分在第一级被放大了约 α_1 倍, 在第二级被放大了约 β_1 倍。所以输出的基波成分为 $\alpha_1\beta_1A(\cos\omega_1t + \cos\omega_2t)$ 。(2) 第一级产生的 IM_3 项, 即 $(3\alpha_3/4)A^3[\cos(2\omega_1 - \omega_2)t + \cos(2\omega_2 - \omega_1)t]$, 在第二级输出端同样被放大了 β_1 倍。

(3) 第二级的输入端接收到 $\alpha_1A(\cos\omega_1t + \cos\omega_2t)$, 所以产生下列 IM_3 项: $(3\beta_3/4)(\alpha_1A)^3[\cos(2\omega_1 - \omega_2)t + \cos(2\omega_2 - \omega_1)t]$ 。(4) $y_1(t)$ 中的二阶非线性在频率 $\omega_1 - \omega_2$, $2\omega_1$ 和 $2\omega_2$ 处产生了分量。它们在第二级经历了同样的非线性, 因而转化成 $2\omega_1 - \omega_2$ 和 $2\omega_2 - \omega_1$ 。更具体地说如图 2.12 所示, $y_2(t)$ 包含有诸如 $2\beta_2[\alpha_1A\cos\omega_1t \cdot \alpha_2A^2\cos(\omega_1 - \omega_2)t]$ 和 $2\beta_2(\alpha_1A\cos\omega_2t \cdot 0.5\alpha_2A^2\cos 2\omega_1t)$ 的项。结果中的三阶交调项可以表示为 $(3\alpha_1\alpha_2\beta_2A^3/2)[\cos(2\omega_1 - \omega_2)t + \cos(2\omega_2 - \omega_1)t]$ 。

由以上观察, 我们可以写出:

¹ $A\cos\omega t$ 的频谱由两个脉冲组成, 每个权重为 $A/2$ 。为了简单在图中我们省略了 $1/2$ 。而最后的结果是正确的。

$$\begin{aligned}
y_2(t) &= \alpha_1 \beta_1 A (\cos \omega_1 t + \cos \omega_2 t) \\
&+ \left(\frac{3\alpha_3 \beta_1}{4} + \frac{3\alpha_1^3 \beta_3}{4} + \frac{3\alpha_1 \alpha_2 \beta_2}{2} \right) A^3 [\cos(2\omega_1 - \omega_2)t \\
&+ \cos(2\omega_2 - \omega_1)t] + \dots
\end{aligned} \tag{2.45}$$

于是得到了同上面相同的 IP_3 。

在许多 RF 系统中，级联中每一级的带宽都很窄。所以，上述第四点所描述的分量落在频带之外，被强烈的衰减。结果式 (2.44) 右边的第二项可以被忽略，得到：

$$\frac{1}{A_{IP3}^2} \approx \frac{1}{A_{IP3,1}^2} + \frac{\alpha_1^2}{A_{IP3,2}^2} \tag{2.46}$$

这个式子可以很容易的推广到三级或更多级的级联情况：

$$\frac{1}{A_{IP3}^2} \approx \frac{1}{A_{IP3,1}^2} + \frac{\alpha_1^2}{A_{IP3,2}^2} + \frac{\alpha_1^2 \beta_1^2}{A_{IP3,3}^2} + \dots \tag{2.47}$$

其中 $A_{IP3,3}$ 表示第三级的输入 IP_3 。所以如果每一级都有大于 1 的增益，那么后面级的非线性将变得越来越重要，因为每一级的 IP_3 实际上将以它之前所有级的总增益按比例缩小。我们要强调式 2.47 只是近似式。实际中，必须进行更准确的计算或者模拟来估计总的 IP_3 。

图 2.12 由两个非线性级组成的级联电路中的交调机制

2.2 符号间干扰

一个线性时不变系统如果没有足够的带宽也会使一个信号失真。一个低通滤波器对一个周期方波高频成分的衰减就是熟知的这种情况的一个例子[图 2.13(a)]。但是有限的带宽对随机比特流(random bit stream)会有更大的危害。为了说明这一问题，首先回忆一下当一个理想的矩形脉冲通过一个低通滤波器时，输出会表现出一个指数形式的带尾，而且这一现象随着滤波器带宽的减小而变得更加显著。这一现象是必然的，因为一个信号不能同时受时间限制和带宽限制：当时限脉冲通过带宽限制系统时，输出在时域上必然会延伸无限长的时间。

图 2.13 低通滤波器对 (a) 周期方波，(b) 随机的 1 和 0 系列信号的响应

现在假设一个数字系统的输出是由一个随机的 0、1 序列构成的，每一个 1 代表一个理想的矩形脉冲，每个 0 代表没有这一脉冲。如果这一序列通过一低通滤波器，输出可由每个输入位的响应叠加而成[图 2.13(b)]。我们注意到每个位的电平都被它之前位产生的衰减带尾破坏了。这一现象叫做“符号间干扰”(intersymbol interference--ISI)，它在检测通过有限带宽信道传输的随机波形时会产生较高的误码率。

在无线通讯中 ISI 的问题就更棘手了，因为这时分配给每个信道的带宽相当

窄。此时减小 ISI 的方法有在发射机中的脉冲整形(pulse shaping) (“Nyquist signaling”)和在接收机中的“均衡”(equalization)。我们在这里简要介绍 Nyquist 信号发送, 关于均衡法减小 ISI 请读者参考更多的文献[3, 4]。

为了减小 ISI, 脉冲的形状可以选择成不易受其平移后产生的信号的干扰。在 Nyquist 信号发送中, 每个脉冲都允许和前后的脉冲重叠, 但是脉冲形状选择成使在特定的时间点 ISI 为 0。如图 2.14, 这一想法就是要使所有其他的脉冲在现行脉冲达到它的最高点时都为 0。所以如果比特流在 $t = kT_s$ 时刻采样, 就没有 ISI。

图 2.14 没有 ISI 时的脉冲形状

通过一个简单的计算可以得到 Nyquist 信号发送的基本条件。为使一个脉冲形状 $p(t)$ 不产生任何 ISI。我们有

$$p(kT_s) = 1 \quad \text{如果 } k = 1 \quad (2.48)$$

$$= 0 \quad \text{如果 } k \neq 0 \quad (2.49)$$

采用一个冲击脉冲(impulse)序列来对这个脉冲采样, 我们得到

$$p(t) \cdot \sum \delta(t - kT_s) = \delta(t) \quad (2.50)$$

对两边进行 Fourier 变换, 我们有

$$p(f) * \frac{1}{T_s} \sum \delta(f - \frac{k}{T_s}) = 1 \quad (2.51)$$

就是:

$$\frac{1}{T_s} \sum p(f - \frac{k}{T_s}) = 1 \quad (2.52)$$

这一方法最初由奈奎斯特 (Nyquist) 提出, 如图 2.15 所示, 其结果表明 $p(f)$ 平移后产生的各项相加后必须形成一个平坦的频谱。比如一个 sinc 波形就满足这一条件, 因为它的 Fourier 变换是一个矩形。

图 2.15 不会产生 ISI 脉冲形状频谱的奈奎斯特条件

但是一个 sinc 脉冲波形会使系统设计十分困难。如果需要一个边沿陡峭的矩形频谱, 那么产生这个矩形频谱的滤波器将很难设计。而且在频谱边沿附近实际的信号能量使发射和接受通路上的滤波器设计要求更加复杂。另外, sinc 波形随时间衰减很慢, 当采样命令中存在时序错误时将引入显著的 ISI。

在 Nyquist 采样中常用的一个脉冲波形与“提升余弦”(raised cosine)的频谱有关。如图 2.16 所示, 这一函数的时域、频域表达式分别为:

$$p(t) = \frac{\sin(\pi t / T_s) \cos(\pi \alpha t / T_s)}{\pi t / T_s \quad 1 - 4\alpha^2 t^2 / T_s^2} \quad (2.53)$$

$$\begin{aligned}
\text{及} \quad P(f) &= T_s \quad 0 < |f| < \frac{1-\alpha}{2T_s} \\
&= \frac{T_s}{2} \left[1 + \cos \frac{\pi T_s}{\alpha} \left(|f| - \frac{1-\alpha}{2T_s} \right) \right] \quad \frac{1-\alpha}{2T_s} < |f| < \frac{1+\alpha}{2T_s} \quad (2.54) \\
&= 0 \quad |f| > \frac{1+\alpha}{2T_s}
\end{aligned}$$

图 2.16: 在 (a) 时域, (b) 频域中的提升余弦脉冲

其中 $0 < \alpha < 1$ 叫做“滚降”系数。要注意 (1) $p(t)$ 比 sinc 函数衰减快, (2) 当 $\alpha = 0$ 时, $p(t)$ 退化为一个 sinc 函数, (3) $p(f)$ 与矩形频谱类似, 但是具有光滑的边缘。

α 的选择要在时域的衰减速度和频域的带宽增量 (相对于矩形频谱而言) 之间折衷。 α 的典型值为 0.3 到 0.5。

提升余弦信号发送还可以从图 2.17 中看出。在这里, 每个比特位用一个冲击脉冲来表示, 数据流通过一个滤波器, 该滤波器的传输函数由式 (2.54) 表示。从这一点来看, 这一操作叫做“提升余弦波滤波”。

图 2.17 提升余弦波滤波

在第三章中将解释, 许多应用都采用了传输函数为式 (2.54) 的平方根的滤波器。

2.3 随机过程和噪声

随机过程是整个通信中的一个组成部分, 用来表示信号和噪声。在这一节里, 我们将按照 RF 设计的要求和适合 RF 设计的语言简要地回顾一下随机过程和噪声。主要的目的是建立一个对这些现象和操控这些现象的关系的直观的理解。在这里我们认为读者已经熟悉诸如随机变量和概率密度函数 (PDF) 等概念。

2.3.1 随机过程

随机过程的复杂性在于它们是随机的! 那些习惯于处理定义得很好的, 完全确定的“铁的事实”的工程师们往往觉得随机这个概念很难掌握, 特别是当它必须用数学式子来表达的时候。下面我们从一个直观的角度来看这个问题, 以消除对随机性的恐惧。

我们说一个现象是随机的因为我们不知道或者不需要知道它的所有信息。我们只用几个参数和函数来表征这一过程, 并且不需要任何其他有关这一过程的信息来解决大多数的问题。经验表明在很多应用中, 包括 RF 设计, 这一方法是可行的, 也是足够的。换句话说, 我们很幸运, 因为在 RF 设计中遇到的大多数

随机过程都可以用很简单的模型来描述。

根据我们的目的，一个随机过程（random, 或更确切的 stochastic）可以定义为一组时间的函数。如果我们今天测量一个电阻两端的噪声电压随时间的关系，它将与明天测量的波形不同²（图 2.18）。为了得知有关噪声电压的所有信息，我们要做无穷多个测量，每一个都要无穷长时间。因为对单个波形的测量往往不能提供这一过程足够的信息，即使最简单的随机过程也在两个维度上展开，也就是说它们需要收集一组测量，也即一组时间的函数。这就是随机信号和确定信号的根本区别，也是使我们混淆的根源。在使用一个普通的确定信号源时，我们一般都认为输出是一个可预测的明确定义的波形（除了在上电阶段，而它往往是不重要的）。而对于随机信号，例如通过电话线的声音，我们并不能这样认为，而只能依靠多次测量得到的统计信息。

图 2.18 将噪声看作是一组时间的函数

我们必须强调随机信号理论之所以很实用是因为在大多数我们感兴趣的情况下这些信号可以用简单的统计函数来建立模型，这些函数指出了幅度随时间有多快和多大的变化，当然还有其他的信息。另外，统计模型可以和确定信号和系统的理论联合使用，这往往使我们可以暂时忽略随机性而采用更加直观的分析技术。

那么如何描述一个随机过程的特性呢？它的统计性质中哪些方面是重要的呢？如何在系统分析中使用这些方面呢？我们将首先建立一些简化的假设而后回答这些问题。

统计系综 (Statistical Ensembles) 如上面提到的，要充分表征类似于电阻两端噪声电压这样的（连续时间）随机信号的全部特性需要两个无限测量才能得到：一是无限次数的测量；二是每次测量要有无穷长的时间（图 2.18）。现在我们假设有很多个完全相同的电阻而不是一个，并且同时来测量它们的噪声电压（仍然是经历很长的时间）。我们可以想到这两组实验可以得到相同的结果。这一大量电阻噪声电压的集合叫做“系综(ensemble)”，每一个这样的波形叫做“采样函数”(sample function)。

那么我们如何测量一个电阻噪声电压的平均值呢？我们熟知的方法是在一个长时间 T 内，测量噪声 $n(t)$ ，并用下式来计算平均值（即直流成分）

$$\langle n(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{+T/2} n(t) dt \quad (2.55)$$

这一随机信号直流分量的概念叫做“时间平均值”。

图 2.19 对取样函数进行平均

另一种关于平均值的定义基于同时对一个系综中所有的波形进行采样（图 2.19）。这里我们把采样数据相加并归一至总的波形数目来计算平均值。这一方法叫做“系综平均”或者“统计平均”，可作如下正式定义

$$\overline{n(t)} = \int_{-\infty}^{+\infty} n(t) P_n(n) dn \quad (2.56)$$

² 这一点不要与系统的时变性相混淆。

其中 $P_n(n)$ 是随机过程的概率密度函数。

上面的定义有两个问题。第一，今天测量的时间平均值与明天测量的相同吗？不一定。统计性质不随时间平移改变的随机过程叫做平稳过程（stationary, strict-sense stationary）。因此时间平均的概念适用于平稳过程，比如说在一定的温度下电阻两端的噪声电压。幸运的是在 RF 系统中大多数的随机现象可以认为是平稳的。

第二个问题是：一个平稳过程的时间平均和系综平均相等吗？不一定，但对于本书中大多数我们感兴趣的随机过程我们可以认为是相等的，这样可以避免使用系综平均。

上面定义的时间平均和系综平均都是一阶的。我们还可以定义高阶的平均。我们特别感兴趣的是二阶平均，因为他们体现了信号的功率。在时域：

$$\langle n^2(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{+T/2} n^2(t) dt \quad (2.57)$$

如果 $n(t)$ 是电压量，上式也叫做“均方”功率（mean square power）（对于 1Ω 电阻）。二阶系综平均是

$$\overline{n^2(t)} = \int_{-\infty}^{+\infty} n^2(t) P_n(n) dn \quad (2.58)$$

本书中我们认为 $\langle n^2(t) \rangle = \overline{n^2(t)}$ 。

概率密度函数（Probability Density Function） 当在时域中考虑一个随机信号时，我们一般需要知道它的幅度有多频繁地出现在一定界限内。例如，如果一个二进制数据序列被掺杂的噪声破坏了（图 2.20），那么找出一个逻辑 1 被误解成逻辑 0 或者反过来的概率是多少是很重要的，也就是噪声幅度有多频繁地超过信号幅度的一半。一个随机信号 $x(t)$ 的幅度统计值是由概率密度函数 $P_x(x)$ 表征的，其定义如下：

$$P_x(x) dx = x < X < x + dx \text{ 的概率} \quad (2.59)$$

其中 X 是 $x(t)$ 在一些时间点的测量值。为了估计 PDF，我们在许多点上对 $x(t)$ 采样（在系综中是对许多函数采样），于是建立了许多小宽度的盒子，盒子的高度等于采样值落在这个盒子宽度内的点的数目，并用总的采样点数目归一化盒子的高度。注意 PDF 并不提供随机信号在时域中变化有多快的信息。

图 2.20 二进制信号被噪声所损坏

一个重要的 PDF 的例子就是高斯（或正态）分布（Gaussian distribution）。中心极限定理（central limit theorem）表明如果许多独立的有着任意 PDF 的随机过程叠加，那么总的 PDF 接近于高斯分布。所以很多自然现象都显示出高斯统计特性就不奇怪了。例如，因为电阻的噪声来自大量电子的随机“走动”，每一个都有相对独立的统计特性，总的幅度就遵循高斯 PDF。

高斯 PDF 定义为

$$P_x(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \quad (2.60)$$

其中 σ 和 m 分别是该分布的标准差和平均值。

从一个随机信号幅度的 PDF 我们还可以回答下面的问题：如果采样数目很大，落在 x_1 和 x_2 之间的百分比是多少？这一问题可由 x_1 到 x_2 之间 $P_x(x)$ 覆盖的面积给出，对于高斯 PDF：

$$P(x_1 < x < x_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] dx \quad (2.61)$$

对于有限的 x_1 和 x_2 ，上式右边的积分必须用数值方法来计算。这一积分的一个较简单的形式，叫做误差函数，它的值在许多参考书中都列出了表格：

$$\text{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left[-\frac{u^2}{2}\right] du \quad (2.62)$$

要记住对于一个高斯分布，有将近 68% 的采样值落在 $m-\sigma$ 和 $m+\sigma$ 之间，有 99% 落在 $m-3\sigma$ 和 $m+3\sigma$ 之间。

功率谱密度 (Power Spectral Density) 由于我们对于时域中随机信号的了解通常非常有限，所以常常也需要在频域中表征这些信号。实际上，正如我们将在本书中看到的，在 RF 设计中随机信号和噪声在频域的行为比它们的时域特性有用的多。

对于一个确定的信号 $x(t)$ ，频率信息包含在其 Fourier 变换中：

$$X(f) = \int_{-\infty}^{+\infty} x(t) \exp(-j2\pi ft) dt \quad (2.63)$$

虽然对随机信号似乎很自然也可以用相同的定义，但我们必须指出 Fourier 变换只对于有限能量的信号才适用³：

$$E_x = \int_{-\infty}^{+\infty} |x(t)|^2 dt < \infty \quad (2.64)$$

也就是只有当 $|x(t)|^2$ 在 $t \rightarrow \infty$ 时有足够快的衰减才成立。如图 2.21 所示，这一条件对两类信号不成立：周期信号和随机信号。但是在大多数情况下，这些波形具有有限的功率：

$$P = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{+T/2} |x(t)|^2 dt < \infty \quad (2.65)$$

对于 $P < \infty$ 的周期信号，其 Fourier 变换仍然可以通过将每个 Fourier 级数中的分量用频域的冲击脉冲函数表示的方法来定义。反之对于随机信号，因为频域中的冲击脉冲函数相当于存在一个确定的正弦分量，所以这个方法一般是不可能的。另一个实际的问题是即使我们设法对一个随机过程（平稳或者非平稳过程）定义了一个 Fourier 变换，其结果本身仍然是个随机过程[5]。

图 2.21 具有无穷大能量的信号

从上面的讨论中我们推断随机信号的频域特性包含在一个不同于直接 Fourier 变换的函数中。功率谱密度(PSD)（也叫做谱密度）就是这样一个函数。在给出 PSD 的正式定义之前，我们从一个直观的角度来介绍一下它的意义[6]。

³ 若 $x(t)$ 是加在 1 欧姆电阻上的电压则很容易看出这是能量的定义。

一个随机信号 $x(t)$ 的谱密度 $S_x(f)$ 表示在频率 f 周围单位带宽范围内信号的功率。如图 2.22 所示，如果我们使这个信号通过中心频率为 f 、带宽为 1-Hz 的带通滤波器，并测量够长时间（1 秒量级）上的平均输出功率，我们将得到 $S_x(f)$ 的估计值。如果对每个 f 值都进行这一测量，我们就得到了整个的信号功率谱。这正是谱分析仪的工作原理⁴。

图 2.22 频谱测量

PSD 正式的定义如下[3]:

$$S_x(f) = \lim_{T \rightarrow \infty} \frac{\overline{|X_T(f)|^2}}{T} \quad (2.66)$$

其中

$$X_T(f) = \int_0^T x(t) \exp(-j2\pi ft) dt \quad (2.67)$$

这个定义可以通过一个相应的计算算法来理解（图 2.23）：（1）将信号 $x(t)$ 用一个较长的时间段 $[0, T]$ 截断；（2）计算该结果的 Fourier 变换并由此得到 $|X_T(f)|^2$ ；（3）对 $x(t)$ 的许多采样函数重复步骤 1, 2（例如测量电阻的多组噪声电压波形）；（4）对所有的 $|X_T(f)|^2$ 求平均得到 $\overline{|X_T(f)|^2}$ ，并用 T 归一化。这个算法在利用随机噪声波形的时域模拟中是很有用的。

图 2.23 进行功率谱密度估算的算法

由于对于实数的 $x(t)$ ， $S_x(f)$ 是关于 f 的偶函数[3]，正如图 2.24(a)所示， $x(t)$ 在频率区间 $[f_1, f_2]$ 内携带的总功率为

$$\int_{-f_2}^{-f_1} S_x(f) df + \int_{f_1}^{f_2} S_x(f) df = \int_{f_1}^{f_2} 2S_x(f) df \quad (2.68)$$

实际上上式右端的积分是由谱分析仪测量的结果，即谱的负频率分量部分沿纵轴折叠，然后与正频率部分相加[图 2.24(b)]。我们把图 2.24(a) 的表示方法叫做“双边” (two-sided) 谱，图 2.24(b) 的表示叫做“单边” (one-sided) 谱。

图 2.24 (a) 双边及 (b) 单边频谱

在用图形分析频域操作时，用双边谱比较方便，而实际噪声的计算用单边谱更容易。然而，虽然这两种表示方法会引起混乱，但是它们之间并没有本质的区别。

⁴ 构造一个 1Hz 带宽、中心频率在例如 1GHz 的低损耗带通滤波器是不切实际的。因此实际的谱分析仪把功率谱变换至较低的中心频率并测量此 1Hz 宽的频带上的功率。

作为 $S_x(f)$ 的一个例子，我们考虑阻值为 R 的一个电阻两端的热噪声电压。双边 PSD 为

$$S_x(f) = 2kTR \quad (2.69)$$

其中 k 是波尔兹曼(Boltzmann)常数，等于 $1.38 \times 10^{-23} \text{ J/K}$ ； T 是绝对温度。这种平坦的谱叫做白噪声谱，因为它在所有的频率上有相同的功率。

式 (2.69) 引出了两个有趣的问题。第一，一个电阻总的噪声功率 ($S_x(f)$ 覆盖的面积) 是无限的吗？实际上， $S_x(f)$ 只是在 $|f| < 100 \text{ GHz}$ 时才为平坦的，超过这一频率时 $S_x(f)$ 会降低，所以总的功率还是有限的[3]。第二， $2kTR$ 的单位是功率每单位带宽吗 (W/Hz)？不是，实际的单位是均方电压每单位带宽。我们只是约定假设这个电压加在一个 1 欧姆的电阻两端，在 1Hz 带宽下产生 $2kTR$ 的功率。在电路噪声计算中，我们经常写成

$$\overline{V_n^2} = 4kTR \cdot \Delta f \quad (2.70)$$

其中 $\overline{V_n^2}$ 是电阻 R 在带宽 Δf 下产生的均方噪声电压。对于 $\Delta f = 1 \text{ Hz}$ $\overline{V_n^2}$ 叫做“点噪声” (spot noise)，它的单位是 V^2/Hz 。

总结一下 PDF 和 PSD 的概念，我们看到前者是一个随机过程的幅度有多大的机率落在一个给定范围内的统计量，而后者是指在一个小的频率区间内信号会有多大功率的期望值。总的来说，PDF 和 PSD 间没有什么关系：热噪声有高斯分布的 PDF 和白噪声分布的 PSD，而闪烁($1/f$)噪声有相同的 PDF，而其 PSD 与 $1/f$ 成比例。

线性系统中的随机信号(Random Signals in Linear Systems) 定义功率谱密度函数的根本原因是它使许多用于确定信号的频域操作也可以应用到随机过程上。可以证明，如果一个谱密度为 $S_x(f)$ 的信号加到一个传输函数为 $H(s)$ 的线性时不变系统上 (如图 2.25)，那么其输出谱为

$$S_y(f) = S_x(f) |H(f)|^2 \quad (2.71)$$

其中 $H(f) = H(s = j2\pi f)$ [3]。这符合我们的直觉，即信号谱的形状将受到系统传输函数的调制。还可以看到如果 $x(t)$ 是高斯分布，那么 $y(t)$ 也是[3]。

图 2.25 线性系统中的噪声整形

2.3.2 噪声

噪声可以笼统地定义为任何与有用信号无关的随机干扰⁵。这个定义区分了噪声和诸如谐波失真和交调等确定性的现象。同其他随机过程一样，噪声特性用 PDF 和 PSD 描述。

存在于所有电路中的是热噪声，它是由电阻、双极型器件的基极和发射极电阻、以及 MOSFET 的沟道电阻产生的[如图 2.26(a)]。MOS 器件的热噪声可以模拟成一个接在源漏之间的电流源，其 PSD 为

$$\overline{I_n^2} = 4kT\left(\frac{2}{3}g_m\right) \quad (2.72)$$

其中 g_m 是晶体管的跨导。由长沟器件导出的系数 2/3 [7]在沟道长度小于 1 微米时需要采用更大的值[8]。注意 MOSFET 的栅分布电阻也会有热噪声[9]，但是可以通过仔细的版图设计而将它的影响减到最小。

图 2.26 器件中的 (a) 热噪声与 (b) 散粒噪声

除了热噪声，有源器件还有散粒噪声(shot)和闪烁噪声(flicker)。散粒噪声是高斯分布的白噪声，由电荷越过一定的能量势垒（如 pn 结）形成，其 PSD 为

$$\overline{I_n^2} = 2qI \quad (2.73)$$

其中 q 为电子电荷， I 为平均电流。对于双极型晶体管，收集极和基极电流的散粒噪声可以模拟成分别连接在收集极和发射极、基极和发射极之间的电流源[图 2.26(b)]。

闪烁噪声由电荷在 MOSFET 的二氧化硅与硅界面上被随机俘获而产生。它可以模拟为与栅串连的一个电压源，其噪声密度为

$$\overline{V_n^2} = \frac{K}{WLC_{ox}} \frac{1}{f} \quad (2.74)$$

其中 K 是与随机过程有关的常数。虽然闪烁噪声的影响在高频下似乎可以忽略，但我们还必须注意到在诸如混频器或者振荡器等电路中的非线性或者时变性可以将 $1/f$ 形状的谱映射到 RF 域（第六、七章）。

输入端参考噪声(Input-Referred Noise) 两端口系统的噪声可以用两个输入噪声源来描述：一个串联电压源和一个并联电流源（图 2.27）[10]。总的来说，必须要考虑这两个电源的相关性。我们用一个例子来说明这个概念。考虑如图 2.28(a)所示的电路，其中我们假设已有合适的偏置保证 M1 饱和且其漏极电流为 I_D 。这个电路只有一个主要的热噪声源：它由沟道产生，并用 $\overline{I_{nD}^2}$ 表示。对于图

2.28(b)所示的模型，我们通过短路输入端求得 $\overline{V_n^2}$ ，开路输入端求得 $\overline{I_n^2}$ 。因为图

2.28(a)和(b)应该在这两种情况下产生相同的输出噪声，我们有 $g_m^2 \overline{V_n^2} = \overline{I_{nD}^2}$ 以及

$g_m^2 \overline{I_n^2} |Z_{in}|^2 = \overline{I_{nD}^2}$ ，其中 Z_{in} 表示该电路的输入阻抗。于是因 $\overline{I_{nD}^2} = 4kT(2g_m/3)$ ，我

⁵ 常常说要是没有任何噪声的话，也就不会有任何模拟电路设计者了。

们得到 $\overline{V_n^2} = 8kT/(3g_m)$ 以及 $\overline{I_n^2} = 8kT(3g_m|Z_{in}|^2)$ 。由于 V_n 和 I_n 表示相同的噪声机制，它们之间一定相关。

图 2.27 采用输入端噪声发生器的噪声表示

图 2.28 (a) MOS 放大器, (b) 等效输入端噪声发生器

我们注意如果 $|Z_{in}| \rightarrow \infty, \overline{I_n^2} \rightarrow 0$, 这时 $\overline{V_n^2}$ 就足以表示这个噪声。然而在射频范围, $|Z_{in}|$ 很小(一些设计中要求为 50 欧姆左右), 所以就需要同时使用 $\overline{V_n^2}$ 和 $\overline{I_n^2}$ 。

上面例子的关键在于即使实际的电路中没有真正的输入噪声电流, 等效电路的输入相关噪声源中也要包含有 $\overline{I_n^2}$ 。

噪声系数(Noise Figure) 在许多模拟电路中, 信噪比(SNR)定义为信号功率与总的噪声功率之比, 这是一个很重要的参数。而在 RF 设计中, 即使最终的目标是最大化接收到和检测到的信号的 SNR, 许多前端接收模块却用它们的“噪声系数”而不是输入相关噪声来描述。这一方面是为了计算方便, 另一方面因为这是一直以来的传统。

噪声系数有很多种不同的定义方式, 最常用的是

$$\text{noise figure} = \frac{SNR_{in}}{SNR_{out}} \quad (2.75)$$

其中 SNR_{in} 和 SNR_{out} 是分别在输入和输出端测量的信噪比。注意在大多数书中上面的式子叫做噪声因子(noise factor), 而噪声系数定义为 $10\log_{10}(\text{noise factor})$ 。我们在这里不做这种区别。

从物理意义上理解式 (2.75) 是很重要的, 噪声系数是信号通过一个系统后 SNR 变差了多少的度量。如果一个系统没有噪声, 那么无论增益怎样都有 $SNR_{in} = SNR_{out}$ 。这是因为输入信号和输入噪声都按相同的比例放大(或者减小), 而且没有引入额外的噪声。所以, 一个无噪声系统的噪声系数等于 1。实际中, 一个系统存在一定的噪声使 SNR 减小, 因此 $NF > 1$ 。

与输入相关噪声相比, 式 (2.75) 定义的 NF 看起来很复杂: 它不仅与所考虑电路的噪声有关, 还与前级的 SNR 有关。实际上, 如果输入信号不含噪声, 那么即使电路可能含有有限的内部噪声, 也有 $SNR_{in} = \infty, NF = \infty$ 。对于这种情况, NF 不是一个有意义的参数。反之在 RF 设计中, 这种情况是不可能的, 因为即使是接收器第一级的信号也会被天线辐射电阻的噪声破坏。

噪声系数的计算往往比式 (2.75) 所示的简单一些。如图 2.29 所示, 我们假设 SNR_{in} 是输入信号功率和源电阻 R_S 产生的噪声的比, 该噪声模型为 $\overline{V_{RS}^2}$ 。如果从 V_{in} 到电路输入端(节点 P)的电压增益等于 α , 那么在这个节点上测得的 SNR

为

$$SNR_{in} = \frac{\alpha^2 V_{in}^2}{\alpha^2 V_{RS}^2} \quad (2.76)$$

设从 P 节点到 V_{out} 的电压增益为 A_v ，则在输出端测得的 SNR 为

$$SNR_{out} = \frac{\alpha^2 A_v^2 V_{in}^2}{[V_{RS}^2 + (V_n + I_n R_S)^2] \alpha^2 A_v^2} \quad (2.77)$$

$$= \frac{V_{in}^2}{[V_{RS}^2 + (V_n + I_n R_S)^2]} \quad (2.78)$$

其中 V_n 和 $I_n R_S$ 在取平方前相加以考虑它们的相关性。于是有

$$NF = \frac{\overline{V_{RS}^2 + (V_n + I_n R_S)^2}}{V_{RS}^2} \quad (2.79)$$

$$= 1 + \frac{\overline{(V_n + I_n R_S)^2}}{V_{RS}^2} \quad (2.80)$$

NF 一般都是针对某一频率周围的 1Hz 带宽来说的，为了强调非常小的带宽它往往叫做点噪声系数。这一数值可从式 (2.80) 得到：

$$NF = 1 + \frac{\overline{(V_n + I_n R_S)^2}}{4kTR_S} \quad (2.81)$$

其中 V_n 和 I_n 也是在相同频率的单位带宽下测量的。

图 2.29 噪声系数的计算

式 (2.81) 指出了噪声系数是源阻抗 R_S 的函数。一般的，已知一个给定源阻抗的 NF 并不能求出另一个不同源阻抗的 NF ，因为在 (2.81) 分式中 R_S 同时出现在分子和分母中。然而在传统的 RF 系统中，大多数的模块都设计成具有 50 欧姆的输入输出电阻⁶（以及可忽略的电抗值），所以避免了 NF 计算中的不确定。但如我们将要在第六章中看到的，这一问题在特定的情况下还需要注意。

为了模拟需要，式 (2.81) 写成如下形式较为方便：

$$NF = \frac{4kTR_S + \overline{(V_n + I_n R_S)^2}}{4kTR_S} \quad (2.82)$$

$$= \frac{A^2 [4kTR_S + \overline{(V_n + I_n R_S)^2}]}{A^2} \frac{1}{4kTR_S} \quad (2.83)$$

$$= \frac{V_{n,out}^2}{A^2} \frac{1}{4kTR_S} \quad (2.84)$$

⁶ 在 TV 系统中，特征阻抗和终端阻抗均为 75 欧姆。

其中 $A = \alpha A_v$, 而 $V_{n,out}^2$ 表示输出端总的噪声。所以为了计算图 2.29 的 NF, 我们将总的输出噪声功率除以从 V_{in} 到 V_{out} 的电压增益的平方, 并将这个结果用 R_S 的噪声来归一化。

作为一个噪声系数计算的例子, 考虑一个电阻 R_p , 如图 2.30(a)。对源电阻 R_S 这个电路的噪声系数是多少呢? 根据图 2.30(b), 总的输出噪声电压为

$$V_{n,out}^2 = 4kT(R_S \parallel R_p) \quad (2.85)$$

而增益为

$$A = \frac{R_p}{R_S + R_p} \quad (2.86)$$

由式 (2.84) 得到

$$NF = 4kT(R_S \parallel R_p) \frac{(R_S + R_p)^2}{R_p^2} \frac{1}{4kTR_S} \quad (2.87)$$

$$= 1 + \frac{R_S}{R_p} \quad (2.88)$$

所以最大化 R_p 可以最小化噪声系数。有趣的是, 最小化噪声系数的条件和最大化功率传输的条件($R_S=R_p$)不相同。

图 2.30 (a) 电阻 R_p 噪声系数的计算, (b) (a) 的等效电路

作为另一个例子, 我们计算图 2.31(a)所示的放大器的噪声系数。这个电路由一个共源级和一个反馈源跟随器构成以提供等于 R_S 的输入电阻。忽略体效应、沟道长度调制效应、寄生电容以及 I_1 的噪声, 我们利用如图 2.31(b)的等效电路来计算总的输出噪声。 $R_{in}=R_S$ 的条件等效于

$$R_S = \frac{1}{g_{m2}} \frac{1}{1 + g_{m1}R_D} \quad (2.89)$$

M_2 的噪声电流流过 $R_S/2$, 产生了等于 $(I_{n2}R_S/2)g_{m1}R_D$ 的输出噪声电压。 R_D 和 M_1 的噪声电流被该电路的输出电阻放大, 当 $R_{in}=R_S$ 时, 该输出电阻可简化为

$$R_{out} = \frac{R_D}{2}(1 + g_{m2}R_S) \quad (2.90)$$

因此总的输出噪声功率为

$$\begin{aligned} V_{n,out}^2 &= 4kTR_S \left(\frac{1}{4} g_{m1}^2 R_D^2 \right) + \frac{1}{4} I_{n2}^2 R_S^2 g_{m1}^2 R_D^2 \\ &+ (I_{RD}^2 + I_{n1}^2) \frac{R_D^2}{4} (1 + g_{m2}R_S)^2 \end{aligned} \quad (2.91)$$

$$\begin{aligned}
&= kTR_S g_{m1}^2 R_D^2 + \frac{2}{3} kT g_{m2} R_S^2 g_{m1}^2 R_D^2 \\
&+ \left(\frac{4kT}{R_D} + \frac{8}{3} kT g_{m1} \right) \frac{R_D^2}{4} (1 + g_{m2} R_S)^2
\end{aligned} \tag{2.92}$$

所以当符合 $g_{m2} R_S = (1 + g_{m1} R_D)^{-1}$ 的条件时噪声系数为

$$NF = \frac{V_{n.out}^2}{A^2} \frac{1}{4kTR_S} \tag{2.93}$$

$$= 1 + \frac{2}{3} g_{m2} R_S + \left(\frac{1}{R_D} + \frac{2}{3} g_{m1} \right) \frac{(1 + g_{m2} R_S)^2}{g_{m1}^2 R_S} \tag{2.94}$$

图 2.31 (a) 输入匹配了的反馈放大器, (b) (a) 的噪声等效电路

级连级的噪声系数(Noise Figure of Cascaded Stages) 对于级联级, 总的噪声系数可由每一级的 NF 和增益得到。考虑图 2.32 所示系统, 其中输入噪声产生源以及每一级的输入输出电阻如图所示。注意阻抗的电抗部分为 0, A_{v1} 和 A_{v2} 表示这两级的无负载电压增益。第一级输入端的总噪声功率可以写作:

$$V_{n.in1}^2 = \overline{[I_{n1}(R_S \parallel R_{in1}) + V_{n1} \frac{R_{in1}}{R_{in1} + R_S}]^2} + \overline{V_{RS}^2} \frac{R_{in1}^2}{(R_{in1} + R_S)^2} \tag{2.95}$$

同样, 第二级的总输入噪声功率为

$$\begin{aligned}
V_{n.in2}^2 &= \overline{V_{n.in1}^2} A_{v1}^2 \frac{R_{in2}^2}{(R_{in2} + R_{out1})^2} \\
&+ \overline{[I_{n2}(R_{out1} \parallel R_{in2}) + V_{n2} \frac{R_{in2}}{R_{in2} + R_{out1}}]^2}
\end{aligned} \tag{2.96}$$

所以, 级联的总输出噪声功率等于 $A_{v2}^2 V_{n.in2}^2 R_L^2 / (R_L + R_{out2})^2$ 。由于从 V_{in} 到 V_{out} 的总电压增益为

$$A_{v.tot} = \frac{R_{in1}}{R_S + R_{in1}} A_{v1} \frac{R_{in2}}{R_{out1} + R_{in2}} A_{v2} \frac{R_L}{R_{out2} + R_L} \tag{2.97}$$

总的噪声系数为

$$NF_{tot} = \frac{1}{A_{v.tot}^2} A_{v2}^2 \left(\frac{R_L}{R_{out2} + R_L} \right)^2 V_{n.in2}^2 \frac{1}{4kTR_S} \tag{2.98}$$

利用式 (2.95) 和 (2.96), 化简结果有

$$NF_{tot} = \frac{4kTR_S + \overline{(I_{n1}R_S + V_{n1})^2}}{4kTR_S} + \frac{\overline{(I_{n2}R_{out1} + V_{n2})^2}}{A_{v1}^2} \frac{1}{\left(\frac{R_{in1}}{R_S + R_{in1}}\right)^2} \frac{1}{4kTR_S} \quad (2.99)$$

右边的第一项可看作是第一级当输入阻抗为 R_S 时的 NF, 但第二项却不那么明显。在 $R_S = R_{in1} = R_{out1} = R_{in2}$ 的特定情况下, 我们有

$$NF_{tot} = NF_1 + \frac{\overline{(I_{n2}R_S + V_{n2})^2}}{\alpha^2 A_{v1}^2} \frac{1}{4kTR_S} \quad (2.100)$$

$$= NF_1 + \frac{NF_2 - 1}{\alpha^2 A_{v1}^2} \quad (2.101)$$

其中 NF_2 是第二级当输入阻抗为 R_S 时的噪声系数。

图 2.32 带有噪声的电路级的级联

在一般情形中, 我们采用“可用功率增益”(available power gain) A_P 的概念来化简式 (2.99)。这种增益被定义为输出端可得到的功率 (即该电路可以传输给共轭匹配负载的功率) 除以可得到的源功率 (即电源可以传输给共轭匹配电路的功率)。图 2.32 中第一级可得到的输出功率为

$$P_{out,av} = V_{in}^2 \left(\frac{R_{in1}}{R_S + R_{in1}}\right)^2 A_{v1}^2 \frac{1}{4R_{out1}} \quad (2.102)$$

而可得到的源功率为

$$P_{source,av} = \frac{V_{in}^2}{4R_S} \quad (2.103)$$

所以

$$A_P = \left(\frac{R_{in1}}{R_S + R_{in1}}\right)^2 A_{v1}^2 \frac{R_S}{R_{out1}} \quad (2.104)$$

注意到第二级当源阻抗为 R_{out1} 时的噪声系数为

$$NF_{2,R_{out1}} = 1 + \frac{\overline{(I_{n2}R_{out1} + V_{n2})^2}}{4kTR_{out1}} \quad (2.105)$$

我们可将式 (2.99) 写作

$$NF_{tot} = NF_{1,RS} + \frac{NF_{2,R_{out1}} - 1}{A_P} \quad (2.106)$$

其中 $NF_{1,RS}$ 表示第一级当源阻抗为 R_S 时的噪声系数。

类似的，对于 m 级级联，

$$NF_{tot} = 1 + (NF_1 - 1) + \frac{NF_2 - 1}{A_{p1}} + \dots + \frac{NF_m - 1}{A_{p1} \cdots A_{p(m-1)}} \quad (2.107)$$

其中每一级的 NF 都是根据驱动该级的源阻抗计算的。该式叫做 Friis 方程[11]。该方程用每一级的噪声系数来表示总的噪声系数，尤其在当一个接收器使用的是由制造商独立描述的各种不同的现成功能块时，这一方程非常有用。

Friis 方程指出了在前一级的增益增加时，后面各级贡献的噪声减小了，说明了级联级中前几级是最重要的。相反的，如果某一级的作用是衰减，那么该级后面的电路等效到该级输入处的噪声系数将被“放大”。比如如果在一个接收器中一个窄带有损滤波器放在天线和低噪声放大器之间来消除带外干扰时，将发生上述情况（第五章）。

有损耗电路的噪声系数(Noise Figure of Lossy Circuits) 在 RF 接收器中使用的无源滤波器存在一定的带内损耗。除了衰减有用信号外，有损电路一般还将引入噪声，当我们考虑例如一个理想 LC 滤波器[图 2.33(a)]时，这个事实并不明显。但是回想一下，许多 RF 电路都要求有明确确定的阻性输入和输出阻抗以正确匹配。比如说在天线和 LNA 之间放置的滤波器可以看成如图 2.33(b)所示。我们的目的是找出具有阻性输入输出阻抗的无源电路的噪声系数与损耗之间的关系。

图 2.33 (a) LC 衰减器，(b) 在输入和输出端匹配的有损电路

考虑一个如图 2.34(a)所示的具有实输入输出阻抗的线性时不变无源互易网络(reciprocal network)。可以证明[5]如果输出电阻是 R_{out} ，那么噪声的戴维宁(Thevenin)等效电路就如图 2.34(b)所示，其电压源的 PSD 为 $4kTR_{out}$ 。注意 R_{out} 并且因此 $\overline{V_n^2}$ 一般来说依赖于源阻抗 R_S 。

图 2.34 (a) 无源互易网络，(b) (a) 的戴维宁噪声模型

现在考虑图 2.35(a)所示的电路。与可获得功率增益的概念类似，我们定义功率损耗 L 为 P_{in}/P_{out} ，其中 P_{in} 是可获得的源功率， P_{out} 为可获得输出功率。为了计算信号损失，电路可模拟成 Thevenin 等效电路形式，如图 2.35(b)所示。注意这个电路对于信号是等效的，而在图 2.34(b)中的电路是对噪声等效的。因为

$P_{in} = \overline{V_{in}^2} / (4R_S)$ 及 $P_{out} = V_{TH}^2 / (4R_{out})$ ，我们有

$$L = \frac{V_{in}^2}{V_{TH}^2} \frac{R_{out}}{R_S} \quad (2.108)$$

为了计算噪声系数，我们求出输出噪声电压和电压增益。当负载为 R_L 时，由图 2.34(b)得到

$$V_{n,out}^2 = 4kTR_{out} \frac{R_L^2}{(R_L + R_{out})^2} \quad (2.109)$$

图 2.34(b)中从 V_{in} 到 V_{out} 的电压增益为

$$A_v = \frac{V_{TH}}{V_{in}} \frac{R_L}{R_L + R_{out}} \quad (2.110)$$

所以

$$NF = 4kTR_{out} \frac{V_{in}^2}{V_{TH}^2} \frac{1}{4kTR_S} \quad (2.111)$$

$$= L \quad (2.112)$$

因此我们得出对于无源互易网络噪声系数等于如上面定义的损耗。

图 2.35 (a) 用于噪声系数计算的电路, (b) (a) 的等效电路

为了强调上面的概念, 我们计算一个有损滤波器后面接一个低噪声放大器时的总的噪声系数 (图 2.36)。由 Friis 方程有

$$NF_{tot} = NF_{filt} + \frac{NF_{LNA} - 1}{L^{-1}} \quad (2.113)$$

$$= L + (NF_{LNA} - 1)L \quad (2.114)$$

$$= L \cdot NF_{LNA} \quad (2.115)$$

其中 LNA 的噪声系数的计算是相对于滤波器的输出电阻的。

图 2.36 带损耗滤波器和 LNA 的级联

2.4 灵敏度与动态范围

灵敏度(Sensitivity) RF 接收器的灵敏度定义为在可接受的信噪比要求下系统可检测到的最小信号电平。为了计算灵敏度, 我们有

$$NF = \frac{SNR_{in}}{SNR_{out}} \quad (2.116)$$

$$= \frac{P_{sig} / P_{RS}}{SNR_{out}} \quad (2.117)$$

其中 P_{sig} 表示输入信号功率, P_{RS} 表示源电阻噪声功率, 以上都是对应单位带宽。

由此有

$$P_{sig} = P_{RS} \cdot NF \cdot SNR_{out} \quad (2.118)$$

既然总的信号功率分布在信道带宽 B 的范围内, 式 (2.118) 两边都必须在这个带宽内积分以得到总的均方功率。所以, 对于频带平坦的信道

$$P_{sig,tot} = P_{RS} \cdot NF \cdot SNR_{out} \cdot B \quad (2.119)$$

方程 (2.119) 把灵敏度估计为产生一定输出 SNR 的最小输入信号。稍微改变表示法, 并用 dB 或 dBm 表示上面的量, 我们有

$$P_{in,min} |_{dBm} = P_{RS} |_{dBm/Hz} + NF |_{dB} + SNR_{min} |_{dB} + 10 \log B \quad (2.120)$$

其中 $P_{in,min}$ 是达到 SNR_{min} 的最小输入电平, 其中 B 用赫兹表示。注意式 (2.120) 与系统的增益无关。

假设输入是共轭匹配的, 我们得到 R_S 传输给接收器的噪声功率 P_{RS} :

$$P_{RS} = \frac{4kTR_S}{4} \frac{1}{R_{in}} \quad (2.121)$$

$$= kT \quad (2.122)$$

$$= -174 \text{dBm/Hz} \quad (2.123)$$

上式假设在室温下。于是我们得到式 (2.120) 的简化形式

$$P_{in,min} = -174 \text{dBm/Hz} + NF + 10 \log B + SNR_{min} \quad (2.124)$$

注意前三项的和是系统总的累积噪声, 有时被称作“噪声底限”(noise floor)。因为 $P_{in,min}$ 是带宽的函数, 接收器由于有较窄的信道而非常灵敏(但自然是以较低的数据率为代价)。

动态范围(Dynamic Range) 动态范围(DR)一般定义为电路可接受的最大输入电平与电路可提供合理信号质量时最小输入电平的比值。这一定义在不同的应用中有不同的定量。比如, 在诸如运算放大器和模数转换器等模拟电路中动态范围定义为“满程”(full-scale)输入电平与使 $SNR=1$ 的输入电平之比。满程一般是指输入电平超过该范围时将发生深饱和, 并可以很容易的通过检查电路发现, 而最小输入电平由噪声底限决定。

反之, 在 RF 设计中, 情况就更复杂了。设想一个简单的共源放大级。我们如何定义其满程输入呢? 有可能将 FS 定义为晶体管处于线性区边界时的输入电压。但是, 如果一个满程电压摆幅的正弦波作用在该电路上, 输出将会有很严重的失真。另外, 最小的信号一定要使 SNR 大于 1, 比如在式 (2.120) 中的 SNR_{min} 。由于以上原因, 我们根据交调的限制制定动态范围的上限, 而根据灵敏度的要求制定其下限。这个定义叫做“厖空动态范围”(spurious-free dynamic range)(SFDR)。

动态范围的上限定义为在双声测试中使得三阶交调项不超过噪声底限的最大输入电平。用 dBm 表示所有的量, 式 (2.33) 可改写为

$$P_{IIP3} = P_{in} + \frac{P_{out} - P_{IM,out}}{2} \quad (2.125)$$

其中 $P_{IM,out}$ 表示输出端 IM_3 分量的功率。因为 $P_{out} = P_{in} + G$ 及 $P_{IM,out} = P_{IM,in} + G$, 其中 G 是用 dB 表示的电路功率增益, $P_{IM,in}$ 是输入相关的 IM_3 项的大小, 我们有

$$P_{IIP3} = P_{in} + \frac{P_{in} - P_{IM,in}}{2} \quad (2.126)$$

$$= \frac{3P_{in} - P_{IM,in}}{2} \quad (2.127)$$

因此有

$$P_{in} = \frac{2P_{IP3} + P_{IM,in}}{3} \quad (2.128)$$

于是使得 IM 项等于噪声底限的输入电平为

$$P_{in,max} = \frac{2P_{IP3} + F}{3} \quad (2.129)$$

其中 $F = -174\text{dBm} + NF + 10\log B$ 。

SFDR 是 $P_{in,max}$ 和 $P_{in,min}$ 的差（用 dB 来表示）

$$SFDR = \frac{2P_{IP3} + F}{3} - (F + SNR_{min}) \quad (2.130)$$

$$= \frac{2(P_{IP3} - F)}{3} - SNR_{min} \quad (2.131)$$

比如，如果一个接收器的 $NF=9\text{dB}$ ， $P_{IP3} = -15\text{dBm}$ ，且 $B=200\text{kHz}$ ，若要求 $SNR_{min}=12\text{dB}$ ，那么它的 SFDR 约为 53dB。

SFDR 表示在一个较小的输入电平下，一个接收器在产生可接受信号质量的情况下能容忍的相对最大干扰信号。

2.5 无源阻抗变换

在射频范围，我们经常利用无源电路来进行阻抗变换—从高变到低，或者反过来；以及从复变到实，或者反过来。虽然有源器件似乎也能用来实现这样的操作，但是在一些情况下只有无源元件才能达到所要求的性能。比如，一个 3V 的功率放大器可以最多传输 $P = 3^2 / (2 \times 50) = 90\text{mW}$ 的功率给一个 50 欧姆的负载。为了增加这个输出功率，必须在 PA 和负载之间增加一个电路来“放大”电压摆幅而不需要更高的电源电压（第九章）。这是由无源匹配电路实现的。

图 2.37 等效串联与并联 RC 电路

在学习变换技术之前，我们看一下图 2.37 所示的 RC 电路。这一串联组合的品质因数 Q 定义为电容的阻抗除以电阻，即等于 $1/R_s C_s \omega$ ，当 R_s 趋于 0 的时候它趋于无穷大。类似的，并联组合的 Q 等于 $R_p C_p \omega$ 。如果 Q 相当大（约大于 5）而且有用频带很窄，那么一种网络可以转化成另一种。这两个电路在下面的条件

下是等效的

$$\frac{R_p}{R_p C_p s + 1} = \frac{R_s C_s s + 1}{C_s s} \quad (2.132)$$

或者对于 $s = j\omega$, 有 $R_p C_p = 1/(R_s C_s \omega^2)$, $R_p C_p + R_s C_s - R_p C_s = 0$ 。假设 $R_p \gg R_s$, 我们有 $C_p \approx C_s$ 以及

$$R_p \approx \frac{1}{R_s (C\omega)^2} \quad (2.133)$$

其中 $C = C_p \approx C_s$ 。所以这种转换将电阻值按照式 (2.133) 变化, 而电容值基本不变。我们还可以写出 $R_p \approx Q_s^2 R_s$, 其中 Q_s 是串联网络的 Q 。对于相应的 RL 网络可以得到类似的结果。

阻抗变换可以通过变压器来完成。一个具有匝数比 m 的理想变压器将阻抗按照 m^2 的比例变换。但是实际上, 高频变压器常常产生损耗以及初级和次级之间的电容耦合, 引起不希望的谐振, 从而使设计复杂化并要求精确的模型。所以我们要研究其他的阻抗变换方法。

考虑图 2.38(a)所示的电路, 其中的电容分压器用来将 R_p 变换成一个更高的值。我们假设高 Q 值以及窄频带, R_p 和 C_p 的并联组合可以被转换成如图 2.38(b)所示的串联电路, 其中 C_s 约等于 C_p , $R_s \approx 1/[R_p (C_p \omega)^2]$ 。将 C_1 和 C_s 合成为 C_{eq} , 我们得到了如图 2.38(c)所示的电路, 它可以转化成如图 2.38(d)所示的并联电路, 其中 $C_{tot} \approx C_1 C_p / (C_1 + C_p)$, $R_{tot} \approx 1/[R_s (C_{eq} \omega)^2] = (1 + C_p / C_1)^2 R_p$ 。所以电容分压器将 R_p 的值提高了 $(1 + C_p / C_1)^2$ 倍。

图 2.38 采用电容分压器的阻抗变换

图 2.39(a)描述了一个类似的使用电感分压器构成的转换电路, 对于图 2.39(b)的等效电路, 如果 Q 值足够高, 所考虑的频带够窄, 我们有 $L_{tot} \approx L_1 + L_p$, $R_{tot} \approx (1 + L_1 / L_p)^2 R_p$ 。

图 2.39 采用电感分压器的阻抗变换

如图 2.40(a)所示是一种常用的将一电阻值转换成较低值的电路。将 R_p 和 C_p 变成如图 2.40(b)的串联组合, 我们有 $C_p \approx C_s$ 及 $R_s \approx 1/[R_p (C_p \omega)^2]$ 。在谐振点附

近， C_s 和 L_1 谐振，因此该电路相当于一个阻值为 $1/[R_p(C_p\omega)^2]$ 的电阻。

图 2.40 电阻降阻变换

我们要注意高 Q 值的假设并不一定成立，尤其是使用片上电感时。所以，在这些情况下要检查上面推导的准确性。

参考文献

- [1] D. D. Wiener and J. F. Spine, *Sinusoidal Analysis and Modeling of Weakly Nonlinear Circuits*, New York: Van Nostrand Reinhold, 1980.
- [2] S. A. Maas, *Nonlinear Microwave Circuits*, Norwood, MA: Artech House, 1988.
- [3] L. W. Couch, *Digital and Analog Communication Systems*, 4th ed., New York: Macmillan, 1993.
- [4] J. G. Proakis and M. Salehi, *Communication System Engineering*, Englewood Cliffs, NJ: Prentice Hall, 1994.
- [5] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed., New York: McGraw-Hill, 1991.
- [6] R. W. Bennet, "Methods of Solving Noise Problems," *Proc. IRE*, Vol.44, pp.609-638, May 1956.
- [7] A. van der Ziel, "Thermal Noise in Field Effect Transistors," *Proc. IRE*, Vol. 50, pp.1808-1812, August 1962.
- [8] A. A. Abidi, "High-Frequency Noise Measurements on FETs with Small Dimensions," *IEEE Trans. Electron Devices*, Vol. 33, pp. 1801-1805, November 1986.
- [9] B. Razavi, "Impact of Distributed Gate Resistance on the Performance of MOS Devices," *IEEE Trans. Circuits and Systems—Part I*, Vol. 41, pp. 750-754, November 1994.
- [10] P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed., New York: John Wiley, 1993.
- [11] H. T. Friis, "Noise Figure of Radio Receivers," *Proc. IRE*, Vol. 32, pp.419-422, July 1944.

第3章 调制与解调

调制和解调模拟信号与数字信号是通讯系统的基本功能。虽然实现这些功能的新方法在不断涌现，但是一些技术已趋成熟并在许多射频电路当中得到了应用。

本章主要介绍了现代射频电路收发器中采用的调制和解调的基本技术，描述了每种方法的优点和局限，并着重强调了在设计射频集成电路中的各种联系。首先，我们简要介绍调制方法的主要特点，然后回顾模拟的幅度、相位和频率调制。接下来我们研究二相和正交调制技术，并强调它们在射频电路中的作用。最后，我们将介绍各种调制技术的功率效率和非相干解调技术。

3.1 概述

射频通讯中传递的波形通常是经过原始信号调制的高频载波。进行调制的原因有很多：(1)在有线系统中，同轴线在高频时表现出极好的屏蔽特性；(2)在无线系统中，为了得到合理的增益，天线尺寸只能是信号波长的一小部分；(3)在大多数情况下，由于 FCC(美国联邦电讯委员会)的规定，通讯系统只能在某一个频段工作；(4)在某些应用中，当信道含有非理想因素时，调制可以使接收端的解调更加简单。

就调制而言，我们通常定义两种信号。信号频谱在 $\omega = 0$ 附近非 0 而在其它部分可以忽略[1][Fig.3.1(a)]的信号叫做基带信号，如麦克风和摄像机产生的信号。信号的频谱在载波信号频率 ω_c 附近的频带非 0 而在该频带之外可以忽略[1][Fig.3.1(b)]的信号叫做带通信号¹。这样的定义虽然在数学上不是很精确，但对于我们的用途来说已经足够了，因为在射频设计中有用的信号带宽通常要比频率小的多。

图 3.1 (a) 基带和 (b) 带通信号

调制就是把基带信号转化为相应的带通信号。从另一个角度来说，调制就是根据基带信号来改变载波信号(例如一个正弦信号)的某些参数。这两个定义是等价的，因为一个带通信号总可以表示为：

$$x(t) = a(t) \cos[\omega_c t + \theta(t)] \quad (3.1)$$

式中 $a(t)$ 和 $\theta(t)$ 一般都与时间有关[1]。换言之，我们起始于一个周期的载波信号 $A \cos(\omega_c t + \theta)$ 并改变它的幅度和相位，这就是调制。(3.1)中，自变量 $\omega_c(t) + \theta(t)$ 称为“全相位”，而 $\theta(t)$ 称为“额外相位”。我们也可以把即时频率定义为相位相对

¹ 我们还会用到“传输频带”。带通信号由基带信号和低通传输频带组成。

于时间的导数： $\omega_c + d\theta/dt$ 是“全频率”， $d\theta/dt$ 是“额外频率”或者“频偏”。

注意，取决于调制技术，一个带宽为 B 的基带信号被转化为带通波形时，它可能会占据不一样的带宽，这点将在第九章中阐明。

和调制相反的过程就是解调(或者说检波)，它的目的是以最小的噪声、失真和 ISI 等把原始信号提取出来。因此，如图 3.2，一个简单的通讯系统包括调制器/发送器，信道(例如空间或者同轴线)，接收器/检波器。因为调制器和检波器的性能决定了整体性能，所以调制技术往往是与相应的解调器一起学习的。我们把两者的组合叫做调制解调器。

图 3.2 简单的通信系统

调制解调器的主要性能 分析图 3.2 中的系统，首要目的就是求出当信道存在衰减和下级解调器的输入存在干扰噪声的情况下它的输出信号的“品质”(例如信噪比)。当信道衰减(或者叫做“路径损失”)和传输功率一定时，信号的品质由调制解调器的种类决定。这样，如果调制解调器有更高的噪声容限，我们就可以减少传输功率，为移动设备提供更长的“通话时间”，或者能够容忍更高的信道衰减，就可以提供更长的通讯距离。

调制解调技术的另一个重要性能就是带宽需求，假设基带信号具有某一带宽，我们可以计算出每一情况下被调制波形所占据的带宽，这个度量标准被称之为“谱效率”。它在为无线网络这类带宽有限的应用场合选择调制解调器时起着重要的作用。

在射频设计中，要考特别考虑的调制技术的第三个性质就是“功率效率”，它关系到能够应用在发信机中的功率放大器的种类。在这一章的后面我们将会看到某些被调制波形可以依靠非线性放大器来处理而不发生频带扩展或损失信息，但是另外一些则要用线性放大器来处理。由于非线性功率放大器一般比相应的线性功率放大器具有更高的效率(第 9 章)，因此我们希望采用的调制技术能够更适合于非线性放大。

附加白高斯噪声信道 为了比较不同接收器的“噪声容限”，我们假设传输信号的功率恒定，而接收信号受到一个附加的功率谱密度为 $N_0/2$ (图 3.3)的高斯分布白噪声的干扰。加入白噪声来考虑接收器的内部噪声是合适的。但邻近干扰源所引入的噪声通常并不是白噪声，因此需要单独考虑。更进一步，在某些解调器(例如包络检波器)中的噪声是相乘而不是相加的。

图 3.3 附加的高斯噪声假定为具有白色频谱

3.2 模拟调制

3.2.1 幅度调制

对于一个基带信号 $x_{BB}(t)$ ，可以通过把(3.1)式改写为 $x_{AM}(t) = A_c[1 + mx_{BB}(t)]\cos\omega_c t$ 来构成一个幅度调制（AM，调幅）波形，这里 m 叫做“调制指数”。图 3.4 说明了一种产生幅度调制信号的方法并给出了相应的波形和频谱：乘以 $\cos\omega_c t$ 就是将 $x_{BB}(t)$ 的频谱转变为在 $\pm\omega_c$ 附近的频带。因此 $x_{AM}(t)$ 的带宽就是 $x_{BB}(t)$ 带宽的两倍。信号 $A_c \cos\omega_c t$ 是由“本地振荡器” (LO)产生的。

图 3.4 在 (a) 时域，(b) 频域里的幅度调制

一个幅度调制信号可以通过一个乘法器和一个低通滤波器实现解调。如图 3.5(a)所示，这一技术把一个幅度调制频谱转换回基带频谱。这里要注意：接收器必须保证本地振荡器产生的频率和 $x_{AM}(t)$ 的载波频率相同。

如果 $1 + mx_{BB}(t)$ 这项对于任何 t 都是正的，那么图 3.4(a)所示的波形包络就不会超过 0，因此我们可以用一个峰漏检波器来对这样的信号进行解调[图 3.5(b)]。这个电路的结构非常简单，因此它已成为调幅收音机的主要选择。

图 3.5 AM 解调器

以上两种检波器输出信号的信噪比都可以表示为：

$$SNR_{out} = \frac{A_c^2 m^2 \overline{x_{BB}^2(t)}}{2N_0 B}, \quad (3.2)$$

其中 $\overline{x_{BB}^2(t)}$ 和 B 分别是信号 $x_{BB}(t)$ 的均方值和带宽[1]。在采用包络检波器的情况下，式(3.2)仅在输入端信号有足够高信噪比时才成立，这是一个合理的假定，因为只有当信噪比大于 25dB 时，接收信号的质量才是可以接受的。注意：传输的功率等于 $A_c^2[1 + m^2 \overline{x_{BB}^2(t)}]/2$ 。

除了在广播收音机和电视的声音传输中有应用以外，幅度调制在目前的无线系统中的应用很少。这是因为用振幅来装载信息不但使信号易被噪声干扰而且要求发送器采用高线性度的功率放大器。

3.2.2 相位调制与频率调制

相位调制(PM)和频率调制(FM)是十分重要的概念,我们不仅在分析调制解调器的时候要用到它们,而且在分析诸如振荡器和频率综合器这类电路的时候也要遇到它们。

在式(3.1)中,当 $a(t)$ 是常数 A_c ,额外相位 $\theta(t)$ 与基带信号成正比时,我们说载波被相位调制,即 $x_{PM}(t) = A_c \cos[\omega_c t + mx_{BB}(t)]$,其中 m 是相位调制指数。例如,如果 $x_{BB}(t) = \alpha t$,那么 $x_{PM}(t) = A_c \cos(\omega_c + m\alpha)t$,这时意味着一个斜坡基带波形使载波频率移动了一个常数值。

类似的,如果额外频率, $d\theta/dt$,与基带信号成正比,我们说载波被频率调制,即 $x_{FM}(t) = A_c \cos[\omega_c t + m \int_{-\infty}^t x_{BB}(t) dt]$,其中 m 是频率调制指数。例如,对于一个直流基带信号, $x_{BB}(t) = A$,我们得到 $x_{FM}(t) = A_c \cos(\omega_c + mA)t$ 。这样一个频率调制器的直流输入信号使载波频率移动了一个常数值,对于压控振荡器的波形这是一个非常有用的结果(第7章)。

图3.6给出了一个基带信号以及它相应的相位调制和频率调制波形。注意一般当基带信号未知时,我们无法判断一个波形是频率调制还是相位调制的。还有,频偏的最大值在相位调制下等于 mdx_{BB}/dt ,在频率调制下等于 $mx_{BB}(t)$ 。

$x_{PM}(t)$ 和 $x_{FM}(t)$ 与 $x_{BB}(t)$ 间的非线性关系使我们很难计算出调制后波形所占据的带宽。我们在稍后将再回过头来讨论这个问题。

图 3.6 相应于方波基带信号的调相和调频波形

模拟频率调制比模拟相位调制被更广泛的采用,这也许是因为频率调制和解调更容易一些。图3.7说明了一个简单的频率调制器,其中 LC 振荡器的谐振频率随着基带信号而成比例的变化。这里所需要的压控电容通常是一个反偏的变容二极管,它的节电容随 $x_{BB}(t)$ 而变化。

图 3.7 简单的调频器

原则上说,频率解调可以通过任何一个传输函数对频率灵敏的电路来实现。例如,一个高通滤波器(也就是一个微分器)能够产生一个与信号频率成正比的输出信号幅度(图3.8),这实质上就是由频率调制转化为幅度调制。如果

$v_{in}(t) = A_c \cos[\omega_c t + m \int x_{BB}(t) dt]$,则输出信号可以近似为输入信号对时间的导数:

$$v_{out}(t) \approx A_c R_1 C_1 [\omega_c + mx_{BB}(t)] \sin[\omega_c t + m \int x_{BB}(t) dt]. \quad (3.3)$$

这样,如果 $\omega_c + mx_{BB}(t)$ 总保持为正值,我们就可以用一个包络检波器来恢复信号 $x_{BB}(t)$ 。特别注意,输入信号 v_{in} 的幅度变化会使解调波形变差,因此要在微分器之前使用“限制器”。

图 3.8 简单的调频解调器

$x_{FM}(t)$ 与 $x_{BB}(t)$ 间的非线性关系使我们很难计算出调制后波形的带宽。首先，我们来考虑一个简单情况，称为“窄带频率调制”，即这时信号满足： $m \int x_{BB}(t) dt \ll 1 \text{ rad}$ 。则频率调制波形可以近似表示为：

$$x_{FM,NB}(t) \approx A_c \cos \omega_c t - A_c m (\sin \omega_c t) \int x_{BB}(t) dt \quad (3.4)$$

因此，如图 3.9(a)所示，它的频谱包括载波以及 $\int x_{BB}(t) dt$ 转变到 $\pm \omega_c$ 附近的频谱，这意味着 $x_{FM}(t)$ 的带宽比图 3.4(a)中 $x_{AM}(t)$ 的带宽小一些，这是由积分引起的 $1/s$ 衰减造成的。

图 3.9 为 (a) 随机信号和 (b) 正弦信号调制的窄带调频

证明窄带频率调制非常有用的一个特例是正弦基带信号，例如 $x_{BB}(t) = A_m \cos \omega_m t$ ，此时我们有：

$$x_{FM,NB}(t) \approx A_c \cos \omega_c t - A_m A_c \frac{m}{\omega_m} \sin \omega_c t \sin \omega_m t \quad (3.5)$$

$$\begin{aligned} &= A_c \cos \omega_c t - \frac{A_m A_c m}{2\omega_m} \cos(\omega_c - \omega_m)t \\ &\quad + \frac{A_m A_c m}{2\omega_m} \cos(\omega_c + \omega_m)t \end{aligned} \quad (3.6)$$

如图 3.9(b)所示，调制后的频谱由 $\pm \omega_c$ 处的冲击和 $\pm(\omega_c \pm \omega_m)$ 处的“边带”组成。

由方程(3.5)和(3.6) 我们可以总结出三个要点。第一，为使窄带频率调制成立，我们必须有 $m A_m / \omega_m \ll 1 \text{ rad}$ 。第二，当 ω_m 增大时，边带幅度减小；也就是说较高的调制频率在频率调制信号中的权重较小。第三，如果 $x_{FM}(t) = A_c \cos[\omega_c t - (m A_m / \omega_m) \sin \omega_m t]$ ，那么最大的频偏等于 $m A_m$ 。注意，边带和载波频率之间的距离是 ω_m ，并且它和最大频偏值无关。

虽然窄带频率调制在很多通讯电路中是一个非常有用的概念，但是由于它的信噪比很低，因此在调制解调器中的应用很少，这个问题我们在后面说明。接下来我们讨论更加实际的情况，我们考虑并不限制 $m A_m / \omega_m \ll 1 \text{ rad}$ 时，一个正弦基带波形的频率调制信号。如果 $x_{BB}(t) = A_m \cos \omega_m t$ ，那么 $x_{FM}(t) = A_c \cos[\omega_c t - (m A_m / \omega_m) \sin \omega_m t]$ ，它可以展开为：

$$x_{FM}(t) = A_c \sum_{n=-\infty}^{+\infty} J_n(\beta) \cos(\omega_c + n \omega_m)t, \quad (3.7)$$

$J_n(\cdot)$ 是 n 阶的第一类贝塞耳函数， $\beta = m A_m / \omega_m$ 。当 $\beta \ll 1 \text{ rad}$ 时(窄带频率调制)， $J_0(\beta) \approx 1$ ， $J_{\pm 1}(\beta) \approx \pm \beta / 2$ ， $J_n(\beta) \approx 0$ ($|n| > 1$)。这时有， $x_{FM}(t) = A_c \cos \omega_c t + (A_c \beta / 2) \cos(\omega_c + \omega_m)t - (A_c \beta / 2) \cos(\omega_c - \omega_m)t$ ，和式(3.6)中的结果一样。如果 β 接近或者超过 1，则 $|n| > 1$ 时 $J_n(\beta)$ 的系数不可以忽略，这样就会

在 $\omega_c \pm n\omega_m$ 处引入分量。图 3.10 画出了 $\beta = 1$ 和 $\beta = 2$ 时频率调制频谱的例子。更一般情况的频率调制如多重频率调制或高斯调制在文献[2]中讨论。

从上面的分析我们可以得出一个重要结论，频率调制信号的带宽由 m 、 A_m 和 ω_m 决定。事实上，如果带宽 B_{FM} 定义为包含 98% 信号功率的带宽时， $B_{FM} \approx 2(\beta + 1)B_{BB}$ ，其中 B_{BB} 是基带信号的带宽[1]。这称为 Carson 规则。

图 3.10 相应于两个不同 β 值的 FM 频谱

由式(3.6)，我们注意到随着调制信号频率的增加，在频率调制波形中对应分量的幅度减小。因此这时频率调制的解调器通过增大高频增益来补偿这个滚降，如图 3.8 的微分器所示。但是这将导致高频噪声也被放大，从而降低总的信噪比。基于这个原因，大多数频率调制的发送器都通过预先放大高频端来对基带频谱整形(“预矫正”)(图 3.11)。在接收器中则使检测到的信号通过低通滤波器来去除预矫正(“去矫正”)，并同时抑制高频噪声。如果最初的基带频谱具有低通形状时，这种方法通常并不显著增加所要求的带宽。

图 3.11 调频 (FM) 信号的预矫正和去矫正

对于一个没有采用预矫正和去矫正的频率调制检波器，我们可以得到：

$$\frac{SNR_{out}}{SNR_{in}} = 6\beta^2(\beta + 1) \frac{\overline{x_{BB}^2(t)}}{V_p^2} \quad (3.8)$$

其中 V_p 是信号 $x_{BB}(t)$ 的峰值[2]。如果采用预矫正和去矫正，则有：

$$\frac{SNR_{out}}{SNR_{in}} = 2\beta^2(\beta + 1) \left(\frac{B}{f_1}\right)^2 \frac{\overline{x_{BB}^2(t)}}{V_p^2} \quad (3.9)$$

这里 B 是基带信号的带宽， f_1 是用来去矫正的低通滤波器的-3dB 拐角频率。在一些典型应用中，由式(3.9)得到的值比式(3.8)大 10 至 15dB[1]。上述两种情况下的性能和 β 有很大关系，这就揭示了信号品质和带宽效率之间的相互制约关系。

我们要考虑的频率调制系统的最后一个性能是功率效率。因为频率调制信号有固定的包络并且在零交越点上载有信息，因此它可以用非线性功率放大器处理而不影响基带信号和增加所要求的带宽。这样，频率调制信号非常适合于高效率（不可避免为非线性）的功率放大器。

3.3 数字调制

在“数字”射频系统中，载波被一个数字基带信号所调制。正如我们将要在第四章中介绍的，数字调制相比模拟调制具有很多优点，因此在无线系统中被广泛采用。

模拟调制中的幅度调制，相位调制和频率调制在数字系统中分别对应幅移键控 (ASK)，相移键控 (PSK) 和频移键控 (FSK)。图 3.12 举例说明了一个二相基带信号的

波形。在射频应用中，相移键控和频移键控对幅度噪声具有较低的敏感度，因此它们比幅移键控具有更广泛的应用。

在 3.1 节中我们曾经介绍，分析一个调制解调器需要定量三个参数的值：信号质量，谱效率和功率效率。数字信号传送方法的优劣用比特误差率(BER)来衡量，它定义为单位时间内检波器输出端观察到的错误比特的平均数和接收到的全部比特数的比值。因此，我们的目的就是计算存在噪声和其它干扰的情况下信号的**错误几率**。

图 3.12 (a) 幅偏，(b) 相偏，(c) 频偏键控

3.3.1 基本概念

分析数字调制解调器需要了解一些在我们比较熟悉的模拟调制系统中没有用到的概念工具。这里我们回顾其中一些最基本的概念。

二进制和 M 进制信号传输 我们会很自然的选择一个二进制波形作为数字基带信号，它可以表示为：

$$x_{BB}(t) = \sum_n b_n p(t - nT_b), \quad (3.10)$$

式中 b_n 是在时间段 $[nT_b, (n+1)T_b]$ 中的比特值，它可取两个值，例如 0 和 1 或者 -1 和 1 中的一个。对于一个矩形脉冲波形，接收器只要简单的判断波形(峰值)是小于还是大于阈值电压就可以确定其比特值。

在实际中，采用多级数字信号来调制载波往往会有好处，这通常是因为这可以放松对带宽的要求。图 3.13 所示的一个例子中，原有的二进制数据流被分成两个比特一组，每组由数模转换器转变为四个可能等级中的一个。我们注意到由此得到的波形在单位时间内的变化次数减少，但是要求检波器具有较高的幅度分辨率。这一技术成为 M 进制信号传输。它也可以用式(3.10)表示，只要 b_n 取所有的离散电平。为了不和二进制信号传输相混淆，我们把发送的每一个电平称为一个码元而不是比特。图 3.13 中信号的“码元率”是原有比特率的一半。

图 3.13 比特数据流的四级数字表示

基函数 对于式(3.10)表示的基带信号，我们现在来考虑它的数字波形的表示方法。例如，图 3.12(c)中的二进制频移键控信号可以表示为：

$$x_{FSK}(t) = A_c \cos \omega_1 t, \quad \text{if } b_n = 0, \quad (3.11)$$

$$= A_c \cos \omega_2 t, \quad \text{if } b_n = 1, \quad (3.12)$$

这说明每一个传输的比特都可以用一个内积来表示，

$$x_{FSK}(t) = \alpha_1 \phi_1(t) + \alpha_2 \phi_2(t) \quad (3.13)$$

$$= [\alpha_1 \quad \alpha_2] \cdot [\phi_1(t) \quad \phi_2(t)], \quad (3.14)$$

其中 $[\alpha_1, \alpha_2] = [0A_c]$ 或者 $[A_c 0]$ ， $\phi_1(t) = \cos \omega_1 t$ ， $\phi_2(t) = \cos \omega_2 t$ 。换句话说， $x_{FSK}(t)$ 可以用两个函数的线性组合来表示。

这一把一个波形用一组基函数的线性组合来表示的方法应当很熟悉，它和傅立叶级数类似： $x(t)$ (具有有限能量) 可以表示为：

$$x(t) = \sum_{n=0}^{\infty} a_n \cos n\omega t + \sum_{n=0}^{\infty} b_n \sin n\omega t, \quad (3.15)$$

其中 $t_0 < t < t_0 + T$ ， $\omega = 1/T$ 。傅立叶级数的一个重要性质就是基函数的正交性，

$$\int_0^T \cos n\omega t \cos m\omega t = 0, \quad (3.16)$$

$$\int_0^T \cos n\omega t \sin m\omega t = 0, \quad (3.17)$$

$$\int_0^T \sin n\omega t \sin m\omega t = 0, \quad (3.18)$$

以上各式当 $n \neq m$ 时成立。

通过上面的观察，我们现在可以得出一个结论：在数字调制中的每个码元可以用一组正交基函数的线性组合来表示，²

$$x(t) = \alpha_1 \phi_1(t) + \alpha_2 \phi_2(t) + \cdots + \alpha_N \phi_N(t), \quad (3.19)$$

其中

$$\int_0^{T_s} \phi_m(t) \phi_k(t) dt = 0 \quad (3.20)$$

式中 $m \neq k$ ， T_s 是码元的周期。数 N 叫做这组基函数的“维数”。

到这里，读者可能会问为什么我们要做以上的推导。这可以从两方面来理解。第一，把信号波形用一组基函数的线性组合表示，就可以使它们有相同的处理方法，并且可以很容易把许多看来不同的调制技术进行合理的比较。第二，正交基函数提供了一个表示波形的“最小组合”，更重要的是，它们可以用来优化信号的检测。

信号星座图 当维数小于 4 时，我们可以把调制后的信号看成它们内积中系数的形式。这就是表示成 $x(t) = [\alpha_1 \alpha_2 \cdots] \cdot [\phi_1(t) \phi_2(t) \cdots]$ ，我们可以简单的忽略基函数，并把向量 $[\alpha_1 \alpha_2 \cdots]$ 的所有可能值在笛卡尔坐标上画出来。这样，图 3.12(c) 中的二进制频移键控信号可以用图 3.14(a) 来表示。这样的图称为“信号星座图”或者“信号空间图”，它对于我们理解数字调制解调器中的各种影响非常有价值。

图 3.14 (a) FSK, (b) ASK 信号星座

作为另外一个例子，考虑图 3.12(a) 中的二进制幅移键控波形，这里每个比特被调制后的信号可以表示为

$$x_{ASK}(t) = A_c \cos \omega_c t \quad b_n = 1, \quad (3.21)$$

$$= 0 \quad b_n = 0 \quad (3.22)$$

因此， $x_{ASK}(t) = \alpha_1 \phi_1(t)$ ，这里 α_1 的值为 0 或 A_c ， $\phi_1(t) = \cos \omega_c t$ 。这是一维调制的情况，它的信号星座图表示如图 3.14(b)。

² 基函数通常都被归一化使它们在 T_b 时间段上具有单位能量，但这里我们先略去这一细节。

信号星座图的一个直接的应用就是来帮助理解噪声对检波接收信号的影响。假设如图 3.15 所示, 一个二进制幅移键控信号为某一噪声信号 $n(t)$ 叠加, 并通过一个包络检波器。这样得到的信号虽然类似于原来的基带波形, 但是却包含有相当的幅度噪声。因此, 需要增加一个模拟比较器来检测并产生一个含较少噪声的输出信号, 这个模拟比较器的阈值处在预期的最高和最低电平值的中间。³

图 3.15 噪声对于 (a) ASK 信号和 (b) 其星座表示的影响

如果噪声的方均根值增大会发生什么事情? 只要噪声的幅度足够大, 包络检波器的输出便会偶尔错误的超过阈值电压, 这样就会导致比较器的输出产生一个错误的比特值。换句话说, 如果噪声幅度和最高最低电平差的一半可以比拟时, 出错的几率便会迅速上升。

为了观察噪声对信号星座图的影响, 我们长时间的观察包络检波器的输出电平, 并得到一个在各点附近的“噪声带” [图 3.15(b)]。噪声幅度变大会使检测到的电平变化范围增大, 因此增加了错误几率, 而增大两点之间的距离(例如更大的信号幅度)可以减少这个错误几率。因此, 在星座图中点之间的坐标距离标志着被调制信号抵抗噪声的能力。事实上, 具有最小距离的两点对比特误差率的影响最大。我们也可以直观的看出如果图 3.15(b)中两个电平出现的几率相等, 则作决定的阈值电压值应当为 $A_c/2$ 。

作为另一个例子, 考虑图 3.14(a)所示的二进制(正交)频移键控在存在噪声时的星座图。所得到的信号空间如图 3.16 所示。如果噪声功率变大, “信号云”超过了决定边界时, 检波器就会出错。

图 3.16 噪声对 FSK 星座的影响

最优检波 把基带信号从被调制波形中提取出来的办法对整个系统的性能, 特别是对信号质量有很大的影响。由于基带信号是数字信号, 因此检波器的输出必须每比特周期 T_b 被采样, 以决定接收到的数值。如图 3.17 所示。

图 3.17 使用采样法的信号解调

为了理解最优化检波过程中碰到的问题, 我们从图 3.18 所示的简单情况入手。假设一个单脉冲(不一定经过调制)在存在附加噪声的情况下被检测。我们如何为这个脉冲整形才能在采样瞬间得到最大的信噪比? 我们可以假设如果 $h(t) = \delta(t)$ 而且采样同步使脉冲的峰值能被检测到时, 输出的信噪比最大[图 3.19(a)]。

图 3.18 滤波器在解调器中的应用

³ 这一检波技术并不是最优的, 这里是用来说明一个观点。

然而我们可以利用噪声的随机性在一个比特周期内取信号的平均值。如图 3.19(b)所示的例子，先把输入脉冲在时间 0 到 T_b 内积分，然后在 $t = T_b$ 时采样。这时，在一个周期内变化剧烈的噪声就能被平均掉了。

图 3.19 采用 (a) 峰值取样，(b) 在取样前对一个比特周期进行积分的解调器

图 3.19(b)的例子给我们引入了“匹配滤波器”的概念。可以证明对于一个受附加白噪声破坏的脉冲信号 $p(t)$ ，总存在一个最优的滤波器使在采样瞬间的信噪比最大[1]，这一电路叫做匹配滤波器(图 3.20)，它的冲击响应为 $h(t) = p^*(T_b - t)$ ，式中右边是由 $p(t)$ 取共轭然后对时间取反，再向右移动 T_b 得到的。在我们感兴趣的大多数情况下， $p(t)$ 是实数，所以 $h(t) = p(T_b - t)$ 。注意， $h(t)$ 和 $p(t)$ 的傅立叶变换的幅度是相等的。

图 3.20 采用匹配滤波的解调器

为了对匹配滤波器有一个直观的认识，我们假设 $p(t)$ 是一个如图 3.21 所示的矩形脉冲。在没有噪声的情况下，滤波器输出为：

$$y(t) = p(t) * h(t) \quad (3.23)$$

$$= \int_{-\infty}^{+\infty} p(t - \tau)h(\tau)d\tau, \quad (3.24)$$

这是一个在 $t = T_b$ 时具有最大值的三角脉冲。我们注意到这是一个积分器的冲击响应(在 $t = T_b$ 时取输出)。这肯定了我们在前面用积分平均噪声的想法。

如果一个匹配滤波器的输入 包含有一个脉冲 $p(t)$ 和一个功率密度为 $N_0/2$ 的附加白噪声，则最大信噪比等于[1]

$$SNR_{\max} = \frac{2E_p}{N_0}, \quad (3.25)$$

其中

$$E_p = \int_{-\infty}^{+\infty} |p(t)|^2 dt. \quad (3.26)$$

我们说 E_p 是信号的能量。如果 $p(t)$ 表示电压，则 E_p 是一个 1Ω 的电阻被 $p(t)$ 驱动时所消耗的能量。方程(3.25)可以导出两个有趣的结论。第一， SNR_{\max} 与脉冲的能量有关，却与它的形状无关。第二， SNR_{\max} 与脉冲或匹配滤波器占据的带宽无关。

图 3.21 矩形脉冲的最优化解调

根据以上的推导，我们现在可以来研究如何对一个被调制信号进行最优检波。首先我们认识到如果相应于一个比特的输入信号是一个实函数 $p(t)$ ，那么匹配滤波器在这个比特周期结束时的输出可以写作：

$$y(T_b) = \int_{-\infty}^{+\infty} x(\tau)h(T_b - \tau)d\tau \quad (3.27)$$

$$= \int_{-\infty}^{+\infty} x(\tau)p(\tau)d\tau, \quad (3.28)$$

其中 $x(t) = p(t) + n(t)$ 。如果 $p(t)$ 在时间段 $[0T_b]$ 以外为 0，则有

$$y(T_b) = \int_{-\infty}^{\tau=T_b} x(\tau)p(\tau)d\tau. \quad (3.29)$$

方程(3.29)是一个很重要的结果，它意味着输入信号同时(包括信号和噪声)和已知的脉冲波形 $p(t)$ 在一个比特周期内必须进行“相关运算”。如图 3.22，当 $p(t)$ 为矩形脉冲时，这个技术就简化为图 3.21 的拓扑结构。

图 3.22 采用关联器的最优化解调器

函数 $p(t)$ 可以取任何波形，包括图 3.12 中所示调制后的波形。但重要的是要使 $p(t)$ 和 $x(t)$ 脉冲的起始点同步，因为时间上不对准会导致相关器的输出产生小的峰值，这个问题可以用“比特(或码元)同步”来解决。

现在来回顾一下，一个经过数字调制的波形可以表示为一组基函数的线性组合。为了检测被传输信号的电平值，我们可以将相关器的概念一般化。图 3.23 所示的是一个二维信号空间的例子，这里我们首先计算出输入信号和基函数之间的相关性，然后把结果正确组合以得到基带波形。如果我们考虑图 3.24 中所示的我们熟悉的二进制频移键控情况，则这个接收器的行为就可以被很直观的理解了。图 3.24 中，每个相关器通过在一个比特周期积分 $x_{FSK}(t)\cos\omega_1t$ 和 $x_{FSK}(t)\cos\omega_2t$ 来发现 $x_{FSK}(t)$ 与 $\cos\omega_1t$ 和 $\cos\omega_2t$ 的相似程度。然后加法器判定哪一个采样值更大，并相应的产生输出二进制值。

图 3.23 二维信号空间的关联接收器

相干与非相干解调 图 3.24 所示的二进制频移键控接收器可以用来区分两种检波方法。首先，我们注意到两个基函数 $\cos\omega_1t$ 和 $\cos\omega_2t$ 由接收器例如通过两个振荡器产生；而信号 $x_{FSK}(t)$ 在发送器中合成。现在我们假设，由于初始相位的不确定性，其中一个振荡器实际产生 $\sin\omega_1t$ 而接收到的波形为 $A_c\cos\omega_1t$ 。如果 T_b 是 $2\pi/\omega_1$ 的整数倍，这两个信号在一个比特周期内的相关性为 0。因此接收器没能产生一个正确的输出。这个例子说明在接收信号的载波和接收器的振荡器输出之间必须 *相位同步*。要求相位同步的检波技术叫做“相干检波”。

图 3.24 相干 FSK 解调器

相反，一些被调制波形可以采用“非相干”检波的方法解调。一个简单的例子就是图 3.8 所示的频率调制检波器。另一个例子是图 3.25 所示的二进制频移键控检波器。这个电路采用两个中心频率分别在 ω_1 和 ω_2 的窄带滤波器以及随后的包络检波器来确定在一个比特周期内接收到的是哪一个频率。

图 3.25 非相干 FSK 解调器

相干检波器通常基于匹配滤波器的原理，它能够提供具有比非相干检波器更低的比特误差率。尽管如此，非相干检波在射频电路中却有更广泛的应用，因为它们较不复杂。(3.5节)

带宽定义 在我们研究模拟频率调制时，我们注意到 98% 的信号功率处在 $2(\beta+1)B$ 的带宽上。在数字调制中，我们通常定义“99%带宽”，也即该带宽上包含了 99% 的信号功率。尽管这样的百分选择多少是随意的，但是对于所有的调制技术，这个百分比的选择必须是一致的。作为另一个选择，信号带宽可以用临近信道中出现的相对功率数量来定义。如图 3.26 所示，这一想法是用带宽的定义来表示相邻信道能够相隔有多近，也就是每个信道受邻近信道影响的程度。例如，我们说对于带宽为 30kHz，信号显示出的“相邻信道功率”为 -50dBc，这意味着在邻近的 30kHz 信道中的功率除以在带有被调制信号信道中的功率等于 -50dB。

图 3.26 相邻信道功率

3.3.2 二进制调制

二进制基带波形的数字调制方法有幅移键控，相移键控和频移键控三种。因为幅移键控很少在射频中应用，所以在这里我们只讨论二进制相移键控(BPSK)和二进制频移键控(BFSK)。其目的是研究每种调制技术的比特误差率、谱效率和功率效率。我们假设接收信号被附加的高斯白噪声干扰，并在每种情况下求出相干检波器的比特误差率，如果可能的话，也求出非相干检波器的比特误差率。

在讨论任何具体的调制技术之前，我们先来分析一个一般的相关器，其输入波形被等几率的二进制电平信号所调制。与图 3.22 所示的简单相关器不同，这里我们用两个(实)脉冲波形 $p_1(t)$ 和 $p_2(t)$ 来代表二进制数据。这是因为在某些情况下，逻辑 0 并不作为一个简单的“低”电平来传输。例如，在二进制频移键控中， $p_1(t) = A_c \cos \omega_1 t$ ， $p_2(t) = A_c \cos \omega_2 t$ ，其中 $0 < t < T_b$ 。于是如图 3.27(a)所示，最优接收器使输入信号分别与这两个脉冲信号相关，并比较结果来决定收到的比特值。⁴这个结构可以用图 3.27(b)来简化，它说明等效滤波器是和 $p_1(t) - p_2(t)$ 相匹配的。因此，由式(3.25)：

$$SNR_{\max} = \frac{2E_d}{N_0} \quad (3.30)$$

图 3.27 (a) 相干二进制接收器，(b) (a) 的简化版本

其中

$$E_d = \int_{-\infty}^{+\infty} [p_1(t) - p_2(t)]^2 dt \quad (3.31)$$

我们还可以得到，当 $p_1(t) = -p_2(t)$ 时， E_d 的值最大。

⁴ 注意在这一情形中 $p_1(t)$ 和 $p_2(t)$ 不一定正交。

当没有噪声时，图 3.27(b)中积分器在 $t = T_b$ 时的输出信号值为两个不同的值 A_1 和 A_2 (例如 $A_1 = +1$, $A_2 = -1$) 中的一个，分别与 $x(t) = p_1(t)$ 和 $x(t) = p_2(t)$ 相对应。作决定的阈值电压设置在 $(A_1 + A_2)/2$ 。反之当存在噪声时，采样值等于 $A_1 + n(T_b)$ 或 $A_2 + n(T_b)$ ，这里 $n(T_b)$ 是积分器在 $t = T_b$ 时输出端的噪声幅度。注意， $n(T_b)$ 也许不是白噪声，但由于相关器是线性的，所以它仍然是高斯噪声。

现在回想一下概率理论，当两个不相关的随机变量相加时，它们的几率密度函数按卷积相乘。因此，如图 3.28 所示，量 $A_1 + n(T_b)$ 和 $A_2 + n(T_b)$ 表现为两个高斯分布相互重叠，意味着如果 $A_1 + n(T_b)$ 大于阈值电压并且传输信号为 $p_1(t)$ 时，或者 $A_2 + n(T_b)$ 小于阈值电压并且传输信号为 $p_2(t)$ 时就会出错。因为传输 $p_1(t)$ 的几率为 1/2，因此发生第一种出错情况的几率为：

$$P_{e1} = \frac{1}{2} \int_{(A_1+A_2)/2}^{\infty} \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp \frac{-(u - A_1)^2}{2\sigma_n^2} du, \quad (3.32)$$

其中 σ_n 是 $n(t)$ 的标准差。发生第二种出错情况的几率相同。因此总的出错几率为 $P_e = 2P_{e1}$ 。

图 3.28 带有附加噪声的二进制数据 PDF (几率密度函数) 计算

改变变量值，即令 $v = (u - A_1)/\sigma_n$ ，我们有

$$P_e = \int_{(A_1+A_2)/2\sigma_n}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \frac{-v^2}{2} dv \quad (3.33)$$

$$= Q\left(\frac{A_2 - A_1}{2\sigma_n}\right), \quad (3.34)$$

其中 $Q(\cdot)$ 为

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp \frac{-u^2}{2} du. \quad (3.35)$$

虽然上式没有一个显式表示，但当 $x > 3$ 时，式(3.35)可以近似表示为

$$Q(x) \approx \frac{1}{x\sqrt{2\pi}} \exp \frac{-x^2}{2}, \quad (3.36)$$

注意，(3.34)中的结果适用于任何滤波器，尽管 $(A_2 - A_1)/(2\sigma_n)$ 与滤波器的种类有关。

(3.34)中对于 Q 函数的自变量是在积分器输出端两个量(信号和噪声)的比值。更方便的做法是把这个比值用在检波器输入端处的信号和噪声来表示。如果我们从以下两个方面来观测，可以看出对于相关器(匹配滤波器)实现这样做是可行的。第一，线性叠加意味着 $A_2 - A_1$ 是接收器对输入信号 $x(t) = p_1(t) - p_2(t)$ (尽管在实际的系统中并不存在这样的输入信号)的响应。第二，因为滤波器(或者是相关器)和 $p_1(t) - p_2(t)$ 匹配，所以当输入 $x(t) = p_1(t) - p_2(t)$ 时，在采样瞬间的信噪比就等于

$(A_2 - A_1)^2 / \sigma_n^2$ ，它达到式(3.30)所表示的最大值。于是，
 $(A_2 - A_1)^2 / \sigma_n^2 = 2E_d / N_0$ ，因此有：

$$P_e = Q\left(\sqrt{\frac{E_d}{2N_0}}\right) \quad (3.37)$$

式(3.37)是一个重要的结果，它说明误差率仅与差分信号能量和噪声谱密度有关。我们强调一下式(3.37)成立的前提条件是附加高斯白噪声的信道和采用匹配滤波器(相干)检波。

至此得出的结果适合于任何二进制调制技术。我们现在来研究一下上面的分析在具体调制解调器中的含义。

BPSK(二进制相移键控) 在二进制相移键控中，二进制基带数据从载波的两个相反相位中选出一个相位。如图 3.29(a)所示，被调制信号可以表示为 $x_{BPSK}(t) = A_c \cos(\omega_c t + \phi)$ ，其中 $\phi = 0$ 或者 180° 。为了得到星座图，我们重写上式： $x_{BPSK}(t) = \alpha \cos \omega_c t$ ，其中 $\alpha = +A_c$ 或者 $-A_c$ [图 3.29(b)]。因为 $p_1(t) = -p_2(t)$ ，检波器中的相关信号为 $p_1(t) - p_2(t) = 2p_1(t) = 2A_c \cos \omega_c t$ [图 3.29(c)]。这样，当没有噪声时，积分器在 $t = T_b$ 时的输出等于：

$$V_{\text{int}} = \int_0^{T_b} \pm 2A_c^2 \cos^2 \omega_c t dt \quad (3.38)$$

当 ω_c 很大时， $V_{\text{int}} \approx \pm A_c^2 T_b$ ，由此可以得到基带二进制数据。

图 3.29 BPSK (a) 调制器，(b) 信号星座，(c) 解调器

为了利用式(3.31)和(3.37)计算误差几率，我们认识到：

$$E_d = \int_{-\infty}^{+\infty} [p_1(t) - p_2(t)]^2 dt \quad (3.39)$$

$$= \int_0^{T_b} (2A_c \cos \omega_c t)^2 dt \quad (3.40)$$

$$= A_c^2 T_b, \quad (3.41)$$

于是有

$$P_e = Q\left(\sqrt{\frac{A_c^2 T_b}{N_0}}\right) \quad (3.42)$$

为了和其它调制技术进行合理的比较，我们定义每个比特的(平均)能量为 $E_b = A_c^2 T_b / 2$ ，这样得到：

$$P_{e,BPSK} = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \quad (3.43)$$

注意， E_b 包括信号的周期和幅度，意味着增大信号功率或者降低数据传输速率可以降低误差率。 E_b / N_0 的值标志着接收器信噪比的大小。

通过两方面的观察我们可以得到一个二进制相移键控波形的频谱。首先，可以把图 3.29(a)中的调制器改成如图 3.30 所示，此时基带数据在+1 和-1 间随机变化，然后它与 $A_c \cos \omega_c t$ 相乘：

$$x_{BPSK} = x_{BB}(t) \cdot A_c \cos \omega_c t \quad (3.44)$$

图 3.30 简化了的 BPSK 调制器

这样，只要把信号 $x_{BB}(t)$ 的频谱移到 $\pm \omega_c$ 附近就得到信号 $x_{BPSK}(t)$ 的频谱。其次，采用 $+p(t)$ 和 $-p(t)$ 来分别表示 1 和 0 时的一个随机二进制信号 $x(t)$ ，具有如下功率谱密度，

$$S_x(\omega) = \frac{1}{T_b} |P(\omega)|^2 \quad (3.45)$$

其中 $P(\omega)$ 是 $p(t)$ 的傅立叶变换[1]。如果 $p(t)$ 是时间间隔为 T_b 单位幅度的矩形脉冲，则有 $P(\omega) = [2 \sin(\omega T_b / 2)] / \omega$ ，且

$$S_{BPSK}(\omega) = \frac{A_c^2 \sin^2[(\omega + \omega_c) T_b / 2]}{T_b (\omega + \omega_c)^2} + \frac{A_c^2 \sin^2[(\omega - \omega_c) T_b / 2]}{T_b (\omega - \omega_c)^2} \quad (3.46)$$

二进制相移键控的功率效率将在 3.4 节中讨论。

BFSK(二进制频移键控) 在二进制频移键控中，二进制基带数据从两个具有相同幅度的载波频率中选出一个频率[图 3.31(a)]。⁵被调制信号可以表示为 $x_{BFSK}(t) = \alpha_1 \cos \omega_1 t + \alpha_2 \cos \omega_2 t$ ，其中 $[\alpha_1 \ \alpha_2] = [0 \ A_c]$ 或者 $[0 \ -A_c]$ 。当两个基函数在一个比特周期内正交时，必有：

$$\int_0^{T_b} \cos \omega_1 t \cos \omega_2 t dt = 0 \quad (3.47)$$

图 3.31 BFSK (a) 调制器，(b) 信号星座，(c) 解调器

对于 $\omega_1 + \omega_2 \gg \omega_1 - \omega_2$ ，上式简化为 $[\sin(\omega_1 - \omega_2) T_b] / (\omega_1 - \omega_2) = 0$ ，于是 $(\omega_1 - \omega_2) T_b = n\pi$ ，当 ω_1 和 ω_2 之间取最小距离时， $(\omega_1 - \omega_2) T_b = \pi$ ，或者 $f_1 - f_2 = 1/(2T_b)$ 。

与相移键控不同，正交的频移键控是二维的调制技术，其信号空间如图 3.31(b) 所示。这里 $p_1(t) = A_c \cos \omega_1 t$ ， $p_2(t) = A_c \cos \omega_2 t$ ；也就是说 $p_1(t) \neq p_2(t)$ 。正如我们在下面将要看到的，这导致了二进制相移键控和二进制频移键控在示错特性之间出现一些有趣的差别。

如图 3.31(c) 所示，一个二进制频移键控检波器将输入与 $p_1(t) - p_2(t) = A_c \cos \omega_1 t - A_c \cos \omega_2 t$ 相关。结果在没有噪声情况下，如果 $x_{BFSK} = A_c \cos \omega_1 t$ ，积分器在 $t = T_b$ 时的输出等于 $A_c^2 T_b / 2$ ，而 $x_{BFSK} = A_c \cos \omega_2 t$ 时，输出为 $-A_c^2 T_b / 2$ 。

⁵ 这一方法称为“不连续”FSK，因为其相位突变。

由式(3.31)和(3.37)，我们可以计算出相干二进制频移键控检波器的比特误差率，因为 $p_1(t)$ 和 $p_2(t)$ 是正交的，

$$E_d = \int_0^{T_b} [p_1^2(t) + p_2^2(t)] dt \quad (3.48)$$

$$= A_c^2 T_b \quad (3.49)$$

于是有，

$$P_e = Q\left(\sqrt{\frac{A_c^2 T_b}{2N_0}}\right) \quad (3.50)$$

与二进制相移键控一样，我们定义二进制频移键控每比特的平均能量为 $E_b = A_c^2 T_b / 2$ ，这样得到

$$P_{e,BFSK} = Q\left(\sqrt{\frac{E_b}{N_0}}\right) \quad (3.51)$$

比较式(3.43)和 (3.51)表明，当给定误差率和噪声密度时，二进制频移键控的比特能量一定是二进制相移键控的两倍。什么是导致差别的原因呢？从图 3.29(b) 和 3.31(b)所示的信号空间可以看出，对于给定的载波幅度 A_c ，二进制相移键控中干扰一个比特使相关器输出错误的超过阈值所需要的噪声数量比在二进制频移键控中要高。换句话说，二进制相移键控信号空间中两点间的最小距离比二进制频移键控的大。从另一个角度来看，回想一下在 $SNR_{\max} = 2E_d / N_0$ 时，若 $p_1(t) = -p_2(t)$ ，则 E_d 的值达到最大，这在二进制相移键控时成立，而在二进制频移键控时不成立。

以上两者之间的差别经常概括为“二进制相移键控相对于二进制频移键控有 3-dB 的优势”。尽管如此，二进制频移键控仍广泛应用于低数据率的应用中(例如寻呼机)，这时采用较长的 T_b 使 E_b 达到最大。二进制频移键控之所以如此流行，是因为它易于检波，特别是在非相干同步系统中(3.5 节)，以及它具有较高的功率效率(3.4 节)。

频移键控信号的频谱通常很难计算[1]。由 Carson 定律，98% 带宽近似等于 $B_T = 2(\Delta f + 1/T_b)$ ，其中 $\Delta f = f_1 - f_2$ [1]。对于正交的频移键控， $\Delta f = 1/(2T_b)$ ， $B_T = 3/T_b$ 。

二进制频移键控的功率效率将在 3.4 节中讨论。

3.3.3 正交调制

在很多应用中，在进行调制之前，把一个二进制数据流分成两比特一组，并用四个电平中的一个来表示，对比特是一个很有利的办法。例如，一对比特 b_m ， b_{m+1} 可以按如下加在同一个载波上：

$$x(t) = b_m A_c \cos \omega_c t - b_{m+1} A_c \sin \omega_c t \quad (3.52)$$

这之所以可行是因为 $\cos \omega_c t$ 和 $\sin \omega_c t$ 是正交的。

我们称这种处理方法为“正交调制”或“正交复路传输”，如图 3.32 所示，这里一个串联至并联转换器把连续的比特流分成两路，一路通过上臂，另一路通过下臂。每组两个比特(每臂中各一个)构成一个码元，因此码元率是比特率的一半(3.31 节)，这表明所需要的带宽是二进制相移键控的一半。这是正交调制得到广泛应用的主要原因。

图 3.32 正交调制

为了得到信号空间，我们假设 b_m 和 b_{m+1} 位是高度为 ± 1 的矩形脉冲，并把调制后的信号写为： $x(t) = \alpha_1 \cos \omega_c t + \alpha_2 \sin \omega_c t$ ，其中 α_1 和 α_2 每个均可以取 $+A_c$ 或 $-A_c$ 值。信号空间如图 3.33 所示。

图 3.33 正交调制的信号星座

正交调制包括两大类：正交相移键控($\pi/4$ -QPSK)和最小位移键控(MSK)。QPSK 包括更多具体的亚种类，如位移正交相移键控(OQPSK)和 $\pi/4$ -QPSK；MSK 则包含一个有着广泛应用的子类：高斯最小位移键控(GMSK)。

为什么正交调制包含这么多的种类？回顾一下对于 RF 应用有三个参数决定了一种调制技术的性能：BER，谱效率和功率效率。以上各种调制技术在这三个参数间体现了不同的利弊。

QPSK 在式(3.52)中，如果比特波形是一个矩形脉冲，我们就得到了一个 QPSK 信号。从另一个角度来看，和 BPSK 类似，根据码元可以从一个正弦信号的四个相位中选择出一个相位：即 $x_{QPSK}(t) = \sqrt{2}A_c \cos(\omega_c t + k\pi/4)$ ， $k=1, 3, 5, 7$ 。

QPSK 信号的相干检波可以按图 3.34 所示进行。这里，该电路将输入信号与 $\cos \omega_c t$ 和 $\sin \omega_c t$ 相关以从集合 $[\pm 1 \pm 1]$ 中选出最有可能的码元值。一个并联至串联(P/S)的转换器接着把一个码元中的每个比特复合，从而产生二进制基带数据。

图 3.34 相干 QPSK 解调

通过比较图 3.29(b)和图 3.33，我们也许会得出 QPSK 比 BPSK 有更高的误差概率的结论，因为图 3.33 所示信号空间中的每个点有可能被错误的取作为另外 3 个点，而这两种技术中各点间的最小距离是相等的。然而更仔细的分析表明，情况并不如此。第一，因为一个公平的比较要求发送器的输出功率相同，所以我们将两个调制后的波形分别表示为： $x_{BPSK}(t) = \pm A_c \cos \omega_c t$ 和 $x_{QPSK}(t) = A_c \cos(\omega_c t + k\pi/4)$ ，这样它们的平均功率相同均为 $A_c^2/2$ 。由此， $x_{QPSK}(t) = \pm(A_c/\sqrt{2})\cos \omega_c t \pm (A_c/\sqrt{2})\sin \omega_c t$ ，于是得到图 3.35(b)所示的信号空间图。图中各点间更加靠近，似乎更加肯定了我们对于 QPSK 具有较高 BER 的猜想。但是，一个公平的比较还要求两种技术总的比特传输率是相等的。这就意味着，对于一个周期为 T_b 的二进制基带数据，BPSK 接收器中的相关器[图 3.29(c)]将

输入信号积分了 T_b 秒，而 QPSK 检波器(图 3.34)中的相关器将输入信号积分了 $T_s = 2T_b$ 秒。因此第一种情况的采样结果(也就是信号能量)等于 $A_c^2 T_b / 2$ ，而第二种情况的结果为 $(\sqrt{2}A_c / 2)^2 (2T_b) = A_c^2 T_b$ 。因此，BPSK 的比特能量低于 QPSK 的码元能量。

图 3.35 相同输出功率的 BPSK 和 QPSK 信号星座

我们把上面的观察总结如下：与 BPSK 相比，QPSK 信号空间中的点更加邻近，具有更小的最小距离，但是有更高的码元能量。如果把这些因素全部考虑进来，可以证明 BPSK 和 QPSK 具有几乎相等的误差几率；也即 $P_e = Q(\sqrt{2E_b / N_0})$ [1]。

因为一个 QPSK 波形可以看成是两个 BPSK 信号的和，所以 QPSK 波形的频谱和它相应的 BPSK 波形的频谱类似，但是比特流的速率在正交复合前被减少了一半(串并转换)。因此，QPSK 的频谱密度可以将式(3.46)中的 T_b 换为 $2T_b$ 即可求出。

QPSK 的一个很大的缺点就是每个码元结束时都会有很大的相位变化。如图 3.36 所示，如果串并行转换器的输出波形同时发生变化，例如，从 $[-1 \ -1]$ 到 $[+1 \ +1]$ ，那么载波就会产生 180° 的相位跃变，这相当于在信号空间中处在对角线相反位置上的两点之间发生转换。正如我们在 3.4 节中将看到的，如果波形要被过滤且随后被一个非线性功率放大器处理的话，这个转变是不希望有的。

图 3.36 QPSK 波形中的相变

QPSK 的一个变型 OQPSK 能够克服上述缺点。如图 3.37 所示，数据流在串并行转换之后，在时间上被平移了半个码元周期，从而避免了在节点 A 和 B 处的波形同时发生变化。因此相位跃变仅仅为 $\pm 90^\circ$ 。图 3.38 画出了相位在时域和信号空间中的变化。OQPSK 的比特误差率和频谱与 QPSK 完全相同。

图 3.37 偏移 QPSK 调制器

图 3.38 OQPSK (偏移 QPSK) 中的相变

尽管 OQPSK 具有相位变化较小的优势，但它却存在一个很重要的缺点：它不适合于“差分编码”[4]。我们将推迟到 3.5 节讨论这个问题，但在这里指出差分编码在如今 RF 应用最流行的非相干接收器中起着重要的作用。

QPSK 的另一个变型是“ $\pi/4$ -QPSK”[3, 4]。与 OQPSK 不同，这一技术可以进行差分编码并应用在美国和日本的电话系统中(第 4 章)。在 $\pi/4$ -QPSK 中，信号集合包含两种 QPSK 信号，一种的相位相对于另一被旋转 45° ：

$$x(t) = A_c \cos\left(\omega_c t + k \frac{\pi}{4}\right) \quad k \text{ 为奇数}, \quad (3.53)$$

$$x(t) = A_c \cos\left(\omega_c t + k \frac{\pi}{4}\right) \quad k \text{ 为偶数}, \quad (3.54)$$

如图 3.39 所示, 交替的从两个 QPSK 发生器中取出输出信号来实现对信号的调制。

图 3.39 $\pi/4$ -QPSK 信号的示意性产生

为了更好的理解这一操作, 让我们来研究一下 $\pi/4$ -QPSK 发生器的简单实现。经过串并行转化之后, 数字信号电平被放大和移项表现为在上臂 QPSK 调制器中的 ± 1 , 而在下臂中的值为 0 和 $\pm\sqrt{2}$ 。因此输出信号等于 $x_1(t) = \alpha_1 \cos \omega_c t + \alpha_2 \sin \omega_c t$, 其中 $[\alpha_1 \ \alpha_2] = [\pm A_c \ \pm A_c]$ 及 $x_2(t) = \beta_1 \cos \omega_c t + \beta_2 \sin \omega_c t$, 其中 $[\beta_1 \ \beta_2] = [0 \ \pm\sqrt{2}A_c]$ 和 $[\pm\sqrt{2}A_c \ 0]$ 。于是信号空间在如图 3.40 所示的两个信号空间之间交替变化。现在来考虑一个基带信号序列 $[1,1,0,1,1,0,1,0]$ 。如图 3.41 所示, 第一对 $[1,1]$ 在上臂中被转化为 $[+A_c \ +A_c]$, 产生 $y(t) = A_c \cos(\omega_c t + \pi/4)$ 。第二对 $[0,1]$ 在下臂中被转化为 $[0 \ -\sqrt{2}A_c]$, 产生 $y(t) = -\sqrt{2}A_c \cos \omega_c t$ 。考虑整个序列的 $y(t)$ 值之后, 我们注意到从两个信号空间选出来的点如图 3.4(a) 那样表现为时间 t 的函数。这里的关键之处在于, 由于没有任何两个连续的点来自同一个信号空间, 因此最大相位跃变为 135° , 这比 QPSK 的相位跃变少了 45° 。图 3.41(b) 说明了这点。因此就最大相位变化而言, $\pi/4$ -QPSK 是 QPSK 和 OQPSK 的折中情况。

图 3.40 $\pi/4$ -QPSK 信号的产生

图 3.41 (a) $\pi/4$ -QPSK 在时域中的演变, (b) 在信号星座中的可能相变

$\pi/4$ -QPSK 与 QPSK 的频谱和 BER 完全相同。我们将在 3.4 节中讨论 QPSK 及其变型的功率效率。

以上我们在处理 QPSK 系列时, 假设了基带比特用矩形脉冲来表示。然而在限制带宽的应用中, 我们希望采用频谱在相邻信道下降很快的脉冲波形。例如, 2.2 节中介绍的升余弦信号, 不仅提供了密集的频谱, 还提供了无码间干扰(ISI)的操作。

在实际应用中, 升余弦滤波器被分为两部分, 一部分放在发送器中, 而另一部分放在接收器中(图 3.42)。这两部分以式(2.54)的平方根为传输函数, 两部分在一起可以完成 Nyquist 信号发送, 同时在接收器中的那一部分还可以起到匹配滤波器的作用。这一技术在[5]中有详细介绍。

图 3.42 提升余弦滤波器分解成两个部分

MSK 最小位移键控的概念可以从几个不同的角度来描述, 每个角度都有助于我们对基本概念有更深入的理解。首先, 我们通过 OQPSK 来认识 MSK, 但是我们最终要把 MSK 看成一种频移键控方法, 并由此来理解 GMSK。

回顾一下式(3.52), 在正交调制中, 每两个连续的比特 b_m 和 b_{m+1} 被加在一个载波的正交相位上。如果 b_m 和 b_{m+1} 都是矩形脉冲, 那么在调制后的信号上就会产生

突然的相位变化，这会导致频谱变宽，因而使功率放大器的设计发生困难(3.4节)。这一考虑使有理由采用“连续相位调制”技术[6, 7]。MSK 就是这样的技术之一。

假设在 OQPSK 中，我们用半正弦而不是矩形脉冲信号来表示和载波信号相乘的信号电平。更明确的，如图 3.43 所示，我们把上臂中的信号电平乘以 $\cos \omega_1 t$ ，下臂中的信号电平乘以 $\sin \omega_1 t$ ，其中 $\omega_1 = \pi/(2T_b)$ [9]。这样，

$x(t) = a_m \cos \omega_1 t \cos \omega_c t - a_{m+1} \sin \omega_1 t \sin \omega_c t$ ，其中 a_m 和 a_{m+1} 是在 +1 和 -1 之间翻转的矩形脉冲，其宽度为 $2T_b$ ，并且它们之间相差一个 T_b 。如图 3.43 所示， $x(t)$ 在相位上或等效的说在幅度的斜率上没有任何突变。为了理解为什么这样，我们首先注意到 a_m 和 a_{m+1} 分别只在 $t = (2k+1)T_b$ 和 $t = 2kT_b$ 时变化。现在我们假设 $a_{m+1} = 1$ 时， a_m 从 +1 变为 -1。于是， $x(t)$ 从 $2 \cos(\omega_1 + \omega_c)t$ 变化为 $2 \cos[(\omega_1 + \omega_c)t + \pi]$ ，其中 $t = (2k+1)T_b$ 。结果，相位变化为

$\Delta\phi = (\omega_c - \omega_1)t + \pi - (\omega_c + \omega_1)t = -2\omega_1 + \pi = -2k\pi$ ，因为 $\omega_1 = \pi/(2T_b)$ 。 a_m 和 a_{m+1} 的其它三种可能的变化情况也可以证明有相同的结果，这表明相位总是连续的。

图 3.43 MSK (最小位移键控) 调制和信号星座

虽然 MSK 与 QPSK 具有相同的误差率，但它的频谱比矩形脉冲的 QPSK 系列下降得更快。MSK 中平滑的相位变化虽然可以降低频谱边带中的信号功率，但这是以加大主带带宽为代价的。作为一个经验公式，如果带通信号的相位具有有界的直至 n 阶且包括 n 阶的导数，那么它的信号频谱就以正比于 $f^{2(n+1)}$ 的速率衰减[8]。事实上，对于 MSK，我们有

$$S_{MSK}(f) = \frac{16A_c^2 T_b}{\pi^2} \left\{ \frac{\cos^2[2\pi T_b(f - f_c)]}{[1 - 16T_b^2(f - f_c)^2]^2} + \frac{\cos^2[2\pi T_b(f + f_c)]}{[1 - 16T_b^2(f + f_c)^2]^2} \right\} \quad (3.55)$$

这就是说衰减速率正比于 f^4 。我们应当把这个结果和式(3.46)进行比较。

上面的经验公式说明，如果可以使相位变化比 MSK 更加平缓，我们就可以实现更高谱效率的调制技术。这个结论已经产生了更一般的 MSK 信号[8]，高斯最小位移调制(GMSK)就是其中的一个例子。为了理解 GMSK，我们首先不加证明的指出图 3.43 的 MSK 信号还可以表示为

$$x_{MSK}(t) = \sqrt{2}A_c \cos \left[\omega_c t + \int_{-\infty}^t \sum_m b_m p(t - mT_b) dt \right] \quad (3.56)$$

其中累加项代表二进制基带信号；即 $b_m = \pm 1$ ， $p(t)$ 是宽度为 T_b 的矩形脉冲。反之在 GMSK 中， $p(t)$ 是一个较光滑的函数。注意，式(3.56)可以看做是频率调制，当 $p(t)$ 为矩形脉冲时，就是 FSK。

在 GMSK 中，我们使基带矩形脉冲信号通过具有高斯冲击响应 $h(t) = \exp(-\alpha t^2)$ 的滤波器得到脉冲波形 $p(t)$ 。如图 3.44 所示，为了得到滤波器的输出 $x_{BB}(t) * h(t)$ ，我们可以先计算 $(dx_{BB}/dt) * h(t)$ 然后对它进行积分。这个方法说

明对于窄带高斯响应(α 很大), 滤波器的输出近似宽度为 T_b 的矩形脉冲信号, 并且随着 $h(t)$ 变宽, 输出也变宽。

图 3.44 GMSK 调制

推导出 GMSK 功率谱密度的表达式很困难, 但是我们可以凭直觉预见到, 如果时域调制脉冲 $p(t)$ 变得更加光滑, 功率谱的衰减越快。例如, 如果 $\alpha=1$, 则 99% 带宽小于 $1.2/T_b$ [10]。

虽则我们的分析表明, α 的值越小, 频谱就越窄, 但 α 的选择还取决于另一个因素: 码间干扰。与以往的调制技术不同, GMSK 允许调制脉冲的宽度大于 T_b , 因此会受到 ISI 的影响。随着 α 值的减小, ISI 的影响变得更明显。在一些典型的系统中, α 的值在 0.3 左右, 因此要求减轻接收器中的 ISI。MSK 和 GMSK 的功率效率将在 3.4 节中讨论。

对各种调制技术的频谱和误差几率的讨论见[11, 12, 13]。

3.4 调制技术的功率效率

至今为止, 我们在研究不同的调制技术时, 曾多次提到功率效率是一个重要的特点。我们将在本节对它进行讨论。

3.4.1 等幅包络和变幅包络信号

对于一个调制后的波形 $x(t) = A(t)\cos[\omega_c t + \phi(t)]$, 当 $A(t)$ 不随时间变化时, 我们称 $x(t)$ 为等幅包络信号, 否则称之为变幅包络信号。

等幅包络和变幅包络信号在非线性系统中的表现截然不同。假设 $A(t) = A_c$ 且系统表现为 3 阶无记忆非线性, 即:

$$y(t) = \alpha_3 x^3(t) + \dots \quad (3.57)$$

$$= \alpha_3 A_c^3 \cos^3[\omega_c t + \phi(t)] + \dots \quad (3.58)$$

$$= \frac{\alpha_3 A_c^3}{4} \cos[3\omega_c t + 3\phi(t)] + \frac{3\alpha_3 A_c^3}{4} \cos[\omega_c t + \phi(t)]. \quad (3.59)$$

式(3.59)中的第一项代表在 $\omega = 3\omega_c$ 附近的调制后的信号。因为原有信号 $A(t)\cos[\omega_c t + \phi(t)]$ 的带宽一般远小于 ω_c , 由 Carson 定律可知, 由 $\cos[3\omega_c t + 3\phi(t)]$ 所占的也是很小的。这样, ω_c 附近的频谱形状保持不变。

现在考虑在上述的非线性系统中加入一个变幅包络信号。此时 $x(t)$ 可以写为:

$$x(t) = x_I(t)\cos\omega_c t - x_Q(t)\sin\omega_c t, \quad (3.60)$$

其中, $x_I(t)$ 和 $x_Q(t)$ 分别是基带信号的同相幅度分量和正交幅度分量, 于是我们有:

$$y(t) = \alpha_3 [x_I(t) \cos \omega_c t - x_Q(t) \sin \omega_c t]^3 + \dots \quad (3.61)$$

$$= \alpha_3 x_I^3(t) \frac{\cos 3\omega_c t + 3 \cos \omega_c t}{4} - \alpha_3 x_Q^3(t) \frac{-\sin 3\omega_c t + 3 \sin \omega_c t}{4} + \dots \quad (3.62)$$

这样，输出包含了中心频率为 ω_c 的 $x_I^3(t)$ 和 $x_Q^3(t)$ 的频谱。由于这些分量的频谱一般比信号 $x_I(t)$ 和 $x_Q(t)$ 的频谱宽，因此我们说当一个变幅包络信号通过一个非线性系统时，它的频谱会“增长”。

但是当基带脉冲波形为矩形时，上面的情况则会不同。由于一个矩形脉冲波形在较高功率时仍然是矩形，所以对于任何非线性调制后信号的频谱形状都是 $\sin^2 \omega_c T_b$ 。

3.4.2 频谱再增长

考虑图 3.45 所示的简单发送器。这里调制器前面的低通滤波器限制了信号的带宽，因此抑制了频谱漏入相邻的信道。功率放大器不仅把信号放大还起到了缓冲器的作用，它把所要求的信号发送给天线。

对数字调制后的信号进行滤波会产生什么样的影响？直觉上，我们希望通过限制带宽可以消除信号在时域上的突变。例如，一个 QPSK 波形在经过滤波器后如图 3.45 所示，当滤波器的带宽变窄时，它的包络变化表现的更加明显。我们还注意到，当相位跃变较大时，由滤波引起的包络变化也变得更加显著。

图 3.45 带宽限制对 QPSK 波形包络的影响

信号包络的变化意味着什么？带宽限制的频谱和时变幅包络间的一一对应关系表明如果功率放大器要使频谱保持在有限带宽上，它也必须使信号包络的变化线性放大。如果我们把调制后的波形表示为 $x(t) = x_I(t) \cos \omega_c t - x_Q(t) \sin \omega_c t$ 就能清楚的看出这点。如果功率放大器呈现出明显的非线性，那么 $x_I(t)$ 和 $x_Q(t)$ 的波形就会改变，频谱也不会限制到所希望的带宽。我们称这种效应为“频谱再增长” [14, 15]，它可以用相对的邻近信道功率来量化。

上面的讨论可以总结如下。数字调制后的信号例如 QPSK 中如果有相位突变，那么当有滤波器限制带宽时，就会导致信号包络的变化。这一变化又要求采用线性功率放大器以避免频率再增长。第 9 章中将会阐明，线性功率放大器一般比非线性功率放大器的效率低。对于 40% 的效率(这在线性放大器已很难达到)和 1W 的输出功率，线性功率放大器消耗的总功率为 2.5W，这相对于一个移动电话其它部分消耗的功率是一个相当大的数量。非线性放大器则不同，它的效率可以高达 60%。因此，我们希望采用的调制技术在用非线性放大器处理时不会产生频谱再增长。调制形式的这一性能叫做功率效率。

功率效率高的调制技术的一个简单例子就是 FM。从式

$x_{FM}(t) = A_c \cos[\omega_c t + m \int x_{BB}(t) dt]$ 可以看出，频率调制波形没有任何突然的相位变化

同时又表现为等幅包络信号。因此 FM 和 FSK 信号可以用非线性功率放大器来放大且不产生频谱再增长。

从图 3.45 所示中，我们可以得出结论即经过滤波的 QPSK 信号很容易产生频谱再增长。例如，如果在基带中采用开方升余弦滤波来得到窄带信号，那么所引起的包络变化要求采用较为线性的功率放大器[17, 16]。这个缺点在 OQPSK 和 $\pi/4$ -QPSK 中较不严重，因为它们允许较小的相位跃变。

比 QPSK 系列具有更高功率效率的是 MSK 信号，因为它的等幅包络允许采用非线性功率放大器处理。然而，当采用升余弦滤波时，MSK 信号一般比 QPSK 波形要求更大的带宽。

总而言之，数字调制技术在谱效率和功率效率之间有一个综合考虑的问题。如果信号带宽受到滤波或者脉冲整形较多的限制时，功率放大器就必须有更高的线性度以避免频谱再增长。不同的无线系统总是选择两个极端之一：采用窄信道和线性功率放大器的脉冲整形 QPSK 调制或者是采用宽信道和非线性功率放大器的 MSK 调制(第 4 章)。

3.5 非相干解调

在 3.31 节中，我们看到匹配滤波器—它的一个例子就是相干检波器(相关器)—提供了最高的信噪比以及最低的误差概率。然而，相干检波器要求接收器中本地振荡器的相位与接收信号相位之间有严格的关系要求，所谓“载波恢复”，即建立这样的相位对准要求电路具有相当的复杂度，而且当存在干扰和信号衰减时，就更难达到要求。因此，很多射频系统采用非相干检波，尽管这种方法的性能多少差些。

非相干 FSK 解调 在 3.31 节中，我们简要的研究了一个非相干 FSK 检波器。如图 3.46 所示，这一结构采用两个带通滤波器(用于二进制调制)，每个滤波器后面连有一个包络检波器以决定接收到的信号频率。因为 $\omega_1 - \omega_2$ 一般远小于 ω_1 和 ω_2 ，输入信号的频谱首先被转换为较低的频率，以降低对品质因数 Q 及带通滤波器中心频率精确度的要求。

图 3.46 非相干 FSK 解调

推导这种检波器的误差概率十分困难。我们只是给出 ω_1 和 ω_2 相差足够大，所以它们的频谱重叠可以忽略的实际情况下的结果[1]。

$$P_e = \frac{1}{2} \exp \frac{-E_b}{2T_b B_p N_0}, \quad (3.63)$$

其中 B_p 是带通滤波器的(有效)带宽， E_b 是比特能量， $E_b = A_c^2 T_b / 2$ 。当 $B_p = 1/T_b$ 时，有

$$P_e = \frac{1}{2} \exp \frac{-E_b}{2N_0}, \quad (3.64)$$

如果误差率为 10^{-3} 量级，那么式(3.64)要求 E_b/N_0 的值比相干 FSK 检波器中的值大 1.5dB[式(3.51)]。在很多应用当中，较低的电路复杂度补偿了性能上的损失。

差分相移键控 虽然在 FSK 中，有可能确定与每一比特相应的频率(用带通滤波器或者是频率解调器)，但是在 PSK 中相位和时间原点有关，因此没有任何“绝对的”意义。例如，一个 QPSK 波形 90° 的相移把原来的信号空间变成类似的信号空间，但是却使所有的码元都被不正确的解调。因此，简单的 PSK 信号不能采用非相干检波方式。然而，如果信息基于从一个比特(或码元)到下一个比特(或码元)时相位的变化，那么就不需要时间原点，于是非相干检波就是可行的。这是通过基带信号分别在调制和解调之前进行“差分”编码和解码来完成的。

现在我们来考虑二进制差分相移键控(DPSK)。差分编码的原则是如果当前输入比特值为 1，则编码器的输出状态不变，反之亦然。这就要求有一个额外的初始比特(任意值)。为了更好的理解这个概念，我们看一下图 3.47(a)实现的例子。一个异或非门(XNOR)通过比较当前的输出比特 $D_{out}(mT_b)$ 和当前输入比特 $D_{in}(mT_b)$ 来决定下一个输出状态，

$$D_{out}[(m+1)T_b] = \overline{D_{in}(mT_b) \oplus D_{out}(mT_b)} \quad (3.65)$$

图 3.47 (a) 差分编码，(b) 差分解码，(c) 已编码与已解码序列实例

这意味着如果 $D_{in}(mT_b)=1$ ，则有 $D_{out}[(m+1)T_b]=D_{out}(mT_b)$ ，而如果 $D_{in}(mT_b)=0$ ，则有 $D_{out}[(m+1)T_b]=\overline{D_{out}(mT_b)}$ 。刚才提到的额外初始比特相应于数据序列开始之前触发器的状态。

一个系统接收到的 DPSK 信号可以通过反操作来解码，如图 3.47(b)所示。显然我们可以看出在被编码信号中的相移并不改变接收器的差分解码输出。

用差分方法传输信号并不具有最低的比特误差率，因为它并不基于匹配滤波。从另一个角度来看，在相干解调中，信号是和一个低噪声参照(振荡器)相比较，而在差分检波中是两个被噪声干扰的信号相互比较[12]。可以证明 DPSK 的误差几率为

$$P_e = \frac{1}{2} \exp \frac{-E_b}{N_0}, \quad (3.66)$$

其中 $E_b = A_c^2 T_b / 2$ 。这个结果说明了当误差率低于 10^{-3} 时，DPSK 要求的 E_b/N_0 值比相干 PSK 的对应值高 3dB[式(3.43)]。

参考文献

- [1] L. W. Couch, *Digital and Analog Communication Systems*, 4th ed., New York: Macmillan, 1993.
- [2] H. E. Rowe, *Signals and Noise in Communication Systems*, Princeton, NJ: Van Nostrand, 1965.
- [3] P. A. Baker, "Phase-Modulated Data Sets for Serial Transmission at 2000 and 2400 Bits per Second," Part I, *AIEE Trans. Communication Electronics*, pp. 166-171, July 1962.
- [4] Y. Akaiwa and Y. Nagata, "Highly Efficient Digital Mobile Communication with a Linear Modulation Method," *IEEE J. of Selected Areas in Communications*, Vlo. 5, pp. 890-895, June 1987.
- [5] R. D. Gitlin, J. F. Hayes, and S. B. Weinstein, *Data Communications Principles*, New York: Plenum Press, 1992.
- [6] C. E. W. Sundberg, "Continuous Phase Modulation," *IEEE Comm. Mag.*, Vol. 24, pp. 25-38, April 1986.
- [7] J. B. Anderson and C. E. W. Sundberg, "Advances in Constant Envelope Coded Modulation," *IEEE Comm. Mag.*, Vol. 29, pp. 36-56, December 1991.
- [8] R. E. Ziemer and R. L. Peterson, *Digital Communication and Spread Spectrum Systems*, New York: Macmillan, 1985.
- [9] S. Pasupathy, "Minimum Shift Keying: A Spectrally Efficient Modulation," *IEEE Comm., Mag.*, Vol. 19, pp. 14-22, July 1979.
- [10] K. Murota and K. Hirade, "GMSK Modulation for Digital Mobile Radio Telephony," *IEEE Trans. Comm.*, Vol. 29, pp. 1044-1051, July 1981.
- [11] T. S. Rappaport, *Wireless Communications, Principle and Practice*, Englewood Cliffs, NJ: Prentice Hall, 1996.
- [12] B. Sklar, *Digital Communications*, Englewood Cliffs, NJ: Prentice Hall, 1988.
- [13] K. Feher, *Wireless Digital Communications*, Englewood Cliffs, NJ: Prentice Hall, 1995.
- [14] D. H. Morais and K. Feher, "The Effects of Filtering and Limiting on the Performance of QPSK, Offset QPSK, and MSK Signals," *IEEE Trans. Communications*, Vol. 28, pp. 1999-2009, December 1980.
- [15] J. F. Sevic and J. Staudinger, "Simulation of Adjacent Channel Power for Digital Wireless Communication Systems," *Microwave J.*, pp. 66-80, October 1996.

- [16] S. Ariyavisitakul and T. P. Liu, "Characterizing the Effects of Nonlinear Amplifiers on Linear Modulation for Digital Portable Radio Communications," *IEEE Trans. Veh. Tech.*, Vol. 39, pp. 383-389, November 1990.
- [17] J. Boccuzzi, "Performance Evaluation of Nonlinear Transmit Power Amplifiers for North American Digital Cellular Portables," *IEEE Trans. Veh. Tech.*, Vol. 44, pp. 220-228, May 1995.

第 4 章 多址访问技术和无线通信标准

第 3 章描述的调制和解调技术提供了单个发送器和接收器之间的通信基础,而在一个拥有极大数量收发器的网络中,就需要额外的方法来确保多个用户之间的正确通信,这种方法叫做“多址访问技术”。

无线系统的复杂性要求它们遵循一定的“标准”。除了调制和多址技术,一个无线标准还包括例如频带、时序和数据编码等技术细节,还要定义一套精确测试收发器性能的方法。

在对移动通信系统特性进行一个简单回顾之后,我们在本章中将叙述 RF 系统里通常采用的三种通用多址技术。它们是:频分多址(FDMA),时分多址(TDMA),和码分多址(CDMA)。我们然后将研究几种流行的无线标准,它们是北美数字通信(NADC)标准,全球移动通信系统(GSMC),Qualcomm CDMA,和欧洲数字无绳电话(DECT)标准。

4.1 移动射频(RF)通信

一个移动通信系统是指用户在这个系统里面可以一边和别人相互通信,一边可以在物理位置上进行移动。例如:传呼机、蜂窝电话和无绳电话。正是 RF 通信的可移动性,才使得它异常的有用和受欢迎。用户所持的收发器叫做“移动单元”(简称“单元”)、“终端”或者“手持单元”。无线系统底层结构的复杂性往往要求移动单元只通过一些固定的、较昂贵的被称作“基站”的设备进行通信。每一个移动单元从基站接收或者发送到基站的信息是通过两个 RF 信道进行的,分别称作“前馈信道”或“下行链路”,和“反馈信道”或“上行链路”。这本书的主要内容是针对移动单元的,因为,和基站相比,手持单元在市场上占有得多的份额,并且它们的设计也很类似于其它类型的 RF 系统。

蜂窝系统 在一个有限可用的频谱上(例如,900MHz 附近的一个 25MHz 的通道),一个拥挤的大都市里的成千上万的人之间是如何互相通信的呢?为了回答这个问题,让我们首先考虑一种较简单的情况:在一个国家里可以有几千个 FM 无线广播电台在同时以 88-108MHz 的频率工作。这是可能,因为在物理位置上相隔足够远的电台可以使用同一载波频率(所谓“频率复用”),而相互间的干扰可以被忽略(除非是处在两个电台的中间位置上,并且它们的发送功率相同)。两个可以采取相同载波频率的电台之间的最小距离是由每一个电台发射的信号功率所决定的。

在移动通信系统中,频率复用是通过一种“蜂窝”状的结构来实现的,其中每一个蜂窝是六边形的,其周围环绕着 6 个其它的蜂窝[图 4.1(a)]。频分复用的概念是:如果中央位置的蜂窝使用频率 f_1 来通信,那么与其相邻的 6 个蜂窝就不可以使用这个频率,但是在与之直接相邻的蜂窝的外面的蜂窝就可以再次使用 f_1 。实际上,更有效的频率分配方式是如图 4.1(b)所示的“七蜂窝”复用模式。注意:实际上每个蜂窝是使用了一组频率。

图 4.1 (a) 简单的蜂窝系统, (b) 7-单元复用模式

图 4.1(b)的每一个蜂窝中的移动单元都有一个基站来提供服务,而所有的基站则由一个“移动电话交换机构(MTSO)”来调度。

同信道干扰 在蜂窝系统里,一个重要的问题是:两个采用相同频率的蜂窝之间有多大的干扰,这一干扰叫做同信道干扰(CCI)。它取决于两个同信道蜂窝之间的距离与蜂窝半径的比,而与发送功率无关。在频率复用的模式下,对于图 4.1(b)所示的“七蜂窝”方式,这个比大约为 4.6[1]。可以证明,这个值产生的信号与同信道干扰的比等于 18dB[1]。

图 4.2 共用信道干扰

跨区 当一个移动单元从蜂窝 A 漫游到蜂窝 B 时将会发生什么事情？因为这时从蜂窝 A 的基站发射来的信号强度不足以维持通信，因此移动单元必须改变为由蜂窝 B 来提供服务。而且，由于相邻的蜂窝并不使用同一组频率，因此，移动单元还必须切换信道。这一过程叫做“跨区”（hand off），是由 MTSO 来完成的。一旦蜂窝 A 里的基站接收到的信号强度低于某一域值的时候，MTSO 就会把移动单元的服务权由交由蜂窝 B 中的基站，并期望后者的距离足够近。这种策略失败的可能性比较高，可能会导致通话的丢失。

图 4.3 跨区发生的问题

为了改进跨区过程，第二代蜂窝系统允许移动单元测量从不同基站发射来的信号的强度，当发觉与第二个基站间路径的损耗足够小时，它就进行跨区操作[1]。

信道路径损耗和多径衰减 在一个移动通信环境里信号的传播是相当复杂的。这里我们只简单描述一下一些重要的概念。在自由空间里传播的信号，其功率损耗与离开发射源的距离 d 的平方成正比。然而在现实上，信号是同时沿着一条直接路径和一条间接的反射路径进行传播的，如图 4.4 所示。在这种情况下可以证明损耗随距离的 4 次方而增加的。在一个拥挤的区域里，实际的损耗情况可能对于某些距离是与 d^2 成正比，而对于另一些距离是与 d^4 成正比。

图 4.4 信号的间接传播以及产生的损耗曲线

除了图 4.4 描述的总的损耗分布情况之外，还有一种机制会使得接收信号的强度随着距离而波动。由于图 4.4 所示的两个信号一般都经受了不同程度的相移，因此有可能在到达接收端时它们的相位刚好相反但是幅度却大致相等，这样净接收到的信号就可能非常弱。这种现象叫做“多径衰减”，如图 4.5 所示，它使得接收器在移动了波长的几分之一距离以后，接收到的信号的强度出现了很大的起伏。

图 4.5 多径损耗曲线

在现实中，由于传输的信号会受到许多建筑物和移动的汽车反射，因而信号的波动是非常没有规律的。尽管如此，总的接收到的信号可以表示成：

$$(4.1)$$

$$(4.2)$$

注意，当 n 很大时，每个累加项服从高斯分布。将第一、二个累加项分别表示成 A 和 B ，我们有

$$(4.3)$$

图 4.6 雷利 (Rayleigh) 分布

式中 $\phi = \tan^{-1}(B/A)$ 。可以证明，幅度 $A_m = \sqrt{A^2 + B^2}$ 服从 Rayleigh 分布（图 4.6）

[2]。在大约 6% 的时间里，损耗使幅度比平均值降低了大约 10dB。

从上面的讨论中，我们可以得出如下的结论：在一个 RF 系统中，发送器的输出功率和接收器的动态范围必须仔细选取，以容忍总的路径损耗（约与 d^4 成正比）和深的多径衰减效应带来的信号强度的变化。虽然，在理论上，多径衰减有可能导致某一距离处的信号幅度等于 0（损耗无穷大）；但这种情况发生的可能性几乎可以忽略，因为在移动环境里一个移

动目标往往会把多径衰减效应“软化”了。

多元化 信号的衰减效应可以通过在接收或者发送信号时增加冗余量来减小。“空间多元化”，或者叫做“天线多元化”是通过把两个或两个以上的天线相隔波长的几分之一放置，以使得接收无衰减的信号的概率较高。

“频率多元化”是指使用多个载波频率。这一想法是：如果两个频率相隔足够远的话，那么信号的衰减效应不大可能同时发生。

“时间多元化”是另一项技术，即把数据多传输一遍或者接收几遍，以克服暂时性的衰减效应。

延迟展宽 假设在一个多径环境里两个信号有大致相同的衰减，却有不同的延时。这是可能的，因为反射或折射材料的吸收系数和相移相差很大，很可能使得两条路径的损耗相同而延时不同。把上述这样的两个信号相加得：

$$x(t) = A \cos \omega(t - \tau_1) + A \cos \omega(t - \tau_2) = 2A \cos \left[\frac{(2\omega t - \omega\tau_1 - \omega\tau_2)}{2} \right] \cos \left[\frac{\omega(\tau_1 - \tau_2)}{2} \right]$$

式中，第二个余弦因子把衰减效应同“延迟谱” $\Delta\tau = \tau_1 - \tau_2$ 联系起来；这里一个重要问题是衰减效应和频率是相关的。如图 4.7 所示，小的延迟谱产生一个相对平坦的衰减，而大的延迟谱则在频谱上引起一个相当大的变化。

在一个多径环境里，多个信号以不同的延迟到达接收器的，所产生的均方根延迟谱大到几个 mS，因而衰减带宽为几百 kHz。这样，总的通信信道有可能因这一个衰减而受抑制。

图 4.7 (a)平坦的与(b)频率选择的衰减

大的延迟谱也引起另一个困难：如果延迟谱同数字调制波形的比特周期可以比拟，那么信号的具有不同延时的多个拷贝都会被接收到，使得码间干扰相当严重。

交叉操作 多径衰减的本质——出错通常发生在成串的比特位上，可以用来帮助我们在信号处理时缓解这一问题。为了减少这些错误的影响，发送器中的基带比特流在调制前要先进行“交叉”操作。一个交叉操作实际上是根据接收器已知的某种算法把比特流在时序上打乱[1]；或者也可以看作是一种没有额外开销的时间多元化（尽管它必然会伴有一定的等待时间）。

4.2 多址访问技术

4.2.1 时分和频分复用

最简单的多址情形是一个收发器的双向通信问题，这一功能称为“双工”。例如，在早期的手持对讲机中，使用者需要按下“讲话”按钮才能进行发送，而把接收通道禁止掉；当要接听时则释放该按钮，从而将发送通道禁止掉。这可以被看作是“时分复用”（TDD）的简单形式。这里发送（Tx）和接收（Rx）通道使用同样的频带范围，但系统只能用一半时间来发送，而用另一半时间来接收。如图 4.8 所示，TDD 通常足够快，从而对使用者来说是感觉不到的。

另一种复用的方法是发送通道和接收通道采用两种不同的频带范围，叫做“频分复用”（FDD），如图 4.9 所示。这一技术利用带通滤波器来隔离这两个通道，确保发送和接收同时进行。由于两个这样的收发器之间不能直接互相通信，因此 Tx 频带必须在某处转换成 Rx 频带。在无线网络中，这种转换是在基站中进行的。

图 4.8 时分转换

下面我们比较这两种复用技术的优缺点，这是很有意义的。在 TDD 中，一个损耗小于 1dB 的 RF 开关接在天线上，交替的接通或者关闭 Tx 和 Rx 通道。尽管发送器的输出功率可

能比接收器的输入信号大 100dB，这两个通道却不会相互干扰，因为在接收的时候发送器是被关闭的。而且，TDD 允许两个收发器之间直接进行通信（“点对点”），这在短距离局域网应用中是一个特别有用的特性。TDD 主要的缺点是，所有附近的移动发送器所产生的强烈信号会落在接收频带内，因而会降低接收器的灵敏度。

在 FDD 系统中，两个前端带通滤波器联合形成了一个“双工滤波器”。虽然 FDD 使接收器不受其他移动单元发射的强烈信号的干扰，但它也存在一系列问题。第一，发送信号泄漏到接收频带的分量通常只被衰减大约 50dB。第二，考虑到要在滤波器的损耗和品质因数之间的进行折衷，通常复用器的损耗要比 TDD 开关大的多。注意，在 Rx 通道中 3dB 的损耗会使得总的噪声系数提高 3dB(第 2 章)，而在滤波器的 Tx 通道上同样的损耗便意味着只有 50%的信号功率到达天线。典型的复用器虽然在 Tx 和 Rx 通道之间提供了可接受的隔离，但它们也确实有 2-3dB 的损耗。

图 4.9 频分转换

FDD 的另一个问题是发送器的输出中频谱泄漏入邻近的信道。这种情况发生在功率放大器被接通或断开（为了节省能量）的时候，或是在驱动调制器的本振在经历一个瞬态的过渡过程时。作为比较，在 TDD 中，这种瞬态过程可以被安排在天线被切换到功放输出端之前结束。

尽管有着上述缺点，FDD 仍应用在许多 RF 系统中，尤其是在蜂窝通信中，这是因为它能把接收器与其它移动发射器的信号隔离开来。

4.2.2 频分多址

为了允许多个收发器相互之间能同时通信，可用的频带范围被划分成多个信道，每一个信道都被指派给一个用户。这项技术叫做频分多址（图 4.10），它在无线电和电视广播中，应该是为大家所熟悉的，只不过在那里信道的指派并不随时间变化。而在多用户的双向通信中，情况恰恰相反，信道的指派只保持到通话的结束；一旦用户挂机以后，这个信道就可以为其他用户使用。注意，在具有 FDD 的 FDMA 中，每个用户被指派了两个信道，一个用于发送而另一个用于接收。

图 4.10 频分多址

FDMA 的相对简单性使得它成为早期蜂窝网络最主要的接入方式，这叫做“模拟调频”系统。然而，在 FDMA 中，能够同时使用的用户的最小数量是由总的可用频带与每个信道的频宽之比决定的，这经常在拥挤的区域产生容量不足的问题。

4.2.3 时分多址

另一个实现多址网络的方法是，每一个用户都使用同样的带宽，但是是在不同的时间使用（时分多址，TDMA）。如图 4.11 所示，TDMA 周期性地将每个收发器打开一段时间（ T_{sl} ），由所有这些时间段构成的总周期叫做一帧（ T_F ）。换句话说，在每个 T_F 秒中，每个用户都可以接入该信道 T_{sl} 秒的时间。

图 4.11 时分多址

当只有一个用户被允许发送数据（如语音）时，所有其它用户产生的数据怎么办呢？为了不丢失信息，这时这些数据被存储（“缓冲”） $T_F - T_{sl}$ 秒的时间，并在某个 T_{sl} 期间以信号串形式传输出去（“TDMA 信号串”由此得名），因为缓冲操作要求数据为数字形式，因此，

TDMA 发送器须对模拟输入信号进行 A/D 变换。数字化也允许我们采用语音压缩和编码技术。

TDMA 系统比起相应的 FDMA 系统来有一系列优点，首先由于在每一个帧里，每个发送器只打开一小段时间，所以在这一帧的其余时间里可以把功率放大器关掉，从而可以节省很多功耗。在实际中，由于稳定过程的要求在正式的时间段开始之前把功率放大器稍微打开一些。第二，由于数字化了的语音可以在时间上被大大压缩，因此所要求的通信带宽可以较小一点，因而总的容量可以较大。也就是说，由于压缩了的语音可以在一个较短的时间段内被传送，因此在每一个帧内里就可以容纳更多的用户。第三，即使和 FDD 一起工作，TDMA 的信号串也可以通过定时使得每个收发器中的接收和发送通道决不会同时工作。

TDMA 需要 A/D 变换、数字调制、时间段和帧同步，这使得它比 FDMA 要更为复杂。不过由于出现了 VLSI DSP，这个问题已不再是致命的弱点了。

在大多数实际的 TDMA 系统中，TDMA 和 FDMA 通常被结合起来用。换句话说，图 4.10 所示的每一个信道在许多用户之间是时间上共享的。

4.2.4 码分多址

在对 FDMA 和 TDMA 的讨论中，我们实际上包含了这样一个事实：即在这些系统中的发送的信号或者是在频域上，或者是在时域上，避免了相互之间的干扰。本质上说，这些信号在这两个领域中的某个域上是正交的。第三种多址技术允许信号在频域和时域上完全重叠，但它采用“正交信息”技术来避免互相干扰。这个可以用一个类比的方法来帮理解[3]。假定在一个拥挤的聚会中，有许多对话正在进行。为了避免互相串扰，可以要求不同组的人使用不同的声调 (!) (FDMA)，或者在一个时刻只允许一组人交谈 (TDMA)。还有一种方法就是：要求不同组的人使用不同的语言。如果所有语言都是正交的（至少在邻近组之间的语言是正交的），并且语音的大小都差不多，那么每一个听众可以“调谐”到某个合适的语言，在所有其他组的谈话都在同时进行着的情况下正确接收信息。

直接序列 CDMA 在“码分多址”技术中，不同的语言是通过正交数字编码方式来产生。在通信开始时，对每一对发送/接收器分配某一代码；并且基带数据的每一比特在调制前都被“翻译”成那种代码。图 4.12(a)给出了一个例子，其中每个基带脉冲通过乘法被替换成一个 8 位的代码。基于 Walsh 迭代方程我们可以给出一种产生正交码符的方法：

$$(4.4)$$

$$(4.5)$$

其中 \bar{W}_n 是通过把每一个元素替换成它的补码来得到，例如：

$$(4.6)$$

图 4.12(b)举了一个 8 位 Walsh 代码符的例子（即，W8 的每一行）：

图 4.12 (a)在 DS-CDMA 中的编码操作，(b)Walsh 代码举例

在接收器中，解调的信号是通过乘以同样的 Walsh 代码符来解码的。也就是说，接收器将信号和 Walsh 代码进行 *相关* 操作以恢复基带数据。

如果另外一个 CDMA 信号存在的话，那么接收到的信号会受到什么样的影响？假设两个的 CDMA 信号 $x_1(t)$ 和 $x_2(t)$ 在同一频带范围内被接收到。把这两个信号分别写为

$x_{BB1}(t) \cdot W_1(t)$ 和 $x_{BB2}(t) \cdot W_2(t)$ ，其中 $W_1(t)$ 和 $W_2(t)$ 是 Walsh 函数，我们将解调器的输出

表示成 $y(t) = [x_{BB1}(t) \cdot W_1(t) + x_{BB2}(t) \cdot W_2(t)] \cdot W_1(t)$ ，于是如果 $W_1(t)$ 和 $W_2(t)$ 是完全正交，那么 $y(t) = x_{BB1}(t) \cdot W_1(t)$ 。但是在实际上， $x_1(t)$ 和 $x_2(t)$ 可能会经历不同的延时，这样 $x_{BB2}(t)$ 就会将破坏 $y(t)$ ，不过，对于长的代码，其结果表现为随机噪声。

图 4.12(a)的编码操作会按代码中脉冲的数目使数据谱的频带展宽。这与我们至今一直强调的频谱效率是背道而驰的。但是，由于 CDMA 使许多用户展宽后的频带都落在同一个频带范围内(图 4.13)，因此这种多址技术的容量并不比 FDMA 和 TDMA 少。事实上，CDMA 甚至有可能获得比这两者更高的容量[4]。

图 4.13 CDMA 里重叠的频谱

CDMA 是“扩展频谱(SS)”通信的特殊情形。扩展频谱是指每一个用户的基带数据被扩展在整个可用带宽上。就这个意义讲，CDMA 也叫做“直接序列 SS”(DS-SS)通信。为了避免和基带数据相混淆，在扩展序列中的每一个脉冲叫做一个“片”，序列的速率叫做“片速率”。因此，频谱是根据片速率和基带比特率之间的比值进行扩展的。

从扩展频谱的观点(图 4.14)来重新考察一下上述解码操作是有意义的。通过乘以 $W_1(t)$ ，所需要的信号被“收缩”了，它的带宽回到原来的值，反之，其他不希望的信号由于和 $W_1(t)$ 的弱相关性，因此与之相乘以后仍然扩展在整个带宽上。如果有大量的用户，那些不希望的信号的扩展频谱可以被看作是 Gaussian 噪声。

图 4.14 去扩展操作

CDMA 的一个重要的特点是它容量的软限制[1]。虽然 FDMA 和 TDMA 中，一旦定义了信道宽度或者时间段之后，最大的用户数就是固定的，而在 CDMA 中增加的数量只是逐渐(线性)地提高噪声底限而已[1]。

DS-CDMA 的一个关键问题就是功率控制。假定如图 4.15 所示，在某一点，收到的有用信号功率可能比其它发送器发出干扰的功率小得多¹。比如，由于后者的距离要近得多。那么即使在收缩操作后，较强的干扰也会大大提高噪声底限，从而降低了接收有用信号的性能。对于多用户来说，这意味着一个大功率的发送器可能事实上导致其他用户不能进行通信。这个问题的严重程度在 FDMA 和 TDMA 要小的多。正因如此，当多个 CDMA 发送器和一个接收器通信的时候，它们必须调整它们的输出功率，使接收器接收到差不多相同的信号强度。为达到这个目的，接收器一直监视着相应于每个发送器的信号强度，并周期性地给它们发送一个功率调整的请求。由于在一个蜂窝系统里，用户是通过基站，而不是直接进行通信的，因此基站必须承担起功率控制的任务。接收到的信号电平通常被控制在相互间 1dB 以内。尽管功率控制增加了系统的复杂性，但是它通常能够缩减移动单元的平均功耗。为了帮助理解这一点，注意到当没有功率控制的时候，移动单元为了能够与基站通信，总是发送出足够大的能量，而不管路径损耗或者衰减效应有多严重。这样即使当信道衰减最小的时候，移动单元仍然在产生最大的输出功率。反之，有了功率控制的时候，只要信道状况有了改善，移动单元就可以以较低的功率发送；这同样也是减少了对其他用户的平均干扰。

¹ 这种情况在我们的聚会比喻中也会出现，即如果两个人谈话的声音比起其他人来说实在太大的话，那么即使使用不同的语言进行交流也会变得很困难。

图 4.15 CDMA 中强干扰信号的影响

跳频 CDMA 另一种开始在 RF 通信中出现的 CDMA 是“跳频 (FH)” CDMA，如图 4.16 所示，这种多址技术可以看作是具有信道伪随机分配特性的 FDMA。在每个发送器中的载波频率会根据某个选定的代码（类似 DS-SS 中的扩展码）进行“跳跃”。这样，尽管，一个发送器在短时间内其频谱可能会和其它发送器的重叠，但是，总的频谱的轨迹，即 PN 码，却可以把每个发送器与其他发送器区分开来。然而，频谱间偶然的重叠增加了出错的概率。

图 4.16 跳码 CDMA

不过由于频谱很少重叠，跳频 CDMA 与 FDMA 比较类似，因此对收到信号不同功率大小的承受能力比直接序列的 CDMA 要好。但是，跳频要求图 4.16 所示的振荡器控制环具有较快的稳定过程，这是在第 8 章中我们将研究的一个重要设计问题。

4.3 无线通信标准

在对调制和多址技术的讨论中，我们可以看到在一个无线系统中打一个电话需要一系列非常复杂的操作。而且，由于一些非理想因素存在（例如噪声和干扰），要求我们拥有一套关于通信参数的精确指标，如 SNR、BER、占用带宽和额外发射等。一个无线标准定义了用在无线系统收发器设计时所有的细节和限制²。这样的标准虽然提供了“边界条件”，但并没有规定实际的发送和接收通道必须如何实现，这允许我们在选择收发器结构时拥有某些灵活性（第 5 章）。

在这一节里，我们将研究一下关于用在蜂窝电话和无绳电话系统中许多标准的简单内容。

4.3.1 先进移动电话服务 (AMPS)

在最早的无线标准中，先进移动电话服务 (AMPS) 标准是采用了带有模拟调频 (FM) 和 FDD 的 FDMA。这个移动电话的发送和接收带宽分别是 824-849MHz 和 869-894MHz，它们被划分成 30-KHz/信道[图 4.17(a)]。这样，这个系统大约可以同时支持 830 个用户³。两个带宽间相隔 20MHz，因此可以使用较低损耗的双工滤波器[图 4.17(b)]。

图 4.17 (a)AMPS 的“空间接口” (b)双向转换器的特性

除了语音通道，AMPS 还需要一些监控信号来初始化、维护和终止一个会话。这些信号也是在同样的信道中被传送的，具体细节请参考[1]。

4.3.2 北美数字标准 (NADS)

作为美国的第一个数字蜂窝系统，北美数字蜂窝(NADC)系统系用 $\pi/4$ -DQPSK 和 FDD 的 TDMA[图 4.18(a)]。为了与 AMPS 兼容，NADC 的设计使用同样的发送和接收带宽，且信道间隔为 30KHz，但它提供了 6 倍于 AMPS 的容量。其中每个信道的比特率是 48.6Kb/s。

图 4.18 NADC(a)空间接口，(b)帧结构，(c)TX 和 RX 的时间槽

用在 NADC 中的 TDMA 帧结构如图 4.18(b)所示[1]。每一帧有 1994 比特长 (40ms)，并含有 6 个时间槽，每个时间段包含大约 260 比特的数据以及 64 比特的控制和同步信息。注

² 某些标准的文档可是有几千多页那么长。

³ 实际上，只有这些用户数的一半用于每个个长距载波。

意，如图 4.18(c)所示，发送和接收的时间段之间相差 1.85ms[5]。这很大程度上使得收发器的一些设计问题变得容易，例如发送信号泄漏到接收带宽里的问题，以及由功率放大器引起的低噪声放大器（LNA）的灵敏度降低问题。

美国大多数的第一代数字电话，实际上能同时兼容 AMPS 和 NADC，以覆盖更多的用户。这些电话叫做双模 AMPS/NADC 系统，是根据电子协会和通信工业协会（TIA）开发的临时标准 54（IS-54）而设计的。

如果读者需要更详细地知道 IS-54 中控制信号和交互技术的详细情况，请参考[1]。

4.3.3 移动通信全球系统（GSM）

GSM（注 4：最初代表 Groupe Speciale Mobile）开始时是作为欧洲统一的无线标准而开发的，现已成为全世界上最广泛使用的蜂窝标准。除了常规的电话之外，GSM 还支持许多其他服务，如传真和 ISDN[1]，这使得移动电话变成了万能的通讯器。

GSM 标准是一个使用 GMSK 调制的 TDMA/FDD 系统，其发送带宽是 890-915MHz，而接收带宽是 935-960MHz[图 4.19(a)]。为了服务 8 个时分的用户，每个信道宽度是 200KHz，每个用户的数据率是 270kb/s。它的 TDMA 帧结构如图 4.19(b)所示。每一帧是 4.615ms 长，含有 8 个时间段，每个时间段包含 114 个比特数据以及其它控制和训练信息。类似于 NADC，TX 和 RX 的时间段之间相差了 3 个时间段，这样发送器和接收器就不会同时工作。系统的总容量是由 25MHz 带宽内的信道数目和每个信道的用户数决定的，大约为 1000。

图 4.19 GSM(a)空间接口及(b)帧结构

4.3.4 高通（Qualcomm）CDMA

基于直接序列 CDMA 的无线标准是由 Qualcomm 公司提出的，并被北美采纳作为 IS-95 标准。它的空间接口使用 FDD，并采用与 IS-54 相同的发送和接收带宽。在移动单元，9.6kb/s 的基带数据被扩展到 1.23MHz，然后再用 OQPSK 调制。反过来，从基站到移动单元的链路则使用 QPSK 调制。这是由于移动单元必须使用功耗效率高的调制技术（第 3 章），而基站却在同时发送多个信道，因而它必须采用线性功率放大器而不管是哪种调制方式。对于这两个链路方向，IS-95 都要求 *相干* 检测，这是通过在通信开始时发送一个较强的“引导音”（例如，没有经过调制的载波）来实现相位同步的。

图 4.20 IS-95 的空间接口

跟上面研究过的其他标准不同，IS-95 从本质上来说更加复杂，它运用了许多技术来增加容量而同时又确保一个比较合理的信号质量。在这里我们只是简单描述了它的一些特点，如果需要知道更多的细节，请读者的参考[3、6、7]。

功率控制 正如 4.2.4 节所说的那样，在 CDMA 中，由基站接收到的来自不同移动单元功率水平相差不应该超过约 1dB。IS-95 中里，在通信开始时，每个移动单元的输出功率是由一个开环过程来控制的，从而实现一个初步的但是很快的调整。接下来，功率将由一个闭环过程来更加精确地予以确定。在开环控制中，移动单元测量从基站发射来的信号的功率，并调整它自己的发射功率使这两者的总和（用 dB 表示）大约为 -73dBm，如果发送和接收通路拥有大致相同的损耗： k dB，并假设基站发送的功率为 P_{bs} ，那么移动单元的输出功率 P_m

满足： $P_{bs} - k + P_m = -73$ dBm。则基站接收到的功率为 $P_m - k$ ，于是我们有 $P_m - k$

$= -73\text{dBm} - P_{bs}$ ，这是一个可以明确定义的数值，因为 P_{bs} 通常是固定的。移动单元的输出

功率可能在几个毫秒内变化约 85dB。

但是由于上面发送和接收路径中损耗相等的假设只是近似的,因此仍需采用闭环功率控制。在实际中,由于两条路径工作在不同的频带上,因此可能经受不同的衰减,为此基站要测量从移动单元发射来的信号功率,并发送一个要求进行功率调节的反馈信号。这个命令每 1.25ms 发送一次,以保证在一个快速衰减的环境中能够及时调整。

频率和时间的多元化 回忆一下 4.1 节所述多径衰减效应对频率具有选择性,这会导致信号传输函数中有一个几 kHz 宽的陷落。由于 IS-95 把频谱扩展到了 1.23MHz,因而它提供了频率的多元化;对于典型的延迟扩展,它大约有 25%的带宽损失。

IS-95 也通过使用时间多元化来从路径信号中得益。这是通过对所接收到的信号的不同延时的副本进行相关操作来实现的(图 4.21)。这个叫做“耙状接收器”的系统,将延迟的副本与一个适当地加权因子 α_j 组合起来,从而在输出端得到最大的信噪比。即如果某一个相关器的输出被破坏的时候,相应的加权因子就会被减少,反之亦然。

图 4.21 耙状接收器

耙状接收器是 CDMA 特有的性质。由于片速率比衰减带宽高很多,并且仔细设计扩展码,使得它和大于片周期的延时之间的相关性可以被忽略,因此多条路径的影响并不引入码间干扰。这样每个相关器可以被同步到其中一个多径信号上。

可变编码率 人类语音中变化的信息速率可以被用来降低每秒被传输比特的平均数。在 IS-95 中,数据率可以在分立的 4 个水平上变化:9600、4800、2400 和 1200b/s。这种安排可以把速率较低的数据缓冲起来,使传输仍然可以以 9600b/s 的速率进行,只是持续的时间相应缩短。例如语音的速率是 2400b/s,那么可以用 50%的时间来积累数据,而只用剩下的 50%时间来传输。这个方法可以进一步减少移动单元发送的平均功率,同时能够节省电池和降低对其他用户的干扰。

软跨区 回忆一下 4.1 节,我们知道当移动单元被分配至另一个基站时,如果信道的中心频率要发生改变的话(例如,在 IS-54 和 GSM 中),那么通话就很有可能被丢失。反之,在 CDMA 中,在一个蜂窝里所有的用户都在同一的信道上通信;因此,当移动单元移至离一个基站较远,而离另一个基站较近的时候,从这两个基站发射来的信号强度都可以用耙状接收器来监测。当它断定较近的基站具有足够强的信号时,就进行跨区操作。这一方法叫做“软跨区”,它可以被看作“先通后断 (make-before-break)”操作,其结果是降低了跨区期间通话丢失的概率。

4.3.5 数字欧洲无绳电话 (DECT)

数字欧洲无绳电话标准 (DECT) 最初是作为欧洲无绳电话网络的标准。然而,这一标准的灵活性使得它在很多范围内的应用也都非常吸引人。DECT 的设计允许接入其它系统例如 GSM,因此它能为局域网的用户,例如,在大楼内的用户提供移动通信。

DECT 使用了 TDMA/FDMA,并在频带范围 1800MHz~1900MHz 上用 TDD 进行操作[图 4.22(a)]。每一个信道是 1.73MHz 宽,在时间上 8 个用户共享,使总容量达到 120。它的调制技术是 GFSK,这是 GMSK 的一种方式。

图 4.22(b)所示的是 DECT TDMA 的帧结构,它包含 24 个时间段(12 个用于发送,12 个用于接收),总的持续时间为 10ms。每一个时间段包含 32 个前导比特,388 个数据比特,和 60 个隔离比特。

与蜂窝标准不同,DECT 的目的是设计低成本和低功耗的手持电话,而不是频谱效率。例如,信道间很宽的间距简化了频率综合器的设计(第 8 章)。

图 4.22 DECT(a)空间接口及(b)帧结构

参考文献

- [1] T. S. Rappaport, *Wireless Communications, Principles and Practice*, Englewood Cliffs, NJ: Prentice Hall, 1996.
- [2] L. W. Couch, *Digital and Analog Communication Systems*, 4th ed., New York: Macmillan, 1993.
- [3] D. P. Whipple, "North American Cellular CDMA," *Hewlett-Packard Journal*, pp. 90-97, December 1993.
- [4] A. Salmasi and K. S. Gilhousen, "On the System Design Aspects of Code Division Multiple Access (CDMA) Applied to Digital Cellular and Personal Communications Networks," *Proc. IEEE Veh. Tech. Conf.*, pp. 57-62, May 1991.
- [5] T. W. Baker and P. E. Bronner, "Create a Low-Cost Cellular Telephone," *Wireless Systems Design*, pp. 23-33, January 1996.
- [6] R. Kerr, *et al.*, "The CDMA Digital Cellular System, An ASIC Overview," *Proceedings of IEEE CICC*, pp. 10.1.1-10.1.7, May 1992.
- [7] J. Hinderling *et al.*, "CDMA Mobile Station Modem ASIC," *Proceedings of IEEE CICC*, pp. 10.2.1-10.2.5, May 1992.

第 5 章 接收发送器结构

在第四章我们对多址技术和标准的研究是把无线系统放在移动通讯站的层次上进行的。在这一章，我们向下深入到接收发送器结构的层次。复杂度、造价、功耗以及外部元件的数量已成为选择接收发送器结构的主要标准。然而随着集成电路技术的发展，这些标准中每一个的相对重要性也在改变，使某些曾经看上去不可实现的方法现在却又回过头来成为似乎合理的解决方案。

在这一章，我们将介绍外差式、零差式、镜像抑制、数字中频和亚采样式接收器，以及直接变换和两步发送器。虽然现在这些结构中只有一些被应用在实际的产品中，但仔细的了解它们的设计问题和综合考虑有助于帮助确定它们各自在新兴应用中的潜力。

为了给大家一个当前最高技术的概念，我们在这一章最后将学习五个现在正在无线系统中使用的具体射频产品。

5.1 概述

无线通信环境，尤其是在城区，经常被称为“敌人”，因为它会给接收发送器的设计强加严格的约束。也许最大的约束就是分配给每个用户的有限带宽(例如：在 IS-54 中为 30kHz；在 GSM 中为 200kHz)。从仙农理论可以知道，这相当于有限的信息传输速率，因而要求使用复杂的技术例如编码、压缩和宽带效率高的调制，甚至对语音号也是如此。

每个用户可用的窄带宽也影响着射频部分的设计。如图 5.1 所示，发送器为了避免频谱向相邻信道的泄漏必须使用窄带调制、放大和滤波，而接收器则必须能在充分抑制相邻信道强干扰的情况下处理有用的信道信号。

图 5.1 无线接收发送器的 (a) 发送器和 (b) 接收器前端

为了更好的感受抑制干扰的困难，假设一个 900MHz 的接收器采用一个带通滤波器来选择一个 30kHz 的信道并同时抑制 60kHz 以外的干扰信道(图 5.2)。如果一个简单的二阶 LC 滤波器用来在离中心频率 45kHz 处提供 60dB 的衰减，那么它的等效 Q 值在 10^7 的量级，这样大小的 Q 值即使是使用象 SAW (表面声波) 器件的滤波器时也很难实现。需要注意的是典型滤波器的损耗和 Q 值之间是相互矛盾的。现在回忆一下第二章中一个有损电路把后级的噪声系数“放大”衰减倍的情形。例如，在图 5.1(b)中，如果带通滤波器有 2dB 的损耗，而 LNA 的噪声系数是 2dB，那么总的噪声系数将增加到 4dB。因此，带通滤波器的选择要同时根据带外抑制和带内损耗，而后者是更关键的参数。

图 5.2 对一个假想的前端带通滤波器所要求的抑制能力

区分频带和信道十分重要：前者包括一个特定标准的用户允许进行通信的整个频谱范围(例如，GSM 的接收带宽从 935MHz 到 960MHz)，而后者是指系统中仅一个用户的带宽(例如，GSM 中为 200kHz)。我们也经常使用术语“带宽选择”和“信道选择”来分别表示抑制带外干扰和抑制信道外干扰(通常是带内的)的操作。上面对 Q 值的计算说明一个实际的前端带通滤波器只能选择所需要的频带，而把信道选择的任务推迟到接收器中某个其他地方进行(图 5.3)。

图 5.3 接收器前端的频带选择

我们还要注意到，前端滤波器的传输函数有着有限的过度带宽。图 5.4 就是一个例子，其中在离接收带 20MHz 处的带外抑制大约是 30dB。换句话说，一个出现在这个频率上的干扰将只被衰减 30dB，这在设计接收路径和频率综合器时是一个严重的问题(第 8 章)。

图 5.4 典型双工器的特性

由于大的带内干扰甚至在前端带通滤波器之后还会伴随接收到的信号，所以后级的非线性尤其是低噪声放大器和混频器的非线性变得十分重要。如在第二章中解释的和在图 5.5 中所示的那样，奇数阶的非线性产生的交调积进入了有用信道内。由于三阶失真通常占主导地位，所以每一级的 IP_3 要足够高以避免信号被交调积破坏。尽管只是幅度失真，但即使信号只是通过相位或者频率承载信息，这种效应仍然十分重要，因为有用信号的过零点被交调积破坏了。

图 5.5 前端非线性性的影响

前端带通滤波器的损耗也会影响发送路径。如果功率放大器产生 1W 的功率，那么在滤波器中衰减 2dB 相当于 370mW 的功率损耗，这比整条接收路径的典型功率还要大！因此带通滤波器必须有最小的带内损耗。同时能充分的抑制发送信号的谐波和带外寄生信号分量。

上面观察也说明通过正确选择调制技术和功率放大器（第 3 章）控制频谱再生长的重要性。功率放大器产生的信道外交调积不能通过带通滤波器进行抑制而必须通过设计使它足够小。

接收发送器设计中的另一个重要考虑是信号的动态范围。当存在多路衰减和路径损耗时，接收信号要求的动态范围的典型值要大于 100dB。由于最小的可检测信号在微伏的量级，所以不仅接收器的输入噪声，而且串扰也变得很关键。在 FDD 系统中当我们考虑接收频带内发送信号有一定衰减的时候，就会出现一个有趣的问题。如图 5.6 所示，如果 TX 功率放大器发送 1W 功率到 50 欧姆的天线，那么天线上的峰峰电压摆幅是 20V。于是从图 5.4 的双工特性我们可以看出，向接收路径的泄漏在 $30mV_{pp}$ （约为 -26dBm 的量级上）。由于 LNA 的 1-dB 压缩点在 -25dBm 附近，所以泄漏的信号可能使 LNA 严重降灵敏。正如在第四章解释的那样，NADC 和 GSM 系统通过错开发送和接收的时间段来避免这一问题，但是模拟的 FDD 标准例如 AMPS 则要求更高的隔离。

图 5.6 由 PA 输出泄漏所造成的 LNA 降灵敏

动态范围的另一个极端是接收器可能会接收到很大的信号，例如，如果接收器接近发送器，则接收到的信号幅度可能会高达几百毫伏。尽管本质上在调相和调频系统中幅度不是关键点，但是接收路径中仍然要正确地处理这一信号。这个问题常常导致使用自动增益控制（AGC）。

最后一个普遍的问题是功率放大器相关的。在最新一代的射频接收发送器中，功率放大器要周期性的打开和关闭以节省功率。但是功率放大器死去的大电流（峰值可达几个安培）会给电源带来极大的噪声，对于有典型输出阻抗的电池，它可以使电池电压改变几百毫伏。因此，在每一个模块中抗噪声和电源抑制变得很重要。

5.2 接收器结构

5.2.1 外差接收器

就像前面提到的，要滤出一个具有很高中心频率和受很大干扰的窄信道要求滤波器具有高的惊人的 Q 值。在外差¹结构中，信号频带被变换到低得多的频率，从而降低了对信道选择滤波器的要求。如图 5.7(a)所示，变换是用混频器完成的，在本章中我们只把混频器看成

¹ 在本书中，我们不区分外差和超外差结构。术语 heterodyne（外差）来自于 hetero（不同的）和 dyne（混合）的组合。

简单的模拟乘法器。为了把中心频率从 ω_1 移到 ω_2 ，信号首先和余弦信号 $A_0 \cos \omega_0 t$ 混频，其中 $\omega_0 = \omega_1 - \omega_2$ ，由此产生了两个频带一个在 ω_2 附近另一个在 $2\omega_1 - \omega_2$ 附近。然后用一个低通滤波器滤去后者。这一操作被称为“下变频混频”或者简称为“下变频”。由于下变频混频器一般都有高噪声，所以在它的前面要有一个低噪声放大器[图 5.7(b)]。余弦信号由一个本地振荡器产生，它的频率此后用 ω_{LO} （=图 5.7 中的 ω_0 ）来表示。

图 5.7 (a) 简单的外差下变频 (b) 采用 LNA 来降低噪声系数

下变频频带的中心频率（ ω_2 ）被称为“中频”（IF），这是一个很重要的参数，它同其它很多性能间存在一个综合考虑的问题。这在以后会进一步的解释。

镜像问题 怎样选择本振频率和中频频率呢？这里最根本的出发点就是“镜像频率”。要理解这个问题，注意到一个简单的模拟乘法器并不能保持他的两个输入频率差的极性，也就是，对于 $x_1(t) = A_1 \cos(\omega_1 t)$ 和 $x_2(t) = A_2 \cos(\omega_2 t)$ ， $x_1(t)$ 和 $x_2(t)$ 的乘积通过低通滤波后具有形式 $\cos(\omega_1 - \omega_2)t$ ，这和 $\cos(\omega_1 + \omega_2)t$ 没有区别。因此，在一个外差结构中，处在本振频率之上和之下对称分布的两个频带会被下变频到相同的中心频率上（图 5.8）。如果接收到的有用频带中心频率在 $\omega_1 (= \omega_{LO} - \omega_{IF})$ ，则它的镜像在 $2\omega_{LO} - \omega_1 (= \omega_{LO} + \omega_{IF})$ 附近，反之亦然。

图 5.8 外差接收中的镜像问题

镜像问题是一个严重的问题。虽然每一个无线标准都会按它自己的用户对信号的发射强加一些限制，但它却不能控制其它频带上的信号。镜像功率因此可能会比有用的信号高得多，所以需要合适的“镜像抑制”。

抑制镜像信号的最常用的方法就是利用放在混频器前面的一个“镜像抑制滤波器”如图 5.9 所示，滤波器设计成使它在有用频带上有较小的损耗，而镜像频带上则有很大的衰减，这两个要求在 $2\omega_{IF}$ 足够大时可以同时得到满足。

图 5.9 通过滤波器来实现镜像抑制

$2\omega_{IF}$ 可以多大呢？回想一下，外差结构的前提就是使中心频率变到足够低的值，从而使用实际的滤波器进行信道选择是可行的。但是当 $2\omega_{IF}$ 增大时，下变频后的频带中心频率也会变大，从而要求更高 Q 值的 IF 滤波器。图 5.10 所示的是对应高低两种中频值的情形以用来说明这两者之间的权衡关系。高的中频会充分抑制镜像，而低的中频会更好的抑制附近的干扰。因此中频的选择取决于三个参数的折中：镜像噪声的数量，有用频带和镜像频带之间的间隔以及镜像抑制滤波器的损耗。要最小化镜像，我们可以增加中频或者在增加滤波器 Q 值的情况下允许较大的损耗。由于 LNA 的增益典型值小于 15dB，滤波器的损耗不能超过几个 dB，从而使权衡考虑简化为镜像噪声和中频值之间的折中。

图 5.10 (a) 高 IF 和 (b) 低 IF 中的镜像抑制和干扰信号抑制

从上面的讨论中，我们可以推断图 5.9 中的外差结构可以从镜像抑制和信道选择这两方面进行综合考虑。由于镜像信号降低了接收器的灵敏度，所以我们说中频的选择要求从灵敏度和选择性两个方面进行权衡。

还有两个因素影响中频的选择：用于不同频率滤波器的可达性和物理尺寸。采用 SAW 或者晶振器件的滤波器被使用在某些频率上，例如 10.7MHz（在调频收音机中），71MHz 等，因此它们的价格比必须把它们设计为用于其它频率时要低，此外，为了减小移动系统的体积尺寸因数，更小的滤波器并且因此更高的中频更有吸引力。

外差结构的一个重要缺点是镜像抑制滤波器通常用外部的无源部件实现。这会进一步要求前级，也就是 LNA，驱动滤波器 50 欧姆的输入电阻，这不可避免地会在放大器的增益、噪声系数、稳定性和功耗之间提出更加严重的权衡考虑问题（第六章）。在 FDD 系统中，如果中频比较高，双工器也是可以用来抑制镜像，从而允许直接把 LNA 直接连接到混频器上。

本地振荡器的频率 ω_{LO} 可以比有用频带的中心频率高或低。这两种情况分别称为“高边注入”和“低边注入”²，它们会引起不同的问题。一方面，人们希望使用后者来降低本振的频率从而使振荡器的设计比较容易（第七章）。另一方面，如果镜像频带低于或高于有用信号频带时会有不同数量的噪声（例如，一个是用来进行低信号电平通信，而另一个是用于高功率的无线标准），那么 ω_{LO} 必须谨慎选择以避免有噪声的镜像频带。

半中频的问题 “半中频问题” [3] 是外差接收器中的一个有意思的效应。如图 5.11，假设除了位于 ω_m 周围的有用频带外，在 $(\omega_m + \omega_{LO})/2$ 处还存在一个干扰源，也就是，在有用频带和本振频率之间一半中频外的信号，也被接收到了。如果在下变频的路径上，干扰源经历了一个二阶失真而本振也包含一个明显的二次谐波，那么中频的输出中就有一个在 $|\omega_m + \omega_{LO} - 2\omega_{LO}| = \omega_{IF}$ 处的分量。另一个可能就是干扰变换成为 $(\omega_m - \omega_{LO}) = \omega_{IF}/2$ ，并继而经历一个基带的二阶失真，这将使它的二次谐波落入下变频后有用频带上。

图 5.11 外差接收中的半中频问题

为了抑制半中频现象，射频和中频路径中的二阶失真必须最小化，而且本振必须维持 50% 的占空比。镜像抑制滤波器在 $(\omega_m + \omega_{LO})/2$ 的地方有足够的衰减也是必要的。

双中频拓扑结构 在图 5.7 中简单的外差结构里，灵敏度和选择性之间的权衡经常十分严重：如果中频高了，镜像可能被抑制住，但是要实现完全的信道选择就十分困难，反过来也一样。为了解决这个问题，外差的概念可以扩展到多下变频，每个下变频之后进行滤波和放大。如图 5.12 所示，这一技术在中心频率不断变小的过程中进行部分的信道选择，从而降低了对每个滤波器 Q 值的要求。现今大多数的射频接收器使用两级下变频，因此被称为“双中频”。注意第二次下变频仍然有镜像问题。对于窄信道标准如 IS-54，第二中频常等于 455kHz，但是对于宽信道应用例如 DECT，它可能会达到几个兆赫。不管怎么样，在现今系统中这些数值的变化都比较大。

图 5.12 双中频外差接收器

² 它们有时也分别被称为“上差”和“下差”。

回想一下第二章中讲过，在一个级联的增益级中噪声系数在前端是最关键的，而线性度在后端是最关键的。因此优化的设计可以根据每一级之前的总增益把该级的噪声系数和 IP_3 扩大。现在假设图 5.12 中的接收器从 A 到 G 的总的增益例如是 40dB。如果这两个中频滤波器没有提供信道选择，那么中频放大器的 IP_3 就会需要比 LNA 的 IP_3 高出 40dB 以上，例如，在 30dBm 附近。但是，在合理的噪声、功耗和增益的情况下要获得这样高的线性度是十分困难的，尤其是如果电路还要工作在低电源电压的情况下。实际上，由于每个中频滤波器把相邻信道的干扰在一定程度上进行了抑制，所以对滤波器以后各级的线性度要求有所降低。这就是常常被不严格的说成是“一分贝的增益需要一分贝的预滤波”。

图 5.12 中所示的是双中频接收器中不同点处的频谱。前端滤波器选择频带并对镜像进行一定的抑制。在进行了放大和镜像抑制滤波之后得到的就是 C 点的频谱。然后一个线性度足够好的混频器把有用信道和它邻近的干扰源变换到第一中频。在 BPF_3 中的部分信道选择，使可以采用对线性度要求不高的第二混频器。接着该频谱被变换到第二中频，同时由 BPF_4 把干扰抑制到可以接受的较低程度。

在第二中频处发生了什么呢？在模拟调频系统中，解调就是在这个频率进行的，于是重新产生了模拟的基带信号。在数字调制系统中，当第二次下变频把频谱变换到零频率的时候一般同时产生同相 (I) 信号分量和正交 (Q) 信号分量 (图 5.13)。(这被称为“单中频”拓扑结构，因为它并不包含第二个中间频率)

由于接收到的信号经历了多路衰减和 ISI，它在 I/Q 检波后被变换成数字形式以进行后续处理。取决于滤波器的干扰抑制和 AGC 电路提供的增益范围，这里对 A/D 变换器分辨率的要求大约在 4 到 10 比特之间变化。

我们对于外差结构的分析说明了接收路径中每一级的 NF, IP_3 和增益的选择依赖于它前级和后级的这些指标，从而要求在结构和电路层次上进行相当多的反复考虑，以使组成接收器的各个模块达到可以接受的增益分布。此外，就像第六章讲的那样，每一个混频器都产生许多假信号分量，这些分量的频率使与射频和中频信号以及振荡器的频率有关。其中某些分量可能会落入有用信道中，从而降低了信号的质量。因此接收器的“频率设计”在可以达到的整个性能中扮演着关键的角色。

图 5.13 正交下变频

好几十年以来，外差结构的选择性和灵敏度使它一直是射频系统的首选。尽管外差结构很复杂，而且需要大量的外部元件，但是它仍然被认为是最可靠的接收技术。

5.2.2 零差接收器

在我们学习外差接收器的时候，读者可能已经想到为什么不把射频频谱在第一下变频时就直接变换到基带。这类接收器被称为“零差”，“直接下变频”或者“零中频”结构³，它相对于外差结构有很多不同的问题。

图 5.14 是一个简单的零差接收器，它的本振频率等于输入的载波频率。注意到信道的选择只需要一个截止特性较好的低通滤波器。图 5.14 (a) 中的电路只是对双边带调幅信号能正确工作，因为它把输入频谱的正负部分重叠了。对于频率调制和相位调制信号，下变频必须提供正交输出[图 5.14 (b)]以避免信息的损失。这是因为 FM 或者 QPSK 频谱的两个边带带有不同的信息，所以在变换为零频率的时候必须被分离为两个正交相位。

零差结构的简单性相对于外差结构有两个很重要的优点。首先，镜像问题被克服了，因为 $\omega_{IF} = 0$ 。因此，不需要镜像滤波器，而且 LNA 也不需要驱动 50 欧姆的负载。其次，IF

³ Homodyne (零差) 这个术语来自于 homo (相同的) 和 dyne (混合)，历史上只是用于相干接收。我们在这里不做区分。

SAW 滤波器和后续的下变频级可代之以适合单片集成的低通滤波器和基带放大器。

图 5.14 (a) 简单零差接收器, (b) 采用正交下变频的零差接收器

要是零差结构这么简单, 为什么它没有在 RF 系统中很流行呢? 直接变换频谱到零频率需要解决很多的问题, 而这些问题在外差接收器中根本不存在或者并不严重。

信道选择 通过有源低通滤波器抑制信道外的干扰比使用无源滤波器更加困难, 根本的原因是有源滤波器与相应的无源滤波器相比需要噪声-线性-功耗之间更严格的权衡。图 5.14 (b) 中每一个分支的基带处理过程可以采用图 5.15[6]中三种排列中的任何一个。

图 5.15 基带功能块的三种排列方式

在图 5.15 (a) 中, 一个低通滤波器用来抑制信道外的干扰, 从而允许 A1 是一个非线性、高增益的放大器, 而且 ADC 可以有一个中等程度的动态范围 (大约为 4 到 8 比特, 这取决于 RF 域的增益控制和调制方式)。但是低通滤波器使噪声-线性度之间的权衡考虑更加困难。

图 5.15 (b) 中的第二种排列, 虽然要求放大器有较高的性能但却可以降低对低通滤波器噪声的要求。在信道滤波之前可以使用一个线性化的单级差分放大器来增加一些增益。此外, 还可以在 LPF 和 ADC 之间加上另外一个放大器用来克服后者的噪声。

图 5.15 (c) 中的第三种意味着有可能在数字域进行信号滤波。在这种情况下, ADC 必须同时具有高的线性度以实现干扰交调最小的数字化, 以及比信号电平低许多的热噪声和量化的噪声底限, 这大概在几百微伏的范围内。

直流偏移 由于在零差结构中下变频后的频带扩展到零频率, 外部的偏移电压有可能破坏信号, 更重要的是, 它可能使以后各级饱和。要了解偏移的来源和影响, 可以考虑图 5.16 中的接收器, 其中, LPF 后面接着一个放大器和一个 ADC 变换器。让我们做两方面的观察。

图 5.16 (a) 本振信号和 (b) 强干扰信号的自混合 [5]

首先, 在本振端口与混频器和 LNA 输入端口之间的隔离并不是绝对的隔离; 也就是从 LO 端口到 A 点和 B 点之间存在一定量的馈通[图 5.16 (a)]。这一效应被称为“LO 泄漏”, 它来源于电容和衬底耦合, 并且如果 LO 信号是由外部提供的则引线耦合也会造成 LO 泄漏。出现在 LNA 和混频器输入端的泄漏信号现在和 LO 的信号相混合, 从而在 C 点产生了一个直流分量。这种现象叫做“自混合 (self-mixing)”。当大的干扰信号从 LNA 或者混频器的输入端泄漏到 LO 端口时也会和自己相乘从而产生类似的效应[图 5.16 (b)]。

其次, 从天线到点 X 的总增益的典型值约为 80dB 到 100dB, 从而可以把微伏量级的输入信号放大到可以通过低成本和低功耗的 ADC 进行数字化的电平。在这一增益中 LNA/混频器组合提供了 25dB 到 30dB 的增益。

通过上面的观察, 我们可以得到对于自混合造成的偏移的粗略估计从而理解这个问题。假设, 图 5.16 (a) 中 LO 信号的峰峰摆幅值是 0.63V (相当于在 50 欧系统中大约为 0dBm), 当它耦合到 A 点时衰减为 60dB。如果 LNA/混频器组合的增益是 30dB, 那么在混频器输出端产生的偏移在 10mV 的量级。我们还注意到在这一点的有效信号电平可以低到约 30 μ V_{rms}。因此如果这个偏移电压直接被剩下的 50dB 到 70dB 的增益放大的话, 它会使得后面的电路饱和, 从而阻止了有用信号的放大。

如果自混合效应是时变的, 那么偏移的问题会被恶化。这发生在 LO 信号泄漏到天线, 然后被发射出去, 再被移动物体反射回接收器时。例如, 当一辆汽车高速行驶的时候, 反射可能会迅速变化。在这些情况下, 要想区分时变的偏移和实际的信号可能非常困难。

在上面的讨论中，我们推想零差接收器会要求某些消除偏移的手段。虽然高通滤波器在这里看上去是很自然的解决方案[4]，但是有两点阻碍着这个技术的应用。首先，如第三章所解释的，许多常用信号的频谱在零频率处都有一个峰值，也就是说，它在直流附近有显著的能量（信息）分布。对于一个 200kHz 的信道，如果高通滤波器只是滤掉 0 到 20Hz 的频带，那么波特误差率就上升到了 10^{-3} 以上 [5]，这意味要求有很低的拐角频率。其次，除了要求又不现实的大电容外，这种方法仍然不能跟踪快速变化的偏移电压，因此只能做到粗略的消除。

偏移问题可以通过下面两种技术之一减轻。首先，在发送器端的基带信号可以被编码，使得调制和下变频后它在直流附近有很少的能量。这被称为“无 DC 编码(DC-free coding)”，它特别适用于宽信道，例如，在 DECT 中，可以浪费几 kHz 的信道而不会显著降低数据率。

第二种技术是在数字无线标准中采用空闲时间间隔来消除偏移。图 5.17 是一个例子，其中一个电容存储相邻 TDMA 脉冲串之间的偏移，并在接收数据的时候引入一个事实上为零的拐角频率。对于一个典型的几个毫秒的 TDMA 帧，偏移的消除在足够的频率下完成以考虑因移动物体造成的变化。

图 5.17 TDMA 系统中简单的偏移消除

实际上，图 5.17 中由于 S1 造成的被存储在电容上的热噪声 (kT/C 噪声) 是十分显著的，所以需要很大的 C_1 。例如，如果天线接收到的一个 $1V_{rms}$ 的信号在偏移消除以前以一个 30dB 的增益被放大，那么 C_1 必须至少 200pF 使得 kT/C 的平方根 ($\sqrt{kT/C}$) 保持在低于信号 15dB。对于使用差分电路的 I 和 Q 信道，这需要 4 个大的电容。但是，如果偏移在基带预放大以前就被消除了[图 5.15 (b)]，那么偏移存储电容的值可以被减小。

接收器中偏移消除遇到的一个普遍的困难就是干扰可能会和偏移一起被存储。就像前面解释的那样，这一情况的发生是由于 LO 信号被周围物体反射回来以后必须被包括进偏移的消除中，因此天线这个时候不能断开（或短路）。虽则实际信号（TDMA 脉冲串）的时序可以被定义的很好，但干扰可以在任何时候出现。一个减轻这个问题的可行方法就是对偏移（和干扰）进行多次采样然后取平均。

我们还应当注意，在外差结构中偏移的问题并不严重。由于第一个 LO 频率并不等于输入的载波频率，自混合只对干扰源才出现[图 5.16 (b)]，由此产生的直流偏移可以被去除，因为 IF 信号与零频率相距还很远。此外，在模拟 FM 系统中第二 IF 不是零，而在数字调制系统中第一 IF 处的信号放大和（部分）信道滤波简化了第二下变频之后的偏移消除。

I/Q 失配 如图 5.14 (b) 所示，对于相位和频率调制技术，零差接收器必须采用正交混频。这要求 RF 信号或者 LO 输出移相 90 度（图 5.18）。就像本章后面和第七章要解释的那样，相移 RF 信号一般要求严格的噪声-功耗-增益间的权衡，所以更希望采用图 5.18 (b) 中的拓扑结构。在任何一种情况下，与名义上 90 度相移的误差，以及 I 和 Q 信号在幅度上的失配破坏了下变频后信号的分布星座，因此增加了比特误差率。注意就像图 5.19 所示，I 和 Q 路径中的每一部分都会影响增益和相位差。

图 5.18 (a) RF 路径 (b) LO 路径中的正交信号产生

图 5.19 各级对 I/Q 失配的影响

为了更深入的了解 I/Q 不平衡产生的效应，假设接受到的信号是 $x_{in}(t) = a \cos \omega_c t + b \sin \omega_c t$ ，其中 a 和 b 等于 1 或者 -1。现在让我们假设 LO 信号的 I 和 Q 相位是

$$x_{LO,I}(t) = 2\left(1 + \frac{\varepsilon}{2}\right) \cos\left(\omega_c t + \frac{\theta}{2}\right) \quad 5.1$$

$$x_{LO,Q}(t) = 2\left(1 - \frac{\varepsilon}{2}\right) \sin\left(\omega_c t - \frac{\theta}{2}\right) \quad 5.2$$

图 5.20 I/Q 失配对 QPSK 信号星座的影响：(a) 增益误差，(b) 相位误差

其中因子 2 是为了简化计算结果而引入的， ε 和 θ 分别代表幅度和相位的误差。把 $x_{in}(t)$ 乘以这两个 LO 相位，然后对该结果进行低通滤波，我们得到如下基带信号：

$$x_{BB,I}(t) = a\left(1 + \frac{\varepsilon}{2}\right) \cos \frac{\theta}{2} - b\left(1 + \frac{\varepsilon}{2}\right) \sin \frac{\theta}{2} \quad 5.3$$

$$x_{BB,Q}(t) = -a\left(1 - \frac{\varepsilon}{2}\right) \sin \frac{\theta}{2} + b\left(1 - \frac{\varepsilon}{2}\right) \cos \frac{\theta}{2} \quad 5.4$$

图 5.20 是当 ε 和 θ 不为零时结果信号的分布星座。考察下变频后时域的 QPSK 信号可以更好的看出这一效应（图 5.21）。增益误差只是相当于幅度的放大因子不是 1。反之，相位不平衡则使每一个信道都叠加上了其它信道的一部分数据脉冲而破坏，如果 I 和 Q 数据流是不相关的，则这会从本质上降低信噪比。

图 5.21 I/Q 失配对解调后 QPSK 波形的影响：(a) 增益误差，(b) 相位误差

实际上，希望使幅度失配小于 1dB 相位误差小于 5 度，但是这些限制值取决于调制的类型[7]。

虽然外差结构也可以在最后一级采用 I/Q 下变频（图 5.13），但是它对失配的要求要宽松的多。这有两个原因。首先，由于 I 和 Q 相位分开时的频率比相应的外差结构要低一到两个数量级，因此这两个路径对由寄生产生的失配不敏感的多。同时在 IC 设计中，较低低频率也允许使用大的器件来改善失配而又不增加额外的功耗。其次，在外差接收器中，信号在 I/Q 分离以前就被放大了 50 到 60dB，因此在之后只需要再有一到两级。相反，零差结构中的每一信道要采用几级增益和滤波，每一级都会产生失配。我们还要注意，外差结构在数字域中能进行 I/Q 分离以避免失配（5.2.4 节），而零差接收器就不行。

I/Q 的失配问题曾经是离散实现时的障碍，但是随着单片集成实现了更多的零差结构部件而是这个问题趋于改善。此外，由于失配会随时间推移变得可以忽略，因此可以用信号处理技术来校正信号空间中的点。

偶阶失真 到现在虽然只把奇数阶的失真当作一个干扰源（第二章），但是在零差下变频中偶阶失真也会成为问题。假设，如图 5.22 所示，在 LNA 中有两个接近有用信道的强干扰信号经过一个非线性变换如 $y(t) = \alpha_1 x(t) + \alpha_2 x^2(t)$ 。如果 $x(t) = A_1 \cos \omega_1 t + A_2 \cos \omega_2 t$ ，

那么 $y(t)$ 包括项 $\alpha_2 A_1 A_2 \cos(\omega_1 - \omega_2)t$ ，这说明存在偶阶失真时，两个高频干扰信号产生了

低频的拍。在理想混频器中这一项与 $\cos \omega_{LO} t$ 相乘就会变换到高频，也就不重要了。但是，

实际上的混频器从 RF 输入到 IF 输出都存在一定的馈通。就像第六章解释的那样，这是由于混频器一般会有一些不对称从而他们的操作可以被看成是 $v_{RF}(t)(a + A \cos \omega_{LO}t)$ ，其中 a 是一个常数。这样， $v_{RF}(t)$ 的一部分就出现在输出端而没有进行任何频率变换。在典型的差分混频器中，拍信号耦合到输出时只被衰减 30 到 40dB。

我们可以从另一个角度来研究偶阶失真问题。假设接受到的信号除了有相位和频率调制外，还有以些幅度调制。这会发生发生在发送器（第三章）滤波时，或者由于传播时的干扰和衰减。如果 $x_m(t) = (A + \varepsilon \cos \omega_m t)(a \cos \omega_c t + b \sin \omega_c t)$ ，其中 $\varepsilon \cos \omega_m t$ 代表低频的幅度调制信号，那么二阶失真产生了如 $(a^2 + b^2)A\varepsilon \cos \omega_m t$ 这样的项。由于这一项是变换到基带的调制信号，所以人们常说偶阶失真解调了 AM 分量。像上面提到的一样，这个信号通过混频器时有一定的衰减，从而破坏了下变频后的有用信号。

图 5.22 偶阶失真对干扰信号源的影响

在我们说明过程中，假设了只有 LNA 有偶阶失真。实际上，混频器的 RF 端口也受到同样的效应，因此设计时要格外小心，因为加到混频器的信号是 LNA 放大过的信号，因此可能会产生显著的失真。

二阶失真可以用“二阶交调点” IP_2 来表征。类似于定义 IP_3 的方法（第二章），将两个等幅度的干扰加到输入端，并在输出端观察它们产生的低频拍信号。画出拍信号的功率和输入信号的功率，并把它们外推就可以得到 IP_2 。

我们可以使用差分 LNA 和混频器来抑制偶阶失真么？这里有两点需要考虑。首先，天线和双工滤波器通常都是单端的，因为它们必须与发送路径上的单端功率放大器一起工作。因此必须有办法把接收到的信号转换为差分形式。困难在于这些方法例如变压器，一般来说在高频时会有几个分贝的损耗，从而直接加大了总的噪声系数。其次，如果 LNA 被设计成为差分电路，那么它比相应的单端电路需要更高的功耗才能达到相同的噪声系数。

闪烁噪声 就像前面说的，通过前端 LNA 和混频器的增益的典型值约 30dB，因此产生的基带信号电平在几十微伏的范围内。因此，它后面各级比如放大器和滤波器的输入噪声仍然十分关键。尤其是由于下变频后的频谱扩展到零频率，器件的 $1/f$ 噪声从本质上破坏了信号，这在 MOS 实现时是一个很严重的问题。因此，希望在 RF 范围内有相对较高的增益，例如，可以使用有源混频器来代替无源混频器（第六章）。

闪烁噪声的影响可以通过几种技术的组合来减小。因为混频器后的各级工作在相对低的频率上，所以它们可以使用很大尺寸的器件来最大程度的减小闪烁噪声的幅度。而且，周期地消除偏移，例如 TDMA 脉冲串之间消除偏移也可以降低在约 $1/TC$ 以下的噪声分量，其中 TC 是指连续的偏移消除之间的时间⁴。此外，如果使用无 DC 编码，那么下变频后的信号以及因此噪声都可以进行高通滤波。

LO 泄漏 除了会引入 DC 偏移以外，LO 信号到天线的泄漏以及从那里的辐射还会在利用相同无线标准的其它接收器频带中产生干扰。无线标准的设计以及联邦通信协会（FCC）的规定都对带内 LO 辐射量的上界进行了限制，典型值是 -50 到 -80dBm 之间。

5.2.3 镜像抑制接收器

决定着在外差结构中是否使用镜像抑制滤波器的权衡考虑激发了 RF 设计者寻找其它抑

⁴ 这种效应在模拟设计中被称为“相关双采样”。

制镜像的技术。在这一节中我们描述两个这样的技术。

在学习这些结构之前，我们需要先定义一个“相移 90 度”的操作。如果一个窄带信号的频谱乘上 $G(\omega) = -j \operatorname{sgn}(\omega)$ ，其中 $\operatorname{sgn}(\omega)$ 是符号函数，那么这个信号就相移了 90 度⁵。

如图 5.23 所示，这种操作把 $\sin \omega t$ 变换成为 $-\cos \omega t$ ，而把 $\cos \omega t$ 变换成 $\sin \omega t$ ；也就是把 t 用 $t-T/4$ 来代替，其中 T 是信号周期。注意对于一个实信号，移相后的信号仍然是实信号，但在正频率和负频率处有不同的相位变化。

一个信号怎样相移 90 度呢？图 5.24 中的 RC-CR 网络经常被用来达到这一目的。在这里对于一个频率为 ω 的正弦输入， $V_{\text{out1}}(t)$ 和 $V_{\text{out2}}(t)$ 的相移分别是 $\pi/2 - \tan^{-1}(RC\omega)$ 和 $-\tan^{-1}(RC\omega)$ 。因此， V_{out1} 和 V_{out2} 在任何频率下都有 90 度的相位差。在第七章我们还要回来讨论这个电路。

图 5.23 (a) 时域和 (b) 频域的 90 度相移

镜像抑制结构的想法就是对信号及其镜像进行不同的处理，以使镜像为它的负的副本相抵消。可以对信号和镜像进行区分是因为它们分别在 LO 频率的两边。

图 5.24 90 度相移网络

Hartley 结构 图 5.25 是由 Hartley 于 1928 年提出的一个镜像抑制结构[8]，它源于单边带 (SSB) 调制器。Hartley 的电路把 RF 输入和本振的两个正交信号 $\sin \omega_{LO} t$ 和 $\cos \omega_{LO} t$ 进行混合，然后进行低通滤波，并把它们其中之一相移 90 度，后再叠加到一起。为了解它的基本原理，假设输入信号是 $x(t) = A_{RF} \cos \omega_{RF} t + A_{im} \cos \omega_{im} t$ ，其中第一项代表有用信道

而第二项为镜像信号。不失一般性，我们可以假设低边带注入： $\omega_{RF} - \omega_{LO} = \omega_{LO} - \omega_{im}$ 。

把 $x(t)$ 乘以 LO 的两个正交信号并忽略高频分量，我们在 A 点和 B 点分别得到如下信号：

$$x_A(t) = \frac{A_{RF}}{2} \sin(\omega_{LO} - \omega_{RF})t + \frac{A_{im}}{2} \sin(\omega_{LO} - \omega_{im})t \quad 5.5$$

$$x_B(t) = \frac{A_{RF}}{2} \cos(\omega_{LO} - \omega_{RF})t + \frac{A_{im}}{2} \cos(\omega_{LO} - \omega_{im})t \quad 5.6$$

把 (5.4) 写成

$$x_A(t) = -\frac{A_{RF}}{2} \sin(\omega_{RF} - \omega_{LO})t + \frac{A_{im}}{2} \sin(\omega_{LO} - \omega_{im})t \quad 5.7$$

并利用图 5.23(a) 的结果，可以得到

$$x_C(t) = +\frac{A_{RF}}{2} \cos(\omega_{RF} - \omega_{LO})t - \frac{A_{im}}{2} \cos(\omega_{LO} - \omega_{im})t \quad 5.8$$

把 $x_C(t)$ 和 $x_B(t)$ 加起来后，我们在输出端得到 $A_{RF} \cos(\omega_{LO} - \omega_{RF})t$ 。这样，RF 信号就在没有受镜像影响的情况下完成下变频。这里的关键点是，B 和 C 点的信号分量有相同

⁵ 乘以 $G(\omega)$ 等价于进行希尔伯特变换。

的极性，而镜像分量有相反的极性。这是因为 90 度相移操作把 $\omega_{LO} - \omega_{RC} < 0$ 和 $\omega_{LO} - \omega_{im} > 0$ 区分开来了。

图 5.25 Hartley 镜像抑制接收器

Hartley 电路也可以通过图解来分析。假设输入信号由有用信号和它的镜像组成（图 5.26）。把输入信号频谱和 $\sin \omega_{LO}t$ 以及 $\cos \omega_{LO}t$ 的频谱进行卷积，然后进行低通滤波，就分别得到在 A 点和 B 点的频谱 $X_A(\omega)$ 和 $X_B(\omega)$ 。在 $X_A(\omega)$ 中， $\pm j$ 表示该频谱的相应部分和同一个数相乘。90 度的相移操作把 $X_A(\omega)$ 的负频率部分乘以 $+j$ ，正频率部分乘以 $-j$ ，从而产生了图中所示的 $X_C(\omega)$ 。因此 $X_C(\omega)$ 和 $X_B(\omega)$ 的和就与镜像无关了。

图 5.26 Hartley 结构的图解分析

我们要注意实际上 90 度的相移是在一条路径上进行 $+45$ 度相移和在另一个路径上进行 -45 度相移来实现的（图 5.27）。

图 5.27 采用分立相移级的镜像抑制接收器

Hartley 结构的根本的缺点就是它对失配敏感。如果 LO 的两相信号并不完全正交，或者图 5.27 中上下两条路径中的增益和相移不相同，那么上面所说的相消就会不完全而且镜像信号就会破坏下变频后的信号。为了更好的理解这一影响，让我们来推导（5.8）并假设 LO 的两相信号是 $A_{LO} \sin \omega_{LO}t$ 和 $(A_{LO} + \varepsilon) \cos(\omega_{LO}t + \theta)$ ，其中 ε 和 θ 分别表示幅度的失配和相位的失衡。于是图 5.25 中 A 点和 B 点的波形等于⁶：

$$x_A(t) = \frac{A_{LO}A_{RF}}{2} \sin(\omega_{LO} - \omega_{RF})t + \frac{A_{LO}A_{im}}{2} \sin(\omega_{LO} - \omega_{im})t \quad (5.9)$$

$$\begin{aligned} x_B(t) = (A_{LO} + \varepsilon) \frac{A_{RF}}{2} \cos[(\omega_{LO} - \omega_{RF})t + \theta] \\ + (A_{LO} + \varepsilon) \frac{A_{im}}{2} \cos[(\omega_{LO} - \omega_{im})t + \theta] \end{aligned} \quad (5.10)$$

由此得：

$$x_C(t) = A_{LO} \left[\frac{A_{RF}}{2} \cos(\omega_{LO} - \omega_{RF})t - \frac{A_{im}}{2} \cos(\omega_{LO} - \omega_{im})t \right] \quad (5.11)$$

将 $x_B(t)$ 与 $x_C(t)$ 相加同时，产生了有用信号和下变频之后镜像信号的一部分：

⁶ 读者也许注意到 $x_A(t)$ 和 $x_B(t)$ 的量纲都是（伏特）²，混频器需要一个合适的比例系数来产生正确的输出量纲。

$$x_{sig}(t) = \frac{(A_{LO} + \varepsilon)A_{RF}}{2} \cos[(\omega_{LO} - \omega_{RF})t + \theta] + \frac{A_{LO}A_{RF}}{2} \cos(\omega_{LO} - \omega_{RF})t \quad (5.12)$$

$$x_{im}(t) = \frac{(A_{LO} + \varepsilon)A_{im}}{2} \cos[(\omega_{LO} - \omega_{im})t + \theta] - \frac{A_{LO}A_{im}}{2} \cos(\omega_{LO} - \omega_{im})t \quad (5.13)$$

现在让我们来计算输出端的镜像信号比, 也就是 $x_{im}(t)$ 的平均功率 P_{im} 和 $x_{sig}(t)$ 的平均功率 P_{sig} 的比:

$$\frac{P_{im}}{P_{sig}} \Big|_{out} = \frac{A_{im}^2 (A_{LO} + \varepsilon)^2 - 2A_{LO}(A_{LO} + \varepsilon) \cos \theta + A_{LO}^2}{A_{RF}^2 (A_{LO} + \varepsilon)^2 + 2A_{LO}(A_{LO} + \varepsilon) \cos \theta + A_{LO}^2} \quad (5.14)$$

注意, 由于 A_{im}^2 / A_{RF}^2 是输入端的镜像信号比, 我们定义“镜像抑制比”(IRR) 为 P_{im}/P_{sig}

除以 A_{im}^2 / A_{RF}^2

$$IRR = \frac{A^2 - 2AB \cos \theta + B^2}{A^2 + 2AB \cos \theta + B^2} \quad (5.15)$$

其中 $A = A_{LO}$, $B = A_{LO} + \varepsilon$ 体现了两个信道的增益不同。对于 $\varepsilon \ll A_{LO}$ 和 $\theta \ll 1rad$,

(5.15) 可以简化为

$$IRR = \frac{(\Delta A / A)^2 + \theta^2}{4} \quad (5.16)$$

其中 $\Delta A / A = \varepsilon / A_{LO}$ 表示相对增益失配, θ 则用弧度表示。

上面的分析只包括 LO 信号的增益和相位失配。读者可自己证明, 在图 5.25 中的混频器, LPF, 加法器的两个端口上的小失配以及由 90 度相移级引起的附加误差都可以集中用公式(5.16)中的量 $\Delta A / A$ 和 θ 表示出来。对于集成电路中典型的匹配情况, 镜像抑制在 30 到 40dB 的范围内, 这很可能是由于 0.2 到 0.6dB 的增益失配和 1 到 5 度的相位失衡共同作用的结果。

在 Hartley 结构中的一个关键问题就是由 90 度相移操作引起的增益失配。例如在图 5.27 中, 两个相移级只在 $\omega_{IF} = 1/(RC)$ 时才有相等的增益。因此, 对于一个给定的 ω_{IF} , 如果 R 和 C 的绝对值随着温度或者工艺而变化, 那么增益将会不相等。具体地, 如果电阻等于 $R + \Delta R$, 电容等于 $C + \Delta C$, 则相对的增益偏移为

$$\frac{\Delta A}{A} = \frac{(R + \Delta R)(C + \Delta C)\omega - 1}{\sqrt{1 + (R + \Delta R)^2(C + \Delta C)^2\omega^2}} \div \frac{1}{\sqrt{1 + R^2C^2\omega^2}} \quad (5.17)$$

由于在 ω_{IF} 附近, $RC\omega \approx 1$, 所以

$$\frac{\Delta A}{A} \approx \frac{\Delta R/R + \Delta C/C}{\sqrt{2 + \Delta R/R + \Delta C/C}} \div \frac{1}{\sqrt{2}} \quad (5.18)$$

$$\approx \frac{\Delta R}{R} + \frac{\Delta C}{C} \quad (5.19)$$

例如, $\Delta R/R = 20\%$, 那么镜像抑制就只能达到 20dB。

图 5.27 中的两个相移级之间的增益失配也可以由频分引起。这是因为完全的镜像消除只在 $\omega_{IF} = 1/(RC)$ 时发生。例如, 如果信道带宽不比 ω_{IF} 小很多, 那么在近信道的边缘处 IRR 就会严重变差。

到底镜像抑制多少才合适呢? 在大多数 RF 应用中, 总的抑制必须在 60 到 70dB 左右。现在回想一下如果 IF 较高, 那么接收器前端滤波器也会在镜像频率处提供相当可观的衰减。因此, 在正确选择 IF 时若由式 5.19 表示的增益误差足够小, 那么一个镜像抑制结构就可以达到所要求的总的镜像抑制。但是要注意在这一拓扑结构中, 对匹配的要求仍然比在零差接收器中要严格得多。

Hartley 结构的单片实现也引起了其它问题。首先, 由于图 5.25 中的低通滤波器不能很容易地抑制相邻信道上的强干扰, 因此加法器的线性度就变得十分关键, 需要额外考虑噪声-功率间的权衡关系。其次, 相移 90 度级 (如图 5.24 中电路) 的损耗和噪声一般来说都十分严重。

Weaver 结构 在 Hartley 结构中, 我们注意到正交下变频后相移 90 度在两个路径中对有用信号产生相同的极性而对镜像信号产生相反的极性。如图 5.28 所示, Weaver 结构[11]把 90 度相移级代之以第二个正交混频操作, 起到了本质上相同的作用。不失一般性, 我们可以假设 $\omega_2 \ll \omega_1$ 。如图 5.29 所示, A 点的频谱和 $j[\delta(\omega + \omega_2) - \delta(\omega - \omega_2)]/2$ 进行卷积, 在 C 点产生了没有 j 因子的变换了的副本。同样, 在 B 点的频谱和 $[\delta(\omega + \omega_2) + \delta(\omega - \omega_2)]/2$ 进行卷积, 因此它从频率上被同时向下和向上变换了。把 D 点的频谱减去 C 点的频谱后, 我们注意到落在有用信号频带内的两个极性相反的镜像信号正好三相相消, 从而产生了没有镜像干扰的有用信号。由于下变频后的频谱仍然包含有在频率 $+\omega_2 + \omega_{IF}$ 和 $-\omega_2 - \omega_{IF}$ 处的镜像, 因此需要一个输出低通滤波器只把有用的频带选出来。

图 5.28 Weaver 镜像抑制接收器

图 5.28 中的两个 LO 频率 ω_1 和 ω_2 , 可以通过不同的方式选择。在图 5.29 中所示的情形中最终的频谱其中心并不是在零频率上, 因此第二次混频操作会引起“二次镜像”的问题。为了理解这个问题, 假设输入频谱包括一个在 $2\omega_2 - \omega_m + 2\omega_1$ 处的干扰(图 5.30)。在第一次下变频的时候干扰出现在 $2\omega_2 - \omega_m + \omega_1$, 也就是信号相对于 ω_2 的镜像。在第二次下变频的时候这个干扰不能被消除, 因为它本来就和有用信号一起处在 ω_{LO1} 的同一侧。因此, 图 5.28

中的低通滤波器事实上必须被代之以带通滤波器以抑制二次镜像。

LO 频率也可以选择为把频谱变换到基带，也就是 $\omega_1 \pm \omega_2 = \omega_m$ ，从而避免二次镜像的问题并减轻零差结构中的一些问题。特别是，由于 ω_1 可以离 ω_m 较远，第一个 LO 到主输入的泄漏就不严重，并且由自混合引起的直流分量可以由带通滤波器滤掉。这样的滤波器也会把信号通道中二阶失真产生的低频分量滤掉。而且，如果在中频处和第二组混频器处可以提供足够的增益，则基带级的闪烁噪声就可以被忽略。

图 5.29 Weaver 结构图解分析

图 5.30 Weaver 结构中的二次镜像问题

Hartley 和 Weaver 结构都有一个共同的问题：由于增益和相位失配使镜像不完全被抑制。Weaver 电路没有式 (5.19) 所描述过的失衡问题，但是如果第二下变频把频谱变换到一个非零频率，则它受二次镜像的影响。同时第二 LO 频率的谐波也可能把没有经过滤波的干扰从第一 IF 下变频到第二 IF 中。

5.2.4 数字中频接收器

在图 5.12 所示的双中频外差结构中，低频的操作如第二组混频和滤波，可以在数字域中更有效的实现。在图 5.31 所示的一个例子中，第一 IF 信号是数字化的，并与数字化的两正交正弦信号混频，然后经过低通滤波得到正交的基带信号。这种方法常称为“数字 IF 结构”。注意数字处理避免了 I 和 Q 的失配问题。

图 5.31 数字中频接收器

这种方法最根本的问题就是它对 ADC 转换器性能的要求。由于图 5.31 中 A 点的信号电平一般不高于几百微伏，因此 ADC 的量化和热噪声必须不超过几十微伏。此外，如果第一 IF 带通滤波器不能很好的抑制邻近的干扰，那么 ADC 的非线性度必须足够小，以使信号被交调的破坏减到最小。同时 ADC 的动态范围也必须足够大，以能允许信号由于路径损耗和多路衰减引起的变化⁷。此外，ADC 必须有与 IF 相当的输入带宽而只消耗合理的功率。

上面的要求使得在数字 IF 结构中使用一个 Nyquist 速率的 ADC 变得十分困难。典型的 50 到 200MHz 的 IF 值要求采样率在 100 到 400MHz 的范围内，而且线性度、底噪声以及动态范围的要求也会使分辨率必须大于 14 比特。这样的性能要求就算是在功耗和成本不是关键因素的情况下用今天的 ADC 变换器也达不到。

对 ADC 性能的限制可以被部分减低，只要注意到大多数的 ADC 使用采样保持电路，因此可以用来进行下变频。这一结构称为“采样 IF”结构，它表示在图 5.32 中，其中 ADC 的采样频率略低于 f_{IF} 。因此下变频后数字化的频谱就在 $f_{IF} - f_s$ 周围。这一操作之后是正交混频和滤波以把频谱变换到基带。

图 5.32 采样中频结构

这一技术虽然可以把 ADC 的采样速率降低一半，但它仍然同时要求有难以达到的高速度和线性度。尽管如此，它仍然被用在基站上，那里有许多的信道必须同时被接受和处理。

由于 ADC 性能的限制数字 IF 和采样 IF 结构还没有被用在移动终端上。但是它们的潜力已使得它们成为热点研究题目[12,14,13]。

5.2.5 亚采样接收器

到目前为止在我们对接收器的讨论中，都假设（第一）LO 频率是在 RF 频带附近。另外还有一种选择，就是 RF 输入可以以低得多的速率被采样，这是因为窄带信号从一个载波

⁷ 有些信号电平的变化可以通过采用模拟域的 AGC 来修正。

周期到另一个载波周期只有很小的变化。这一被称为“亚采样”的想法是这样的：如果一个带宽为 Δf 的带通信号以等于或者大于 $2\Delta f$ 的速率被采样，那么它就可以被变换到一个低频带。如图 5.33 中所示，对于理想采样，这样的操作产生了没有混叠的频谱副本。

图 5.33 (a) 时域和 (b) 频域中的亚采样

由于下变频频率的极大下降，使用亚采样可以简化本振以及它相应的频率综合环的设计。与一个 RF 混频器相比，采样电路本身也可以从降低了综合考虑的要求中获益。

尽管有这些特性，亚采样也有一个重要的缺点：噪声的混叠。为了理解这个问题考虑图 5.34(a) 中（开关打开时）采样器的等效电路。这个电路的带宽必须至少等于输入信号的频率 f_0 ，以使信号的衰减可以被忽略。这样，电阻 R_{on} 就会产生显著的噪声分量，其范围从直流一直到超过 f_0 ，如图 5.34 (b) 所示，为了简单起见这里画成矩形频谱。对这个频谱进行图 5.33(b) 的亚采样操作，我们得到如图所示的总效果。因此，用 m 倍进行的亚采样使采样电路下变频后的噪声功率增大了 $2m$ 倍。

图 5.34 (a) 采样器的等效电路，(b) 亚采样中的噪声混叠

亚采样也使采样时钟中噪声的影响更坏。可以证明时钟相位噪声功率被“放大”了 m^2 倍[15]。

5.3 发送器结构

一个 RF 发送器要完成调制，上变换和功率放大，有时前两个功能会合并在一起。与形式多样的为 RF 接收而发明的方法不同，发送器结构只有少数几种形式。这是因为对发送器中如的噪声、干扰抑制和频带选择等的要求要比在接收器中宽松的多。

在研究发送器的结构以前，我们先简单看一下在基带信号和 RF 部分之间的接口，以及功率放大器和天线之间的接口。

基带/RF 接口 图 5.35 所示的是一个模拟或者数字 FM 系统中的基带/RF 接口。这里，基带信号直接调制由 VCO 产生的载波频率。回忆一下第三章的内容，一个频率调制器的输出频谱取决于调制信号的幅度和带宽—结调制系数。因此，基带首先通过一个滤波器和/或者一个可变增益级“整形”，以补偿 VCO 特性由制造引起的偏移。同时由于振荡器的频率会随时间以及温度的变化而变化，所以必须用一个反馈环也就是一个频率综合器来稳定（第八章）。

图 5.35 FM 系统中的基带/射频界面

在数字相位调制系统中，基带信号和 RF 发送器之间的接口比较复杂。从第三章中我们注意到，简单的 QPSK 调制可以如图 5.36 描述的那样来实现。但实际上，数据脉冲必须被整形以使码间干扰最小并且/或者限制信号带宽（例如 GMSK）。由于模拟领域尤其在低频的脉冲整形要求体积很大的滤波器，每一个输入脉冲联合运用数字和模拟技术被变换到所需要的形状。图 5.37 所示的一个例子中[16, 17]，输入脉冲产生一系列的地址，如它可以驱动一个计数器，使从两个只读存储器（ROM）中产生出一组电平。这些电平被逐一变换成模拟形式，并在 A 和 B 点产生所需的脉冲波形。

图 5.36 正交调制系统中基带/射频界面

上面方法的一个特定的例子就是图 5.38 所示的 GMSK 电路。从第三章中我们知道一个 GMSK 信号可表示为 $x_{GMSK}(t) = A \cos[\omega_c t + \phi_k(t)]$ ，其中 $\phi_k(t) = \int \sum h(t) * p(t - kT) dt$ ， $h(t)$ 是高斯滤波器的冲击响应。由于 $h(t)$ 的精确定义（和截取）影响信道带宽（也就是相邻

信道功率), 滤波器的数字实现被证明比相应的模拟实现精确。滤波器产生的相位 ϕ_k 通过两个 ROM 被映射到 $\cos \phi_k$ 和 $\sin \phi_k$, 然后该结果根据 $x_{GMSK}(t) = A \cos \omega_c t \cos[\phi_k(t)] - A \sin \omega_c t \sin[\phi_k(t)]$ 被转换成模拟形式。

图 5.37 基带脉冲整形

图 5.36 和图 5.37 调制器的一个重要问题就是 I 和 Q 路径中相位和增益的失配。与在零差接收器中的 I/Q 失配效应类似, 这一缺陷导致由正交载波相位调制的两个数据流之间的串扰。一个常用的用来量化发送器中 I/Q 失配的方法是把两个信号 $V_0 \sin \omega_m t$ 和 $V_0 \cos \omega_m t$ 加载到 I 和 Q 的输入端然后检查加法器产生的频谱。在理想情况下, 有用频带内的输出就是 $v_{out}(t) = V_0 \sin \omega_m t \sin \omega_{LO} t + V_0 \cos \omega_m t \cos \omega_{LO} t = V_0 \cos(\omega_{LO} - \omega_m) t$ 。反之, 在有增益失配 ε 和相位失衡 θ 的时候,

$$v_{out}(t) = V_0 \sin \omega_m t \sin \omega_{LO} t + V_0(1 + \varepsilon) \cos \omega_m t \cos(\omega_{LO} t + \theta) \quad (5.20)$$

$$\begin{aligned} &\approx \frac{V_0}{2} [1 + (1 + \varepsilon) \cos \theta] \cos(\omega_m - \omega_{LO}) t \\ &\quad - \frac{V_0}{2} (1 + \varepsilon) \sin \theta \sin(\omega_{LO} - \omega_m) t \\ &\quad + \frac{V_0}{2} [-1 + (1 + \varepsilon) \cos \theta] \cos(\omega_m + \omega_{LO}) t \\ &\quad - \frac{V_0}{2} (1 + \varepsilon) \sin \theta \sin(\omega_{LO} + \omega_m) t \end{aligned} \quad (5.21)$$

这样在 $\omega_{LO} + \omega_m$ 边带处的功率除以在 $\omega_{LO} - \omega_m$ 边带处的功率就是

$$\frac{P_+}{P_-} = \frac{1 - (1 + \varepsilon) \cos \theta + \varepsilon}{1 + (1 + \varepsilon) \cos \theta + \varepsilon} \quad (5.22)$$

可以用来度量 I/Q 失衡。实际上, 如果上面的测试产生一个低于有用信号 30dB 左右的无用边带, 则这两个数据流之间的串扰可以被忽略。

图 5.38 GMSK 系统中的基带脉冲整形

PA/天线接口 如在 4.1 节中解释的那样, 发送器的输出必须通过一个双工滤波器或者一个 TDD 开关, 以使它能与天线接收到的信号相分离。双工滤波器一般会有 2 到 3dB 的损耗, 这样就会消耗掉 30% 到 50% 的 PA 输出功率并变成热。如果 PA 提供 1W 的功率, 那么就有大于 300mW 会被这个滤波器消耗掉。由于 PA 的效率一般小于 50%, 这一功率浪费相当于从电源吸收了 600mW 的功率, 这可与一个典型接收发送器中其余部分消耗的功率相比拟! 反之, TDD 开关的损耗在 0.5 到 1dB 之间, 这比 FDD 情形有明显更高的总体效率。

5.3.1 直接变换发送器

如果图 5.36 和 5.37 中的被发送的载波的频率与本地振荡器频率相等, 这种结构称为“直接变换”。在这一情形中, 调制和上变换在同一个电路中完成。如图 5.39 所示, 调制器后面是一个功率放大器和匹配网络⁸, 匹配网络的作用就是给天线传送最大的功率并把由放大器的非线性产生的带外分量过滤掉。注意由于基带信号是在发送器中产生的因此它足够强, 所以混频器中的噪声远没有接收器中严重。

图 5.39 直接变频发送器

图 5.39 的结构有一个重要的缺点: 由功率放大器引起的发送本地振荡器扰动。如图 5.40, 这个问题的引起是由于 PA 输出是一个调制波形, 它有很高的功率而且频谱的中心在 LO 频率附近。尽管可以用各种保护技术来隔离 VCO, 但是 PA 的“含噪”输出仍然能破坏振荡器的频谱。这种破坏影响是由一种称为“注入上拉”或者“输入锁定”的机制造成的, 将在第七章中描述。如果为了节能而周期性的开关 PA, 则问题更加严重。

图 5.40 PA 输出到振荡器的泄漏

如果 PA 输出频谱比振荡器的频率高出或者低出足够的值, 则 LO 上拉的现象可以被绕过。对于图 5.36 和 5.37 中的正交调制方法这可以通过给 LO 频率加上“偏移”来完成, 也就是加上或者减去另一个振荡器的频率[18]。图 5.41 是一个例子, 其中 VCO₁ 和 VCO₂ 的输出被混频, 然后被滤波使得载波频率是 $\omega_1 + \omega_2$, 离 ω_1 和 ω_2 都很远。

图 5.41 带有偏移 LO 信号的直接变频发送器

5.3.2 两步发送器

另一个解决发送器中 LO 上拉问题的方法是, 用两步 (或多步) 来上变换基带信号以使 PA 的输出频谱远离 VCO 的频率。作为一个例子, 考虑图 5.42 中的电路。这里基带的 I 和 Q 信道在较低的频率 ω_1 (称为中频) 外进行正交调制, 然后经过混频和带通滤波上变换到 $\omega_1 + \omega_2$ 。第一个 BPF 抑制了 IF 信号的谐波, 而第二个则滤掉了中心频率在 $\omega_1 - \omega_2$ 处的无用边带。

图 5.42 两步发送器

两步上变换与直接变换相比, 优点在于由于正交调制是在较低的频率完成的, I 和 Q 的匹配很好, 从而使两个比特流之间的串扰很少。同时, 一个信道滤波器可以用在第一个 IF 处以限制相邻信道中被发送的噪声和毛刺。

两步发送器中的困难在于第二个上变换后面的带通滤波器必须以一个很大的倍数, 一般是 50 到 60dB, 来抑制无用的边带。这是因为, 简单的上变换混频操作同时产生相同幅度的有用和无用的边带。由于较高的中心频率, 这一滤波器一般是无源的, 而且是相对昂贵的片外器件。

5.4 接收发送器的性能测试

今天的高性能的接收发送器必须通过 100 多个测试, 每个测试都把系统置于一个极端的条件下, 这些极端条件是接收发送器在实际环境中工作时可能遇到的。在这一节, 我们总结测试中常用到的几个重要的性能测试。

灵敏度和动态范围 在大多数系统中, 都会定义一个最小可检测信号电平 (在没有干扰的情况下)。例如 GSM 标准要求一个 -120dBm MDS, 只有 9 到 12dB 的 SNR。SNR 中这一 3dB 的变化量是考虑了各种类型的解调器和基带功能在性能上的差别, 但目标是达到 10^{-3} 的 BER。从 2.4 节可知

⁸ 可以在调制器和功率放大器之间插入一个带通滤波器来抑制谐波。

$$P_{MDS} = -174 \text{ dBm} + 10 \log B + NF + SNR \quad (5.23)$$

这说明最大的噪声系数是 7 到 10dB。

除了总的噪声系数和 IIP3 外，几个其它的接收发送器的特性也影响灵敏度和动态范围。一个通用的测试考察系统对阻塞信号的响应，这是通过测量带内交调、带外交调、二阶交调、交叉调制和互易混合来完成的。GSM 的前四个效应表示在图 5.43 中[19]，而最后一个将推迟到第七章中讲。注意在每一情形中都有有用信道中都有一个-98dBm 的信号。

在图 5.43(a)中一个-49dBm 的未经调制的谐波和一个-50dBm 的 GMSK 调制后的谐波被加在第四和第八个相邻的信道上。在接收器的输出端，有用信号（这里被简单的称为载波）被交调和噪声破坏。在这个测试中，输出载波比上噪声和交调的和[C/(N+I)]不能小于 9dB。

在图 5.43(b)中，两个电平分别为-45dBm 及-35dBm 的带外调制后谐波被相隔成它们的上交调积落在有用信道中而它们的二阶拍与 IF 频带相同。在这些条件下，IF 信号的 C/(N+I) 必须超过 9dB。

在图 5.43 (c) 中的交叉调制测试中，一个-23dBm 的带外幅度调制干扰被加在输入信号中。接收器的非线性必须足够低以使 C/(N+I)大于 9dB。

图 5.43 GSM 中的典型阻塞测试。(a) 带内交调，(b) 带外和二阶交调，(c) 带外互调

无用的发射

由发送器天线辐射的信号必须同时符合无线标准和 FCC 的严格规定。为了保证在相邻的信道中的辐射可以被忽略，每一个标准一般都会定义一个调制“障板”，发送器的输出频谱不能超出它。例如，GSM 就提供如图 5.44 所示的障板。

除了调制模版以外，无限标准一般还对调制后的信号说明相邻信道的功率（第四章）。IS-54 和 IS-95 标准分别要求 ACP 小于 -26dBc 和 -42dBc[1]。

图 5.44 GAM 中的调制障板及典型信号谱

发送器的输出也可能会包含谐波、毛刺和热噪声，它们有可能干扰其他标准的用户。谐波是由调制器和功率放大器产生的，而毛刺来自混频器、振荡器以及 PA 的寄生谐振和非线性。在 FDD 系统例如 AMPS 中，落入接收频带的输出毛刺十分棘手，因为前端双工器只能将这一分量衰减 50 到 60dB。类似的，输出的热噪声也很关键，因为它增大了接收路径中的接收底限[2]。很有意思的是注意到 NADC 和 GSM 系统都减轻了这两个问题，它们把发送和接收的时间段错开了，使发送器和接收器的操作不同步进行。

5.5 实例研究

为了进一步强调至今所描述的接收发送器的设计概念和问题，我们在这一节分析几个用在无线系统中的实际产品。以单片集成为重点，我们从简到繁的逐渐介绍更为复杂的结构。其它 RF 接收发送器的发展在[20-26]中有介绍。

5.5.1 Motorola 调频接收器

MC3362 是 Motorola 公司生产的一个单片双变换的 FM 接收器[27]。如图 5.45 所示，这是整个的接收器包括一些所需的外部部件。在如对讲电话或者第一代无绳电话这样的典型应用中，一个为天线接收到的 50MHz 的信号，在通过一个 LC 匹配电路后被加到第一个下变频混频器。注意为简单，这里既没有使用 LNA 也没有使用镜像抑制滤波器，但是匹配网络在一定程度上对镜像进行了抑制。混频器是为达到低噪声系数设计的，所以它由工作在比输入频率高或低 10.7MHz 的 LC VCO 驱动的。所产生的 IF 信号通过一个陶瓷的带通滤波器，以在一定的程度上抑制干扰，这样就可以采用噪声系数和线性度要求不高的第二个混频器。

第二个混频器是由 10.245MHz 的晶振驱动的，它产生等于 455kHz 的第二 IF。另一个陶瓷滤波器完成信道的选择。在这一点，FM 信号仍然很小，但是它可以被非线性放大，因为由于干扰产生的交调积可以忽略。这样，一个多级“限制”放大器（由六个级联差分对构成）把信号放大到使它能驱动后面的 FM 解调器。注意信号在解调以前放大要比在解调之后放大具有较高的功率效率，因为解调器的间频输出要求有线性和低噪声的增益级。

图 5.45 MC3362 的简化了的结构

图 5.45 中的接收器除了需要匹配网络和陶瓷滤波器以外，还需要其他的片外部件用于振荡器和解调器，从而使得片外部件的总数在 25 个左右。电路设计成工作在低至 2V 电源电压下，当输入 $0.7\mu\text{V}_{\text{rms}}$ 时 SNR 达到 20dB，电流一般在 5mA 左右。

5.5.2 Philips 传呼机接收器

Philips 的 UAA2080T 是一个单片双极型零差接收器，它是为 FSK 传呼应用而设计的 [28]。图 5.46 所示的是接收器和必要的外部部件。在前端，一个 LC 网络同时完成了匹配和单端至双端差分的转换，它的输出驱动一个低噪声共栅差分放大器。与 5.2.2 节中描述的零差结构不同，这个接收器通过两个 LC 网络把 RF 信号而不是 LO 信号分成同相和正交两个分量，正如下面要讨论的，这一方法的合理之处在于所得到的 LO 电路简单性。

信号在下变频到基带后，被放大、滤波并且交流耦合到信道选择滤波器。所产生的 FSK 信号的 I 和 Q 分量通过限制放大器，然后用来驱动 FSK 解调器。

LO 频率一般在 470MHz 左右，它是由一个 235MHz 的振荡器后面接一个倍频器产生的。振荡器实际工作在 78.3MHz 晶振的三次谐波（“overtone”）上。由于传呼信号是在单个固定频率上被接收，所以振荡器的频率无需变化（除了在制造时的微调），因此省去了频率综合器。但这个小巧的、低功耗实现不能很容易地产生精确的正交信号，所以需要在 RF 路径上对两个信号进行分离。

图 5.46 UAA2080T 的简化了的结构

针对 5.2.2 节中描述的零差设计问题考察 UAA2080T 是很有意义的。由于传呼 FSK 信号在零频率附近的能量可以忽略，所以通过交流耦合可消除下变频路径中的偏移。双极型工艺中的匹配特性最小化了 I/Q 失配，同时偶阶失真也通过差分电路被抑制。此外，LO 泄漏被 LNA 和混频器中的共栅结构减小了。有限动态范围的问题在这个应用中不是很严重，这两个原因：（1）FSK 信号使用了较高的调制指数，也就是它偏离载波频率一个很大的量，因此提供了较高的 SNR，但这是以频谱效率为代价的（第三章）；（2）比特误码率（BER）可以高达 3%，因为在数据流中利用的冗余来修正错误。

UA2080T 工作的最小电源电压是 2V，在有效工作时吸收大约 2mA 的电流。对于 BER=3%，数据率为 1.2kb/s，该电路要求输入信号电平大于 -125dBm ($0.126\mu\text{V}_{\text{rms}}$)，并能承受有用信号以上 60dB 的干扰。整个系统采用大约 30 个外部部件。

5.5.3 Philips DECT 接收发送器

Philips 提供一个芯片组可以用来构成一个 DECT 接收发送器。图 5.47 是整个 1.89GHz 的 TDD 系统，其中指明了每个芯片 [29]。在接收路径中，信号被转换成差分形式然后加在一个镜像抑制 LNA/混频器(UA2077)上。为了使镜像抑制达到大于典型的失配限制值 30dB，电路通过几个 DAC 转换器为 I 和 Q 路径中的相位和增益提供了数字控制。第一个 IF 信号为 110MHz，它通过一个 SAW 滤波器然后被下变频到 9.8MHz 的第二个 IF。接着它由两个陶瓷器件进行带通滤波，由一个限制级进行放大，再加到一个 FM 解调器上。注意这里第二 IF 要比图 5.45 中的高的多，因为 DECT 的信道带宽为 1.7MHz，所以需要足够高的中心频率。

图 5.47 Philips DECT 收发器

发送路径由一个 VCO，一个缓冲器和一个功放构成。在发送的时候，基带高斯形脉冲直接调制 VCO，产生高斯频移键控。系统实际的时序如图 5.48 所示，其中接收和发送模式由一个“盲槽”隔开。在接收模式结束的时候，发送 VCO 就被放在反馈回路（频率综合器）中用来稳定频率。由于回路需要大约 250us 才能稳定下来，所以盲槽被置于信号发送之前以避免频谱泄漏到邻近的信道中。在盲槽结束的时候，VCO 就从反馈回路中断开，于是它可以被调制，同时功放也开始工作。

图 5.48 说明盲槽操作的时序图

图 5.47 中发送器的根本的问题是由三个干扰引起的 VCO_{tx} 频率的误差。首先，当 VCO 脱离反馈回路的时候，它的控制线上会有一定的电荷注入误差。其次，当功放开始工作时，它的输入阻抗会发生变化，从而改变了 VCO 的负载阻抗，并且因此振荡频率也发生变化。第三，功放的有效工作电流大约 250mA，这可以使电池电压下降几百个毫伏，从而影响 VCO 的输出频率。这些误差的和必须不超过 50kHz（相当于在 1.9GHz 时 26ppm）。各种隔离和电源调节技术被用来减小这些效应减到最小[29]。实际上，系统时钟频率的时效问题则会把预算限制到 30kHz 以下。

5.5.4 Lucent（朗讯）GSM 接收发送器

Lucent Microelectronics 前身是 AT&T Microelectronics) 提供一个单片解决方案，它连同—个低噪声放大器和一个功率放大器可以组成一个完整的 GSM 接收发送器（直到基带接口）。图 5.49 描述了这整个系统。接收路径把 900MHz 的输入变换到 71MHz 的 IF，由一个 SAW 滤波器完成部分的信道选择，由一个可编程增益把信号放大，然后被下变频成正交的基带分量。与图 5.45 和图 5.47 中的接收器不同，这种结构只需要在接收路径中使用两个外部的滤波器（不包括双工器），但是，如果 IF SAW 器件必须把相邻信道过滤到足够低的电平，那么它往往会有较高的损耗（及较高的成本）。

发送路径把基带高斯形的数据直接上变换到 900MHz。为了避免 VCO 上拉，所需的载波信号是由 VCO₁ 和工作在 117MHz 的第二个振荡器 VCO₂ 的频率相加产生的（5.3.1 节）。调制器后面的缓冲器把 0dBm 的功率传输给 50 欧姆的负载。

图 5.49 Lucent Technologies 的 GSM 收发器

这个结构中使用的三个 VCO 是嵌入在综合器环路中的（第八章）。VCO₃ 所需的正交信号和发送的载波信号是使用有相当的工作频率和复杂度的不同的电路技术产生的。这些技术将会在第七章中描述。

用 12GHz 的双极型工艺制造的 GSM 芯片工作在 2.7 伏电源时吸收大约 60mA 的电流。

5.5.5 Philips GSM 收发器

Philips' semiconductor 提供一对 RF 和 IF 芯片用于 GSM 接收发送器。图 5.50 所示的是整个系统[30]。接收路径包括两个 LNA，以允许使用两个低成本的要求不高的镜像抑制滤波器。每一个 LNA 的增益都可以通过数字编程进行控制，覆盖范围从+21dB 到-38dB。放大后的信号通过与一个 1.3GHz 的 VCO 混频被变换到一个 400MHz 的 IF。当镜像信号在 1.7GHz 时，LNA 和混频器的输入级被设计成在镜像频率处的累积增益损失约小于 30dB，从而放宽了对滤波器截止带抑制的要求。IF 信号然后被滤波并下变频到基带的正交信道。由于大多数的信道选择都是在基带上通过五阶集成低通滤波器完成的，所以对 IF SAW 滤波器有比较宽松的要求。

图 5.50 Philips GSM 收发器

发送路径中使用两步上变换。第一步，高斯形的基带数据被调制到 400MHz 的载波上，然后由 LC 电路滤波。第二步，信号被分成正交的两相并加到由一个 1.3GHz 振荡器驱动的

单边带混频器上。这个混频器可以把无用边带衰减 20dB，从而放宽了对前级滤波器的抑制要求。然后 900MHz 的信号被缓冲后再加到功放上。

图 5.50 所示的结构只使用了两个振荡器来完成接收和发送路径中所有的频率转换，从而简化了对可能由耦合和交调引起的各种毛刺的预测。这种策略在这里之所以可行是因为系统时分时（而且频分）复用的，这使得这两个路径有可能共享振荡器。（回忆一下第四章，GSM 中发送和接收时间槽相隔 3 个时间槽。）

由于两个振荡器都是片外的，所以它们的输出泄漏到电路的其它部分就成为问题。特别是，如果 V_{LO2} 是 400MHz，那么，就像零差中的下变频一样，自混合就会由于 DC 偏移而破坏基带信号。因此， V_{LO2} 是在 800MHz 产生的，并且这个频率在 IF 芯片上被二分频。

图 5.50 接收发送器中有一个有趣的规定就是它和另一个标准 DCS1800[30]部分兼容。这个标准和 GSM 类似，但是工作在 1.8GHz 周围。 V_{LO1} 的频率被选择在 GSM 频带和 DCS1800 频带的中间，这允许它们两个都使用 400MHz IF，但是前者要求高边注入，后者要求低边注入。因此在原理上，这两个标准可以共享发送器中除 LNA、镜像抑制滤波器和功放以外的所有部分。

用 13GHz 的 BiCMOS 工艺制造的 GSM 接收发送器工作在 2.7V 电源电压的时候，发射模式下吸收的电流是 105mA，接受模式下吸收的电流是 50mA。

参考文献

- [1] J. S. Kenney and A. Leke, "Power Amplifier Spectral Regrowth for Digital Cellular and PCS Applications," *Microwave J.*, pp.74-92, October 1995.
- [2] B. Maoz and A. Adar, "GaAs IC Receivers for Wireless Communications," *Microwave J.*, pp. 102-114, January 1996.
- [3] R. C. Sagers, "Intercept Point and Undesired Responses," *IEEE Trans. Vehicular Tech.*, Vol. 32, pp. 121-133, February 1983.
- [4] J. F. Wilson et al., "A Single-Chip VHF and UHF Receiver for Radio Paging," *IEEE Journal of Solid-State Circuits*, Vol. 26, pp. 1994-1950, December 1981.
- [5] A. A. Abidi, "Direct-Conversion Radio Transceivers for Digital Communications," *IEEE Journal of Solid-State Circuits*, Vol. 30, pp. 1399-1410, December 1995.
- [6] B. Razavi, "Design Considerations for Direct-Conversion Receivers," *IEEE Trans. Circuits and Systems*, Part II, Vol. 44, pp. 482-435, June 1997.
- [7] K. Anvari, M. Kaube, and B. Hriskevich, "Performance of a Direct Conversion Receiver with $\pi/4$ -DQPSK Modulated Signal," *Proc. IEEE Vehicular Tech. Conf.*, pp. 822-827, 1991.
- [8] R. Hartley, "Modulation System," U.S. Patent 1,666,206, April 1928.
- [9] M. D. McDonald, "A 2.5 GHz BiCMOS Image-Reject Front End," *ISSCC Dig. Tech. Papers*, pp. 144-145, February 1993.
- [10] J. W. Archer, J. Granlund, and R. E. Mauzy, "A Broadband UHF Mixer Exhibiting High Image Rejection over a Multidecade Baseband Frequency Range," *IEEE Journal of Solid-State Circuits*, Vol. 16, pp. 385-391, August 1981.
- [11] D. K. Weaver, "A Third Method of Generation and Detection of Single-Sideband Signals," *Proc. IRE*, Vol., 44, pp.1703-1705, December 1956.
- [12] A. Hairapetian, "An 81 MHz IF Receiver in CMOS," *ISSCC Dig. Tech. Papers*, pp. 56-57, February 1996.
- [13] J. E. Eklund and R. Arvidsson, "A 10 b 120 MS/s Multiple Sampling Single Conversion

- CMOS A/D Converter for I/Q Demodulation,” ISSCC Dig. Tech. Papers, pp. 294-295, February 1996.
- [14] S. Jantzi, K. Martin, and A. Sedra, “A Quadrature Bandpass $\Sigma-\Delta$ Modulator for Digital Radio,” ISSCC Dig. Tech. Papers, pp. 126-127, February 1997.
- [15] D. H. Shen et al., “A 900MHz RF Front End with Integrated Discrete-Time Filters,” IEEE Journal of Solid-State Circuits, Vol. 31, pp. 1945-1954, December 1996.
- [16] K. Feher, Wireless Digital Communications, Englewood Cliffs, NJ: Prentice Hall, 1995.
- [17] R. Steele, ed., Mobile Radio Communications, Piscataway, NJ: IEEE Press, 1992.
- [18] T. D. Stetzler et al., “A 2.7-4.5 V Single Chip GSM Transceiver RF Integrated Circuit,” IEEE Journal of Solid-State Circuits, Vol. 30, pp. 1421-1429, December 1995.
- [19] J. Fenk, “RF Heterodyne Receiver Design Techniques,” MEAD Short Course on RF Integrated Circuit Design, April 1996.
- [20] J. Sevenhans et al., “An Integrated Si Bipolar RF Transceiver for a Zero IF 900 MHz GSM Digital Radio Front-End of a Hand Portable Phone,” Proc. CICC, pp. 7.7.1-7.7.4, May 1991.
- [21] V. Thomas, J. Fenk, and S. Beyer, “A One-Chip 2 GHz Single Superhet Receiver for 2 Mb/s FSK Radio Communications,” ISSCC Dig. Tech. Papers, pp. 42-43, February 1994.
- [22] J. Crols and M. S. J. Steyaert, “A Single-Chip 900 MHz CMOS Receiver Front-End with a High-Performance Low-IF Topology,” IEEE Journal of Solid-State Circuits, Vol. 30, pp. 1483-1492, December 1995.
- [23] T. Antes and C. Conkling, “RF Chip Set Fits Multimode Cellular/PCS Handsets,” Microwaves & RF, pp. 177-186, December 1996.
- [24] J. C. Rudell et al., “A 1.9 GHz Wideband IF Double Conversion CMOS Integrated Receiver for Cordless Telephone Applications,” ISSCC Dig. Tech. Papers, pp. 304-305, February 1997.
- [25] S. Heinen et al., “A 2.7 V 2.5 GHz Bipolar Chip Set for Digital Wireless Communications,” ISSCC Dig. Tech. Papers, pp. 306-307, February 1997.
- [26] R. G. Meyer, W. D. Mack, and J. E. M. Hageraats, “A 2.5 GHz BiCMOS Transceiver for Wireless LAN,” ISSCC Dig. Tech. Papers, pp. 310-311, February 1997.
- [27] Motorola Semiconductor, Motorola Communication Device Data, 1996.
- [28] Philips Semiconductor, Philips RF/Wireless Communications Data Handbook, 1996.
- [29] R. Mohindra, “Isolator for DECT Open-Loop Modulator,” RF Design, pp. 30-42, January 1996.
- [30] C. Marshall et al., “2.7 V GSM Transceiver ICs with On-Chip Filtering,” ISSCC Dig. Tech. Papers, pp. 148-149, February 1995.

第六章 低噪声放大器与混频器

在第四章和第五章中我们介绍了无线通信标准和无线收发机的结构，接下来的这章和后续章节我们将进一步深入到电路层次。从在接收路径中首先遇到的电路开始，我们介绍低噪声放大器和下变频器，重点放在可用 VLSI 工艺技术集成的方法上。在对电路的稳定性和输入匹配等一些一般性问题做了回顾之后，我们研究用双极型和 CMOS 实现的 LNA 设计，并且介绍一些最新的电路拓扑结构。接下来，我们讨论混频器：定义噪声因子和转换增益，描述用双极型和 CMOS 技术实现混频器的方法。最后，我们分析了混频器以及级联级中的噪声。

6.1 低噪声放大器

6.1.1 概述

除在第一章中提到的 RF 设计六边形外，一些其它因素也影响低噪声放大器的设计。表 6.1 列出了用于衡量在外差体系结构中的 LNA 性能指标的可接受的典型参数值。有必要来考虑每一个参数值以及该值对 LNA 总体性能的影响。

表 6.1 外差体系结构系统中的典型 LNA 参数

<i>NF</i>	2 dB
<i>IIP</i> ₃	-10 dBm
增益	15 dB
输入及输出阻抗	50 欧姆
输入及输出返回损耗	-15 dB
逆向隔离	20 dB
稳定性因子	> 1

因为 LNA 是接收路径中的第一个增益级，它的噪声因子直接被加到系统的噪声因子上去。和一个 NF 为 2dB 的双工机 (duplexer) 一起，如果忽略后续电

路产生的噪声，则最终系统的噪声因子是 4dB。因而，当带宽为 200KHz，所需信噪比（SNR）为 8dB 时，输入灵敏度大约为 -109dBm（见 2.4 节）。

为了更好地理解一个 2dB 的噪声因子会是什么样子的，如图 6.1 所示，假设我们用一个串联电阻来代表一个双极型晶体管以输入端为参考的噪声。也就是， $NF = 1 + R_{eq}/R_s$ （忽略器件的电容和其它电阻）。因为单位带宽中以输入端为参考的噪声，其噪声电压可用下式表示：

$$\overline{V_n^2} = 4kT \left(r_b + \frac{1}{2g_m} \right) \quad (6.1)$$

$$= 4kT \left(r_b + \frac{V_T}{2I_C} \right). \quad (6.2)$$

其中， I_C 为集电极电流， $R_{eq} = r_b + V_T/(2I_C)$ 。由于 $NF = 2dB$ ，因此 R_{eq} 一定不能大于 29Ω 。因而， Q_1 必须相对大些，而且其偏置电流要大。在实际情况中，由于负载电阻和随后各级也对噪声产生有贡献，要求 $r_b + V_T/(2I_C)$ 的值比 29Ω 还应低。

图 6.1 用一个等效电阻对双极型晶体管的噪声建模

由表 6.1 给出的 IIP_3 值为 -10dBm 和 $NF = 2dB$ ，可以导出一个真空动态范围。我们计算两个邻近干扰信号的最大值，它们的交调积低于噪声基底（floor）。由 2.4 节可知：

$$SFDR = \frac{2}{3}(IIP_3 - F) - 12dB \quad (6.3)$$

$$= 61dB. \quad (6.4)$$

外差结构的 LNA 的最小增益由三个参数决定：镜像抑制滤波器的损耗，以及混频器的噪声因子和 IIP_3 。对于滤波器的损耗为 4 到 5dB，混频器的噪声因子为 10dB，以及 IP_3 为 +5dBm，增益为 20 dB 的 LNA 可以有效地抑制输入端作参考的噪声，同时保持一个合理的等效 IIP_3 。在本章的最后我们将更严密地考查级

联级的性能。注意 LNA 必须在驱动 50Ω 的负载时提供这样大的增益。

在零差体系结构中，不需要用镜像抑制滤波器，因而将缓解 LNA 对增益和驱动的要求。但是，基带电路中的闪烁噪声问题却使得它的增益要在 RF 范围内最大化。

在天线和 LNA 之间的界面造成了区分模拟电路设计师和微波工程师的一个有趣的现象。视 LNA 为一电压放大器，我们会希望它的输入阻抗的理想值为无穷大。从噪声角度来分析，在 LNA 前面需要一个阻抗变换网络以获得最小的 NF。从信号功率的角度来分析，在天线和 LNA 之间需要利用共轭匹配。虽然每一种选择都有各自的优缺点，但是在现今的系统中最后一种占有优势。也就是说，LNA 的输入阻抗设计成 50Ω 电阻。这是因为接在天线后的带通滤波器（如双工器）常常要用于各种不同的无线收发系统，因此必须工作在标准的终端阻抗，其典型值为 50Ω 。如果从滤波器看对的源和负载阻抗远远偏离 50Ω ，这时滤波器的传输频带和抑止频带的特性将会呈现出严重的衰减和大幅度的波动。

输入匹配的品质由输入“返回损耗”来表示，定义为 $20\log|\Gamma|$ ，其中 Γ 是相对于源阻抗 R_0 的反射系数：

$$\Gamma = \frac{Z_{in} - R_0}{Z_{in} + R_0}. \quad (6.5)$$

简化之，令 $Z_{in} = R_0 + \Delta R$ ，得

$$\Gamma = \frac{\Delta R}{2R_0 + \Delta R}. \quad (6.6)$$

因而，在一个返回损耗为 -15 到 -20 dB 的 50Ω 系统中， $\Delta R \approx 15$ 到 9Ω 。事实上，工艺和温度的变化以及残余电抗项都要求 ΔR 局限在一个更小的范围。

对于外差结构，LNA 的输出阻抗也应该等于 50Ω ，使得驱动镜像抑制滤波器有最小的损耗和波动。类似的反向损耗的限制量也适用于此。

LNA 的逆向隔离决定了从混频器到天线的泄漏的 LO 信号的数量。该漏损起因于电容性的路径，衬底耦合和键合线耦合。在一个第一级 IF 值高的外差式接收器电路中，镜像抑制滤波器和前-后双工器在很大程度上抑制了漏损，这是因为 LO 的频率都落在了它们的抑止频带中了。另一方面，在零差拓扑结构中，

漏损主要是就被 LNA 的逆向特性减弱的。

除了上面这些参数外，LNA 的稳定性也是需要考虑的。由于输入和输出间存在反馈回路，在某些源和负载阻抗的组合下，电路会变得不稳定。由于双工器和镜像抑制滤波器的终端阻抗不能在 SPICE 模拟中被精确模拟，因此在正常情况下是稳定的 LNA 设计，会由于在生产过程中出现的极端值或工作在未预料到的高频或低频率，而产生振荡。

通常用一个常数来表征电路的稳定性，该常数称为 Stern 稳定性因子，定义为：

$$K = \frac{1 + |\Delta|^2 - |S_{11}|^2 - |S_{22}|^2}{2|S_{21}||S_{12}|}, \quad (6.7)$$

其中 $\Delta = S_{11}S_{22} - S_{12}S_{21}$ 。如果 $K > 1$ 且 $|\Delta| < 1$ ，则电路无条件稳定。也就是它在任何源和负载阻抗的组合下都不会发生振荡。使用 K 来描述的困难之处在于，要在很大频率范围内计算（或测量）电路的 S 参数，以确保在任何频率下 K 都大于 1。例如在使用 SPICE 模拟时，需要采用一些后处理来计算 K 值。

方程（6.7）表明当 S_{12} 减小的时候稳定性增加，也就是说，当电路的逆向隔离增加的时候稳定性增加。在传统的 RF 设计中，这是通过“中和”输入-输出的电容性路径实现的，如图 6.2(a)所示。即是 L_1 和 C_u 在应用频率下发生共振。但是在 IC 设计中，**浮空**的电感的寄生电容和耦合电容 C_1 馈入输入和输出结点。另外还有一种选择就是，通过共源共栅级联结构来抑制反馈（图 6.2(b)），但是噪声因子会稍稍有些增加。

图 6.2 采用（a）中和及（b）共源共栅级联来实现稳定性。

我们同时也应该注意到， K 是对稳定性的一种偏于保守的度量，因为它允许电路的源和负载阻抗任意变化。另一方面，如果 LNA 的负载阻抗已经调定得很好，例如，在零差或镜像抑制结构中，这时稳定性就很容易达到，只要确保输入阻抗的**实部**在任意频率下都为正即可。该判据为在第七章中提到的振荡器的单端口模型中被证实是正确的。

如其它的放大器一样，由于键合线电感引起的交流电源和地的迴路，可能会使得 LNA 变得不稳定。在 900MHz 和以上的频率范围，甚至一些纳亨的电感都会通过接地点在两级间产生耦合，从而引起振荡。正是因为这个原因，在设计中的预防措施，布线和精确的封装模拟是必要的。

最后要注意的是，我们应该注意到 LNA 中的低噪声要求通常决定了设计中拓扑结构的选择和参数值的使用。在下面我们将会看到，常常只允许一个晶体管（通常为输入器件）作为 NF 的主要决定源，因此在设计时就排除了采用诸如发射极或源跟随器和电阻反馈一类的结构的可能性。

6.1.2 输入匹配

为了说明建立一个 50Ω 的输入电阻的困难性，我们来研究一些电路结构。

考虑图 6.3 中描述的共源级。我们注意到在足够高的频率下， C_F 会引起 M_1 的栅极和漏极端的短路，这时就产生了一个电阻值大致等于 $R_L \parallel (1/g_m)$ 的输入电阻。这样，这个电路在原则上可以在工作频率下提供 50Ω 的输入电阻。具体说来，除 C_{GS} 之外的输入导纳的实部和虚部可表示为：

$$\Re\{Y_{in}\} = R_L C_F \omega^2 \frac{C_F + g_m R_L (C_L + C_F)}{R_L^2 (C_L + C_F)^2 \omega^2 + 1} \quad (6.8)$$

$$\Im\{Y_{in}\} = C_F \omega \frac{R_L^2 C_L (C_L + C_F) \omega^2 + 1 + g_m R_L}{R_L^2 (C_L + C_F)^2 \omega^2 + 1} \quad (6.9)$$

如果 $g_m R_L \gg 1$ ， $C_L \gg C_F$ ，且 $\omega \approx 1/(R_L C_L)$ ，则上面的表达式可以简化成：

$$\Re\{Y_{in}\} = \frac{g_m C_F}{2 C_L} \quad (6.10)$$

$$\Im\{Y_{in}\} = C_F \omega \left(1 + \frac{g_m R_L}{2} \right) \quad (6.11)$$

这表明只要适当地选择 $2C_L/(g_m C_F)$ ，就可以得到 50Ω 的输入电阻。这种拓扑结构的缺点是，由于在输出端带宽的限制，使得在高频下会产生相对较低的电压增益。

图 6.3 共源级

在图 6.4(a)中，一个 50Ω 的电阻并联在输入端，输入阻抗的电容部分被外部电感抵消了。但是，这个并联电阻也产生噪声。事实上，如图 6.4(b)所示且在第二章中已证实，由并联电阻 R_p 和源电阻 R_s 组成的电路级的噪声因子为：

$$NF = 1 + \frac{R_s}{R_p} \quad (6.12)$$

当 $R_p = R_s$ 时，LNA 的噪声因子超过 3dB。这里的关键点是电路应该有一个 50Ω 的输入电阻，但是没有 50Ω 电阻产生的热噪声。

图 6.4 (a) 电阻性终断，(b) 终断电阻的噪声因子计算

在图 6.5(a)中，负的并联反馈提供了一个低的输入阻抗，运用适当的设计可以使得其实部为 50Ω 。图 6.5(b)为这种拓扑结构的实现，其中 M_3 根据输出电压来反馈回一个电流信号。在这个例子中有双重困难。第一，反馈信号可以包含大量的噪声，因而使得噪声因子增大到不能接受的水平。第二，回路的整个相移对某些源和负载阻抗会引起电路的不稳定。[注意图 6.5(b)的 LNA 和三级环形振荡器的相似之处。]

图 6.5 (a) 通过反馈实现的电阻性终断，(b) (a) 的简单实现

图 6.6(a)为共栅级设计以获得 50Ω 的输入阻抗，也就是说， $1/(g_m + g_{mb}) = 50\Omega$ 。通过外部电感补偿输入电容，使之等于零。这种方法主要的缺点是输入晶体管的跨导不能任意的大，从而使得噪声因子有一个下限。只考虑 M_1 的沟道热噪声， $I_n^2 = 4kT\gamma g_m$ (其中对于长沟道器件 $\gamma = 2/3$)，然后使用图 6.6(b)和图 6.6(c)的等效电路，读者可以容易地证明得到噪声因子为：

$$NF = 1 + \gamma \quad (6.13)$$

如果 $\gamma = 2/3$ ，由式 (6.13) 可以得出 $NF = 5/3 = 2.2\text{dB}$ 。同样，如果采用共基结构提供电阻值等于 R_s 的输入阻抗，只考虑集电极的散粒噪声，这时 $I_n^2 = 4kTg_m/2$ ，因此， $NF = 3/2 = 1.77\text{dB}$ 。

图 6.6 (a) 共栅级，(b) 包括沟道噪声在电路中，(c) (b) 的等效电路。

虽然通过上面的计算表明可以在输入匹配的共栅和共基拓扑结构中得到可以接受的噪声因子，但是实际上其它的噪声源大大地降低了电路的性能。在短沟 MOSFET 中，因子 γ 比 $2/3$ 要大得多 [1]。而且在双极型晶体管中，基极电阻的热噪声会比集电极的散粒噪声要大。此外，在这两种情况下，偏置电流源，以及负载也可能，会产生额外的噪声。

图 6.7 给出了另外一种得到 50Ω 输入电阻的方法。忽略栅-漏和源-衬体电容，我们可以得到：

$$Z_m \approx \frac{g_m L_1}{C_{GS}} + L_1 s + \frac{1}{C_{GS} s} \quad (6.14)$$

这样，适当选择的 g_m ， L_1 和 C_{GS} ，可以产生一个 50Ω 的实部。实际上，上式的最后两项在工作频率下可能不会发生共振，这就被迫在输入端使用芯片外的元件。高频时，要求的 L_1 值变得可以与接地焊接引线的电感值相似，这就需要多根（并联）的焊接引线或精确的引线电感模型。同样，由于兼并引起等效跨导的降低，这样可能会增大与 M_1 漏极相连器件对噪声的贡献。

图 6.7 用电感型兼并电路实现的电阻终断

6.1.3 双极型 LNA

我们首先研究一个简单的共发射极级来理解这种结构的局限性。如图 6.8(a) 所示的电路， Q_2 和 I_1 决定了 Q_1 的偏置电流， R_1 将信号路径从 Q_2 的噪声隔离开来，

R_2 维持与由于 Q_1 的不可忽略的基极电流在 R_1 上产生的电压降的同样压降。如果 R_1 远远大于 R_S ，则偏置电路对 LNA 性能的影响可以忽略不计。

图 6.8 (a) 简单的双极型 LNA, (b) 包括基极散粒噪声

由方程 (6.2) 可知, Q_1 必须是一个尺寸大的器件且被相对较大的电流偏置。但是, 增加器件的尺寸, 就会导致输入电容 (C_{je} 和 C_{jc} 的密勒倍率增加) 变大, 其后果是减弱输入信号和“放大” Q_1 和 R_C 的噪声。此外, 更大的集电极-基极电容和集电极-衬底电容会降低电压增益, 同时增大偏置电流会导致基极-发射极扩散电容 ($\approx g_m \tau_F$) 的增大和基极散粒噪声的增加。正是因为这些原因, 噪声因子会在 Q_1 特定的尺寸和偏置电流下, 取得最小值。

下一步是提高方程 (6.2) 式的精确度, 我们用一个电流源来将基极散粒噪声包括在等效电路中 [如图 6.8(b)]:

$$\overline{I_n^2} = 4kT \frac{I_C / \beta}{2V_T} \quad (6.15)$$

对于源电阻 R_S , 整个与输入相关的噪声电压包括 R_S 引起的, 由下式表示:

$$\overline{V_{tot}^2} = 4kT \left(R_S + r_b + \frac{1}{2g_m} + \frac{g_m R_S^2}{2\beta} \right) \quad (6.16)$$

上式忽略了集电极散粒噪声和基极散粒噪声之间的相关性。因此噪声因子等于:

$$NF = \frac{\overline{V_{tot}^2}}{4kTR_S} \quad (6.17)$$

$$= 1 + \frac{r_b}{R_S} + \frac{1}{2g_m R_S} + \frac{g_m R_S}{2\beta} \quad (6.18)$$

当 $R_{S,opt} = \sqrt{\beta(1+2g_m r_b)} / g_m$ 时, 噪声因子达到最小值:

$NF_{min} = 1 + \sqrt{(1+2g_m r_b) / \beta}$ 。这个方程没有考虑寄生电容的影响。但是, 通过一个

合理的高频近似表达式可以简单地估算出在工作频率下 β 的大小, 这个表达式就是 $|\beta| \approx f_T/f$ [2]。NF 的精确值应该由模拟最终得到。

上面得到有关 $R_{S,opt}$ 的表达式表明将天线的输出阻抗用位于天线和 LNA 之间的匹配网络转换为 $R_{S,opt}$, 可以获得最小的噪声因子。但是, 正如前面所提到的, 共轭匹配比“噪声匹配”的应用更为广泛。

来计算图 6.8(a)所示 CE (共发) 级的输入 IP_3 是有启发性的。 Q_1 的集电极电流可以用下式表示:

$$I_C \approx I_S \exp \frac{V_{BE0} + V_{in}}{V_T} \quad (6.19)$$

$$= I_S \exp \frac{V_{BE0}}{V_T} \left[1 + \frac{V_{in}}{V_T} + \frac{1}{2} \left(\frac{V_{in}}{V_T} \right)^2 + \frac{1}{6} \left(\frac{V_{in}}{V_T} \right)^3 + \dots \right] \quad (6.20)$$

比较 (6.20) 式和 (2.11) 式, 我们可以得到 $\alpha_1 = (1/V_T) \exp(V_{BE0}/V_T)$ 和 $\alpha_3 = [1/(6V_T^3)] \exp(V_{BE0}/V_T)$, 因此, 通过 (2.28) 式可以得到, $IP_3 = 2\sqrt{2}V_T$ 。该电压在 50Ω 电阻上对应的功率大小为 -12.7dBm 。因而, 通常需要提高额外的线性度来使得 IP_3 的值大于 -10dBm 。

共发射极级的小信号增益同时依赖于 Q_1 的寄生参数和负载阻抗。如果 LNA 要驱动一个 50Ω 的负载并维持相当的增益, 这时就可能需要第二级。这样, 第一级必须驱动第二级的输入电容和提供足够的增益以减小后级对噪声的贡献。这两级通常的设计成为一个整体以使得总的噪声因子最小化及输入 IP_3 和增益最大化。

上面我们已经研究了共发射极级的局限性, 现在我们考虑一些最新的双极型 LNA。

图 6.9(a)给出了一个 900-MHz LNA 的简化电路, 它包括两个共发射极级, 其中一个有电感兼并支路, 另外一个有电阻性反馈支路[2]。用焊接引线实现并提供第一级与地之间的连接, 电感 L_e 在电路中起到双重作用: 它不仅为输入提供了共轭匹配, 同时还对电路起到了线性化的作用。忽略 C_u 和 r_x 的影响, 我们

得到输入阻抗的表达式如下：

$$Z_{in} = r_b + \frac{g_m L_e}{C_\pi} + L_e s + \frac{1}{C_\pi s} \quad (6.21)$$

因而，适当选择 g_m ， L_e 和 C_π ，我们可以使得 $r_b + g_m L_e / C_\pi = 50\Omega$ ，同时式 (6.21) 中的最后两项互相抵消。

第二级中的反馈支路同时提供使电路线性化和降低输出阻抗的作用，但是以更高的噪声为代价。两级之间的界面是一个普遍问题的例子：适当选择第一级的增益和第二级的噪声和非线性问题。我们注意到，第一级的增益要足够高以减小第二级对噪声的贡献，但是不要高到让 Q_2 的非线性来决定整体 IP_3 的程度。因而，完整的设计要通过反复调试以得到一个可接受的折衷方案。

图 6.9 (a) [2]中的二级 LNA，(b) LNA 的偏置电路

图 6.9(a)所示的 LNA 的直流偏置必须保证一个恒定增益来应对温度和电源电压变化的影响。图 6.9(b)以简化的形式描述了偏置电路[2]。放大器 A_1 和电阻 R_{B1} 提供了 Q_1 的偏置电流以使得 $V_X \approx 1.2\text{ V}$ 。另外，利用能带隙技术产生一个与绝对温度成正比的电压 V_{PTAT} 。这个电压再被加上 V_{BE} 的整数倍，它控制 Q_3 和 Q_4 的基极，使得它们的小信号增益与温度和电源电压无关。例如，因为 $V_{b1} = V_{PTAT} + 2V_{BE}$ 时， Q_1 的集电极电流可以表示成 $I_C \approx (V_{PTAT} + 2V_{BE} - V_{BE3} - V_{BE2}) / R_1 = V_{PTAT} / R_1$ ，其中在 R_E 上的电压降被忽略了。因而，第一级的直流增益等于 $A_{v1} = g_{m1} R_1 = I_{C1} R_1 / V_T = V_{PTAT} / V_T$ 。另一方面，在 900MHz 时， L_e 的阻抗是相当大的，导致一些项不能完全被抵消掉，并出现某些温度依赖关系。

采用 13-GHz BiCMOS 技术，采用 5-V 的电源电压来实现图 6.9 的电路，可以得到噪声因子为 2.2dB，输入 IP_3 为 -10dBm，在 900MHz 时增益为 16dB[2]。

另一个被设计来驱动 50Ω 负载的双极型 LNA 在图 6.10 中给出。此 LNA 使用负反馈通过单片集成变压器来使电路线性化，当电源电压小到等于 V_{BE} 时仍能

工作。有趣的是，变压器同时减小了在低频和高频时放大器的增益，有利于稳定电路。外部电感 L_1 和电容 C_1 在输入端提供了共轭匹配。

图 6.10 包括变压器反馈的双极型 LNA。

采用 11-GHz BiCMOS 工艺实现的图 6.10 的电路从 1.9-V 的电源电压中耗取 2mA 电流，其噪声因子为 2.8dB，1.9GHz 时增益为 9.6dB。变压器反馈使得输入 IP_3 提高到 -3dBm。

另外一种可能的双极型 LNA 拓扑结构是共基极结构（图 6.11）。该电路相对于共发射极电路具有三个优点：更为简单的输入匹配，更高的线性度和更大的逆向隔离。忽略发射极和基极电阻，则输入阻抗为 $(g_m + C_\pi s)^{-1}$ ，当 $I_C \approx 0.5 \text{ mA}$ 时，得到输入阻抗的电阻部分为 50Ω 。 C_π 的影响要么可以忽略不计，要么就被外部电感补偿掉了。

图 6.11 共基结构 LNA

在共基极电路中，源电阻 R_s 线性化电路的输入-输出特性是通过“软化”发射极电流的漂移实现的。这与共发射极电路形成对比，在共发射极电路中，源电阻仅仅反映基极电流的变化。当 $R_s = g_m^{-1} (= 50\Omega)$ ，对于图 6.11 的电路，通过模拟可以预知输入 IP_3 近似为 -6.8dBm。如果基极偏置被适当地交流短接，则这个拓扑结构也表现出高的逆向隔离，该作用被证实在零差体系结构中是很有用的。如 6.1.2 节提到的，共基极电路最主要的缺点是噪声因子相对较大。

6.1.3 CMOS LNA

虽然 MOSFET 在大约十年前还被认为是速度慢，噪声大的器件，但是等比例缩小规则已经大大地提高了它们的性能。不象双极型晶体管，其基极电阻是影

响噪声因子的主导因素，MOSFET 只有一个主要的噪声源，该噪声是由沟道产生的。因此，在亚微米技术中，合理的器件和偏置电流的组合将会提供令人满意的低噪声。

在研究 CMOS LNA 之前，我们应该提到 MOS 器件的分布式栅电阻也会产生热噪声。对于总的栅电阻为 R_g ，可以通过等效集中参数模型得出一个集中的噪声源为 $\overline{v_{ng}^2} = 4kT R_g / 3$ [4]。不过，如果通过布线将晶体管变成很多狭窄器件的并联组合，则栅电阻会减小。

基本的共源和共栅 LNA 结构在性质和局限性上与共发射极和共基极电路有相似之处。但是有一点除外，对在 RF 应用中采用的典型的偏置电流和尺寸，MOS 器件比双极型器件更为线性。

为了得到 CMOS LNA 拓扑结构，我们考虑如图 6.12(a)所示的共源极，其中 C_1 表示 M_1 的寄生电容和下一级的输入电容。由于 MOSFET 的跨导较小，该电路的电压增益相对较小，这样由 R_D 和下一级贡献的噪声是相当重要的。为了提高电压增益，用电流源代替 R_D ，如图 6.12(b)所示。如果 M_2 的输出阻抗足够大，电路的增益大概等于 $g_{m1}/(C_1s)$ 。因为电流源自身会产生噪声，我们可以推想如果 M_2 也能放大信号[图 6.12(c)][5]，则总的以输入为参考的噪声会因此减小。由于 M_2 的偏置电流和漏极电压强烈依赖于电源电压，一个直流反馈网络可用来决定静态工作点[图 6.12(c)]。这里的关键想法是重复使用偏置电流来提供了一个更大的等效跨导： $g_{m1} + g_{m2}$ 。但是这个电路的输入电阻相对较大。

图 6.12 (a) 带有电阻性负载的共源级，(b) 将负载变换成电流源，(c) 利用负载来作为能放大的器件，(d) 文献[5]中的 LNA。

图 6.12(d)的电路被一个相似的电路级所跟随，以便驱动 50Ω 的负载。注意共源级比源跟随器具有更高的驱动能力，这是因为前者的增益等于 $g_m R_L$ ，而后的增益为 $g_m R_L / (1 + g_m R_L)$ [5]。当 $g_m \approx 1/R_L$ 时，共源级的电压增益为单位 1，

而源跟随器却以因子 2 而衰减。

采用 $0.5\text{-}\mu\text{m}$ CMOS 工艺, 电源电压为 2.7-V , 构成如图 6.12(d)所示的 900-MHz LNA, 得到最小噪声因子为 1.9dB (带外部匹配网络时), 增益为 15.6dB , 输入 IP_3 为 -3.2dBm 。消耗功率为 20mW 。

图 6.13 所示的是为采用片上和片外电感的 1.5-GHz CMOS LNA 电路。采用如图 6.7 相似的方式, 该 LNA 结合 L_s 和 L_l 来取得输入端的共轭匹配网。当频率为 1.5GHz 时, 片上电感 L_D 在它的 Q 值小于 4 时仍能提供高的电压增益。与此成对比的是, 采用负载电阻需要一个大的电压降才能提供可与之相比较的增益。

图 6.13 文献[6]中的 LNA

共栅晶体管 M_2 通过增加 LNA 的逆向隔离, 在电路中起了两个重要的作用: (1)它降低了由后级混频器产生的 LO 漏损, (2)通过减小输出到输入的反馈提高了电路的稳定性。注意相同的电路如果没有共源共栅级, 则易于发生振荡。

图 6.13 的 LNA 后接另外一个共源共栅级以驱动 50Ω 负载, 每一级汲取电流 10mA 。采用 $0.6\text{-}\mu\text{m}$ CMOS 工艺, 电源电压为 1.5-V , 电路的噪声因子为 3dB , 增益为 20dB , 输入 IP_3 为 -10dBm 。

图 6.14 所示的 LNA 拓扑结构以前是用 MESFET 实现[7], 现在用 COMS 器件实现[8]。在该电路中, M_1 和 M_2 的工作状态都为共源级形式, 但是它们共用同一个偏置电流。由 M_1 放大的信号通过 C_1 与 M_2 的栅极相连, 而 M_2 的源极另一通过 C_2 交流接地。因而, 这个电路通过偏置电流的复用来节省了功率。在实际实现中, 寄生的 C_1 的底部平板电容会限制结点 X 或 Y 处的 RF 增益。

图 6.14 文献[7]中的 LNA

6.2 下变频混频器

6.2.1 概述

混频器通过两个信号（包括也可能是它们的谐波）相乘进行频率变换。在接收路径上的下变频器有两个区分得很清楚的输入端口，称为 RF 端口和 LO 端口。RF 端口接收将要进行变频的信号，LO 端口接收由本地振荡器产生的周期性波形。如图 6.15(a)所示的简单电路，当开关 S_1 合上时，输出等于 RF 端口的输入信号；当开关 S_1 打开时，输出为零。如第 2 章中提到的，该操作也可以被看成是 RF 信号与矩形方波信号相乘。注意对于 RF 端口来说该电路为线性时变系统，而对 LO 端口而言为非线性时变系统。这对大多数下变频器来说都是正确的。下面我们将看到，总是希望开关的开与关尽可能地陡变。

图 6.15 (a) 简单开关被用作混频器，(b) 用 NMOS 实现的开关

被 LNA 放大的信号（也可能是经镜像抑制滤波器滤波后的信号）被输入到混频器的 RF 端口。因而，该端口应具有足够低的噪声和足够高的线性度；要有高线性度的原因是由于邻近干扰信号也被 LNA 放大，因此会出现更强的交调乘积项。如果图 6.15(a)用一个 MOS 开关[如图 6.15(b)所示]，那么晶体管的导通电阻会产生噪声。此外，当 RF 输入信号变化时 M_1 的栅-源过驱动电压也会变化，引起导通电阻的变化。这在 M_1 和 R_L 之间引入电压分压非线性性。

表 6.2 中列出了典型下变频器的性能参数。如果混频器的噪声因子为 12dB， IIP_3 为 +5dBm 时，混频器的前面可以接一个增益约为 15dB 的 LNA，整个电路的仍然具有可以接受的噪声和非线性指标（第 2 章）。

表 6.2 典型混频器的指标

NF	12 dB
IIP3	+5 dBm
增益	10 dB

输入阻抗（外差）	50 Ω
端口到端口隔离	10-20 dB

无源和有源混频器 图 6.15(b)所示的电路是“无源”混频器的一个例子，因为它不提供任何增益。电路的损耗可以通过如下方法得到：如果 LO 信号的占空比为 50%，那么 RF 输入信号被乘以一个在 0 和 1 之间转换的无量纲的正方形波形。这种波形的基频为 ω_{LO} ，振幅等于 $\sin(\pi/2)/\pi = 1/\pi$ 。因为 IF 信号是输入频谱与这个基频的卷积，电路的电压增益等于 $1/\pi$ ¹。

与此不同，有源混频器一般能提供增益。例如，如图 6.16 所示，RF 输入信号改变 M_1 的漏极电流，而 M_2 和 M_3 作为开关对由 LO 信号驱动。因而，实质上是 M_1 的漏极电流与正方形波形相乘，就象 M_1 的漏极电流被交替地传过 R_1 和 R_2 上。嗣后，我们会详细地研究这个电路的 MOS 和双极型实现，但是我们可以注意到适当选择器件的尺寸和偏置电流能够使这个电路提供相当大的电压增益。

图 6.16 有源混频器

由于它们的增益，有源混频器减少了后级对噪声的贡献，因而被广泛地应用在 RF 系统中。另一方面，无源混频器通常具有高线性度和高速的特征，被应用于微波和基站电路中。

转换增益 混频器的增益必须仔细地定义，以免发生混淆。一个混频器的“电压转换增益”定义为 IF 信号的均方根电压与 RF 信号的均方根电压的比值。注意这两个信号的中心频率为两个不同的频率值。通过输入一个频率为 ω_{RF} 的正弦波并测量频率为 ω_{IF} 的下变频信号分量幅度可以得到电压转换增益。

混频器的“功率转换增益”定义为传送到负载的 IF 功率除以从信号源可以

¹ 这个计算假设开关的导通电阻比负载的阻抗小得多。

得到的（最大）RF 功率。例如，如果混频器的输入阻抗和负载阻抗都等于源阻抗，比如说 50Ω 。这时，如果以分贝来表示，电压转换增益和功率转换增益相等。

在采用了镜像抑制滤波器的外差式接收器中，第一个下变频级中必须有对混频器的输入端的共轭匹配。如前所述，这些滤波器的传输函数通常只在一个标准的端口阻抗条件下被校准的。如果采用别的阻抗，可能会出现波动。另一方面，由于大多数无源 IF 滤波器的输入阻抗为 500 到 $1000\Omega^2$ ，使得大多数混频器的负载不可能等于 50Ω 。在一些如零差拓扑的体系结构中，出于取得最大电压增益的考虑，混频器看到的负载甚至可能更高。

从上面的分析中，我们可以发现混频器的电压和功率转换增益如果用分贝来表示并不一定相等。虽然这个不一致性本身并不是个什么问题，但是这个差异会在计算整个噪声因子，线性度，和采用混频器的级联级的增益上产生混淆。在 6.3 节中我们将返回来继续讨论这个问题。

SSB 和 DSB 噪声因子 混频器的噪声因子通常是产生大的混淆的来源。我们简单地只考虑一个增益为 1 的无噪声混频器。如图 6.17 所示，进入 RF 端口的频谱包含信号部分和进入信号频带和镜像频带中由 R_s 产生的热噪声。通过下变频，信号，信号频带中的噪声，镜像频带中的噪声都被转换成 ω_{IF} 。因而，如果混频器的输入频率响应在信号频带和镜像频带中是相同的，那么输出的 SNR 等于输入的 SNR 的一半。换言之，一个无噪声混频器的噪声因子等于 3dB。

图 6.17 RF 与镜像噪声被折到中频频带

以上的测量提供的是混频器的“单边带”噪声因子（SSB NF）。SSB 这个词表示有用信号频谱只位于 LO 频率的一边，这在外差式系统中是个普遍存在的情况。现在，考虑 AM 信号通过单边无噪声混频器进行零差下变频（图 6.18）。在这种情况下，输入和输出的信噪比是相等的，因此噪声因子为 0dB。这称为“双

² 除非当 IF 特别高时，例如达到 100MHz 以上。

边带”噪声因子 (DSB NF), 以便强调输入信号频谱位于 ω_{LO} 的两边。

图 6.18 一个 AM 信号的下变频

总的说来, 如果在混频器的 RF 端口, 信号频带和镜象频带具有相同的增益, 则混频器的 SSB 噪声因子比 DSB 噪声因子大 3dB。典型的噪声因子测试计只测 DSB NF, 测试计给出的 SSB NF 值就是简单地在 DSB NF 的值上加 3dB 得到。

在讨论混频器的噪声因子时, 我们隐含地假定 RF 输入信号是与正弦波的 LO 信号相乘。如前所述, 混频器中的开关转换操作可以看成是输入信号与矩形波的相乘。从而, 如果本地振荡器的输出占空比为 50%, RF 的输入信号是被 LO 信号所有的奇次谐波相乘。这就表示 RF 在 $3\omega_{RF}$, $5\omega_{RF}$ 等附近的噪声成分, 都下变频到 IF 频带, 进一步增加了输出噪声。虽然如此, 这个效应通常可以忽略, 这是因为(1)混频器的有限输入带宽降低了在更高频率下的转换增益; (2)矩形波的高次谐波的幅度很小。

端口间的隔离 在混频器中每两个端口间的隔离是很重要的。LO-RF 馈送导致了 LNA, 最终是天线中的 LO 漏损, 而 RF-LO 馈送使得 RF 路径中的强干扰与驱动混频器的本地振荡器互相作用。LO-IF 馈送是重要的, 因为如果大量的 LO 信号即使经过低通滤波器后依然存在于 IF 输出信号中, 会使得下一级电路变得不敏感。最后, RF-IF 隔离决定在 RF 路径中多少信号直接出现在 IF 中, 这是一个在零差接收器中的有关偶阶失真关键问题(第 5 章)。

所要求的隔离程度在很大程度上取决于混频器被使用的环境。假若混频器提供的隔离度不够, 其之前或之后的电路可能需要被修改以解决这个问题。

单均衡和双均衡混频器 如果混频器提供差分 LO 信号和单端 RF 信号, 则该混频器称为“单均衡”, 一个例子是如图 6.19(a)所示的拓扑结构。如果混频器的 LO 和 RF 输入都是差分方式, 则称为“双均衡”, 其有源结构实现用了 Gilbert 单元的形式[图 6.19(b)]。

图 6.19 (a) 单均衡混频器, (b) 双均衡混频器

在指定功耗下, 单均衡结构表现出的以输入端为参考的噪声比相应的双均衡结构要低。但是, 正如在 6.2.4 节中解释的, 这个电路更易受到在 LO 信号中的噪声的影响。双均衡产生的偶阶失真小, 因而缓解了外差式接收器的半中频问题, 并降低了零差结构体系中的差拍成分。不过, 经 LNA (可能也包括镜像抑制滤波器) 处理过的 RF 信号通常是单端输出, 因此双均衡混频器输入端口的一端就简单地接到偏置电压。这种接法反过来又导致两个通过在图 6.19(b)中的 M_1 和 M_2 放大的信号相位有不同的传播时间——即相移——, 其后果是偶阶失真不能忽略。

图 6.19(a)所示的单均衡混频器有一个缺点就是 LO-IF 馈送。注意 M_2 和 M_3 形成差分对用来放大 LO 信号。如果 IF 频率不比 LO 频率小多少, 这时接着混频器后的一阶低通滤波器不能在不衰减 IF 信号的情况下, 充分地抑止 LO 馈送。这样, 大的 LO 成分可能会使得 IF 放大器失去灵敏度。另一方面, 在图 6.19(b)所示的双均衡混频器中, 差分对 M_3-M_4 和 M_5-M_6 在放大的 LO 信号中增加了相反的相位, 因此提供了一阶抵消。一般而言, LO-IF 馈送的问题要和混频器后面级的非线性性一起仔细考虑。

在图 6.19 中的所有两个拓扑结构中, 输出可以是差分信号也可是单端信号。除了高转换增益外, 差分输出还为 RF 信号到 IF 输出的馈送提供了更大的抗扰度。为了了解这种作用, 考虑如图 6.20 所示的简化了的单均衡混频器。对两电路的操作都可以看成是 I_{RF} 与正方形波的乘积。不同之处是在图 6.20(a)中正方形波是在 0 和+1 之间变换, 而图 6.20(b)中的正方形波是在-1 和+1 之间变换。这样, 两个混频器的输出电压都可以表示成 $V_{out}(t) = I_{RF}R[a + bS(t)]$, 在图 6.20(a)中 $a = b = 0.5$, 图 6.20(b)中 $a = 0$, $b = 1$, $S(t)$ 为图 6.20(b)中画出的正方形波。从这个方程中可以发现, 单端输出包含一个等于 $0.5I_{RF}R$ 的成分。我们称这种作用为“直接馈送”。

图 6.20 (a) 单端输出及 (b) 差分输出的简单混频器

I_{RF} 高频成分的直接馈送一般都不重要，这是因为在 IF 输出端的低通滤波器抑止了这些成分。另一方面， I_{RF} 的低频成分毫不衰减地出现在输出信号中。回想在第 5 章中讨论的零差结构，在 RF 路径（如 LNA）中的偶阶失真引起了低频差拍或解调信号。这些信号直接出现在混频器的单端输出信号中，可能会破坏下变频频谱。在外差式接收器中，这个影响可以忽略。这是因为虽然在 IF 级需要克服两个干扰信号，但是其中一个通常落在接收频带之外，因而被前端滤波器衰减了。

图 6.20(a)中 I_{RF} 到 V_{out} 的直接馈送也会增加混频器的噪声因子。这是因为 I_{RF} 通常在中频附近也含有噪声成分。事实上，因为带宽的限制，混频器对这种成分的转换增益比 RF 要高许多。结果是，单端 IF 输出信号带有 I_{RF} 中的噪声。

相反，如果输出是用差分信号取出，直接馈送的影响会减小很多。在图 6.20(b)中，由于失配引起的非对称性以及 LO 占空比偏离 50%，都会导致一些馈送存在。但是，对于 IC 中的典型的失配量，馈送可以保持在几个百分点下。

尽管差分输出具有优点，但是大多数混频器还是采用单端输出，这主要是因为 IF SAW（表面声波）滤波器只对单端输出适用。为了减小直接馈送的影响并达到最大增益，常常采用一个外部 LC 网络将混频器输出的差分信号转换成单端形式[9]。图 6.21 就是这样一个例子[10]，其中 LC 网络包括 L_1 ， L_2 和 C_1 ，它将混频器的差分输出电流转换成单端电流 I_{out} 。不过，当差分 SAW 器件出现时，混频器的差分输出端就可以直接与 IF 滤波器相连。

图 6.21 差分电流到单端输出的变换

混频器伪响应 一般而言，混频器会产生 RF 和 LO 信号以及它们的谐波的各种各样的交积。所得的频率成分可以表示成 $|m\omega_{RF} + n\omega_{LO}|$ ，其中 m 和 n 都为整数。在接收器设计中的困难就是确保除 $|\omega_{LO} - \omega_{RF}|$ 外，所有其它成分都不要落在 IF 频带内。由于 RF 路径中的非线性，干扰源的谐波可能会和 LO 信号的谐波差

拍，而破坏了下变频信号。对不同的 m 和 n 值组合下的伪响应通常是用计算机程序进行辅助分析。

6.2.2 双极型混频器

单均衡有源双极型混频器可以用如图 6.22 所示的电路实现，在这里发射极筒并被利用来获得相当的线性度。在每个电路中，RF 输入信号改变了 Q_1 的集电极电流， $Q_2 - Q_3$ 的轮换操作将这种变化乘以正方形波。图 6.22(a) 的电路具有相对较高的输入阻抗，该电路不适合外差式接收器，在那里镜像抑制滤波器的终端必须接 50Ω 的阻抗。另一方面，在图 6.22(b) 中选择 $R_E + 1/g_{m1}$ 近似等于 50Ω [2]。如在有关共基极放大器的讨论中提到的（6.1.3 节），当产生相同的噪声时，源阻抗 R_S 在图 6.22(b) 对 RF 端口线性化的程度远远大于图 6.22(a) 中的。

图 6.22 双极型混频器，其 RF 信号加到输入器件 (a) 基极，(b) 发射极

计算图 6.22(b) 中的混频器的低频电压和功率转换增益是具有指导作用的。计算是在假定当 $Q_2 - Q_3$ 对在输入信号 V_{LO} 过零点的转换是完全瞬时发生的。 Q_1 的小信号集电极电流等于 $I_{c1}(t) = V_{RF}(t) / (R_S + R_E + 1/g_{m1})$ 。如果 LO 波形具有 50% 的占空比，这时 $I_{c2} - I_{c3}$ 等于 I_{c1} 与一个在 -1 和 +1 间反复转换的正方形波的乘积，因而得到

$$V_{out}(t) = \frac{V_{RF}(t) R_C}{R_S + R_E + 1/g_{m1}} \cdot \frac{4}{\pi} \cos \omega_{LO} t \quad (6.22)$$

其中，高次 LO 谐波忽略不计。因为 $V_{RF}(t)$ 与 $\cos \omega_{LO} t$ 相乘等效于将 $V_{RF}(\omega)$ 位移 $\pm \omega_{LO}$ ，所得结果再除以 2，所以 IF 输出信号在频域中的表示如下：

$$V_{IF}(\omega) = \frac{V_{RF}(\omega - \omega_0)}{R_S + R_E + 1/g_{m1}} \frac{2R_C}{\pi} \quad (6.23)$$

电压转换增益等于输出 IF 电压 V_{IF} 除以输入 RF 电压（在结点 X 处），

$$A_V = \frac{(2R_C/\pi)(R_S + R_E + 1/g_{m1})}{(R_E + 1/g_{m1})/(R_S + R_E + 1/g_{m1})} \quad (6.24)$$

$$= \frac{2}{\pi} \frac{R_C}{R_E + 1/g_{m1}} \quad (6.25)$$

当输入匹配时， $R_S = R_E + 1/g_{m1}$ ，上式简化成：

$$A_V = \frac{1}{\pi} \frac{2R_C}{R_S} \quad (6.26)$$

下面，计算功率转换增益，IF 传递到负载端的功率平均值等于：

$$P_{IF} = \frac{V_{IF,rms}^2}{2R_C} \quad (6.27)$$

其中，分母上的因子 2 是因为 $V_{IF,rms}$ 为 $2R_C$ 上的电压降。根据 (6.23) 式，我们可

以得出：

$$P_{IF} = \frac{V_{RF,rms}^2 R_C^2}{(R_S + R_E + 1/g_{m1})^2} \frac{4}{\pi^2} \frac{1}{2R_C} \quad (6.28)$$

$$= \frac{V_{RF,rms}^2 R_C}{(R_S + R_E + 1/g_{m1})^2} \frac{2}{\pi^2} \quad (6.29)$$

可以提供的最大信号源功率为：

$$P_{in} = \frac{V_{RF,rms}^2}{4R_S} \quad (6.30)$$

因而，功率增益为：

$$A_P = \frac{P_{IF}}{P_{in}} \quad (6.31)$$

$$= \frac{8}{\pi^2} \frac{R_S \cdot R_C}{(R_S + R_E + 1/g_{m1})^2} \quad (6.32)$$

当 $R_S = R_E + 1/g_{m1}$ 时，上式可以变成：

$$A_P = \frac{1}{\pi^2} \frac{2R_C}{R_S} \quad (6.33)$$

从等式 (6.26) 和 (6.33) 可以看出, 一般而言, $A_p \neq A_v^2$ 。因而, 当以分贝表示时电压和功率转换增益不相等。只有在极少的特殊情况下, 当差分负载电阻 $2R_C$ 等于源电阻 R_S 时, $A_p = A_v^2$ 才成立。

电压和功率转换增益的对应关系可以用下式来表示:

$$A_p = \frac{V_{\text{out}}^2/R_L}{V_{\text{in}}^2/R_S} \quad (6.34)$$

$$= A_v^2 \frac{R_S}{R_L} \quad (6.35)$$

表明, A_p 和 A_v^2 之间相差 $10\log(R_S/R_L)$ 。

事实上, 由于信号路径上寄生电容的存在使得电路的转换增益比上面计算得到的要小一些。一般更精确的值都是要通过仔细的模拟得到的。

将有关噪声因子的计算推迟到 6.2.4 节, 现在我们来考虑如图 6.22 所示的混频器的非线性。有源混频器失真主要是由于输入端电压到电流的转换器的非线性引起的。例如, 在图 6.22(a)和图 6.22(b)中这个转换器是由 Q_1 , R_E 和 R_S 组成的网络。混频器所要求的高 IIP_3 值决定了在输入端要使用高度的电阻性兼并, 这样就要求在噪声, 非线性和功耗之间作出折衷选择。

图 6.22 所示混频器中的另一种非线性是由如下机理引起的, 即 $Q_2 - Q_3$ 基极-发射极扩散电容随 Q_1 的集电极电流的变化[11]。由于实际上 LO 信号与正弦波很接近, 因而差分对的转换不是突然进行的, 在一个周期内会有一小部分时间 Q_2 和 Q_3 是同时导通的。在这个时间内, $Q_2 - Q_3$ 基极-发射极电容充当了 Q_1 集电极电流的非线性分流, 并使得 Q_2 和 Q_3 的集电极电流失真。可以通过提高 LO 的转换速率 (slew rate) 或降低差分对的基极电阻来降低这种作用的影响[11]。

考察图 6.22(a)和 6.22(b)中 LO-RF 馈送也是很有意义的。在这两种情况下, 如果 LO 信号是全差分信号, 同时 Q_2 和 Q_3 也是完全匹配的, 在结点 P 处的波形在每个 LO 周期内要经历两次全偏移, 因而表现出的频率成分等于 $2\omega_{LO}$ 及更高的

频率。在图 6.22(a)中这个波形通过 Q_1 的基极-集电极的结电容耦合到 RF 输入信号，在图 6.22(b)中它被共基极结构的高的逆向隔离大大削弱了。这里的关键点就是这些泄漏信号包含的频率成分远离所关心的频带，因此它带来的问题比在 ω_{LO} 处的直接馈送带来的问题的要小得多。

但是，实际上混频器在 ω_{LO} 处一定程度受到直接馈送的影响。这主要的原因是在大多数情况下 LO 信号都是单端并且由外部提供的，因而可以通过焊接引线和衬底与 RF 输入耦合。例如，在一个图 6.22(b)电路的实际实现中，单端 LO 信号到 RF 端口的馈送近似等于 -33 dB[2]，虽然 Q_1 已经有效地隔离了输入端与结点 P 处的波形。

虽然相比图 6.22(b)电路，图 6.22(a)表现出较低的 LO-RF 隔离，但是它具有较高的输入阻抗，因而它更适合 LNA 必须与混频器直接相连的情况，也就是说不必要有镜像抑制滤波器的结构。

用 Gilbert 单元实现的双均衡双极型混频器已被广泛应用。Gilbert 单元除了降低 LO-IF 穿透外，而且通过 Schmook 技术进一步改进了 RF 端口的线性度[12]。为了理解这个技术，首先注意，如图 6.23(a)所示，如果双极型差分对采用的两个晶体管的发射区是不同的，分别为 A 和 nA ，那么它的输入-输出特性被水平平移了 $V_T \ln n$ 。现在如果两个这样的差分对是交叉相连的[如图 6.23(b)所示]，这样电路的整体特性会趋向于“更平滑”。图 6.23(c)给出的是复合差分对的等效跨导 G_m 作为输入电压的函数，揭示了当 $n \approx 4$ 时 G_m 中间出现了一个平坦的部分 [12]。

图 6.23 (a) 非对称的差分对，(b) Schmook 的线性化的差分对，(c) 等效 G_m 随着输入电平的变化。

在实际上，为了得到满意的线性度，如图 6.23(b)所示的交叉连接差分对仍然需要采用某些电阻性兼并电路。此外，两种线性化方法的结合能够在给定的 IIP_3 下得比单一电阻简并差分对更低的噪声因子。

图 6.24 所示为另一种双均衡双极型混频器。这里，一个片上变压器不仅担

当了从单端信号变换到差分信号的转换器，同时还提供了输入匹配。起到开关作用的四个晶体管的偏置电流由 I_{EE} 提供，电容 $C_1 - C_3$ 在变压器的主级和次级造成谐振。采用 11-GHz 双极型工艺实现该体系结构在 1.9-GHz 下的电路，可以得出当 IIP_3 等于 +2.3dBm 时，电路的 SSB NF 为 10.9dB，此时电源电压为 1.9-V，功耗为 5mW。

图 6.24 在 RF 输入端采用变压器的混频器

其它有关线性化的技术在[13, 14]中都有描述。

6.2.3 CMOS 混频器

如图 6.19 所示，有源 CMOS 混频器可以通过单均衡和双均衡拓扑结构实现。前面提到的双极型混频器的设计考虑大部分也都可以应用到对应的 CMOS 结构中。例如，特别是在设计 RF 端口电压-电流转换器时，为了满足噪声、非线性度和功耗间的平衡，最后还是要采用一种折衷方案。

让我们考虑图 6.25 所示的单均衡混频器。这个电路中的 LO 电压摆幅直接影响到转换增益和噪声。这是因为 MOS 开关对 $M_2 - M_3$ 与相应的双极型相比需要更高的摆幅来进行完全地开关操作。如果在一个周期内有相当长的一段时间 M_2 和 M_3 同时导通，这时在这个时间内，由 M_1 产生的一些 RF 电流会被当作共模信号“浪费”了。这样会降低转换增益，并且我们在 6.2.4 节中将要看到，增加 M_2 和 M_3 对噪声的贡献。

图 6.25 CMOS 有源混频器

在给定的 LO 驱动下，通过增加 M_2 和 M_3 的宽度或降低它们的漏极电流都可

惟使 M_2 和 M_3 的转换变得更加急剧。前一措施增加了结点 P 处的看到的电容，从而短路了 RF 电流³。后一方法降低了 M_2 和 M_3 的跨导，增大了这些器件从源端看进去的阻抗，因而允许更多的 RF 电流流过结点 P 处的电容。正是因为这些原因，在选择总的偏置电流和 M_2 与 M_3 的宽度时要考虑有以用到的 LO 信号摆幅。

混频器的 RF 端口的线性化可以通过提高输入晶体管的栅-源过驱动电压实现。但是，对于给定的偏置电流，高的过驱动电压导致跨导变低，因而增加了噪声因子，并降低了转换增益。在低电压设计中，混频器的线性度和增益之间的平衡通常都成了难题所在。

CMOS 混频器也可以用无源形式实现。如图 6.26(a)所示是一个下变频混频器，它是通过 LO 的互补信号来驱动 M_1 和 M_2 转换。该电路与对应的有源电路相比，具有两个优点。第一，如果当 M_1 和 M_2 在处于导通状态时经历了一个高的栅-源过驱动电压以致于 V_{RF} 不能显著地改变导通电阻时，该电路它可以达到一个较高的 IP_3 。第二，它不会从电源电压汲取功率。

图 6.26 (a) 无源 CMOS 混频器，(b) 开关导通电阻的变化

但是，图 6.26(a)所示的无源混频器也具有一些严重的缺点。第一，因为电路的增益小于 1，当以 V_{RF} 为参考时，混频器的下一级产生的噪声是被“放大”了。用 6.2.2 节相似的方法进行计算，可以得出，在理想的瞬时转换的情况下，混频器的电压增益等于 $2/\pi$ (≈ -4 dB)。对于正弦 LO 信号，由于 M_1 和 M_2 在一个周期内相当长的时间是同时导通的，所以增益会变得更小。第二，为了得到低的导通电阻， M_1 和 M_2 的宽度要大，这就导致从 LO 到 IF 产生大量的电容性馈送。

无源混频器的线性度的优势在低电源电压时减小了。如图 6.26(b)所示，对于正弦 LO 信号，MOS 开关在一个周期中相当长时间内经受一个相对较小的栅-源过驱动电压。因而，在这段时间内，RF 的输入信号大大改变了转换导通电阻，

³ M_2 和 M_3 的跨导也增加了，但是比起节点 P 处的电容增加还是慢些。

导致了失真。

6.2.4 混频器中的噪声

混频器的时变性和频率的变换使得很难去计算它的噪声因子。如图 6.27 的概念性图示说明，在经过下变频器前噪声成分位于 RF 范围内，而经过下变频后噪声成分位于 IF 范围内。通过转换级的 RF 噪声成分的转换是不能在 SPICE 中直接使用小信号交流和噪声分析的，要求模拟必须在时域中进行。但是，现在大多数 SPICE 版本都不支持时域噪声分析。另一个困难是开关级贡献的噪声呈现出时变统计特性。

图 6.27 混频器的概念图

定性分析 在计算混频器的噪声之前，我们先对图 6.28(a)的单均衡拓扑结构进行一些定性地观察。在这个电路中什么是噪声源？在 RF 路径中，由 Q_1 基极电阻和发射极电阻 R_E 产生的热噪声，以及集电极 Q_1 产生的散粒噪声构成了主要的成分。在 IF 路径中，电阻 R_{C1} 和 R_{C2} 引入热噪声。

图 6.28 (a) 单均衡混频器，(b) 当 Q_3 截止时 Q_2 的噪声贡献

需要更加关注的是由 Q_2 和 Q_3 贡献的噪声。首先，假定我们假设晶体管的转换都是在瞬间进行的。接着，如图 6.28(b)所示，在 LO 周期内每个器件都有一半的时间是导通的，将噪声注入到输出端。这是因为结点 P 处存在的寄生电容 C_p 提供了到地的阻抗。这个电容起源于 Q_2 和 Q_3 的基极-发射极结和 Q_1 的集电极-基极和集电极-衬底结，如果为了减小晶体管的基极电阻而采用了大尺寸的晶体管，这个电容会变得很大。因而，由基极电阻和 Q_2 的集电极电流产生的 RF 噪声被该晶体管的开关操作转移到 IF 信号里。

现在我们考虑更为实际的情况，LO 信号不再是一个正方形波，在一个周期

内有一部分时间时 Q_2 和 Q_3 会同时导通。在同时导通的时间内，两个晶体管放大了它们基极电阻产生的热噪声并将它们集电极的散粒噪声注入到输出端。从另外一个角度来讲，在这种情况下 Q_2 的操作与图 6.28(b)描述的情况相似。不同的是在目前讨论的情况下 Q_2 的发射极看到一个低的对地的阻抗 ($\approx 1/g_{m3}$)。

当 Q_2 和 Q_3 同时导通，都向输出端注入噪声，这时 I_{C1} 中的噪声影响就会很小，这是因为它主要以共模信号的形式出现。换句话说，当 Q_2 和 Q_3 的噪声变得很大时， I_{C1} 的噪声就变得不重要了，反之亦然。

通过上面的观察，我们推断可以采用以下的方式来减小 Q_2 和 Q_3 贡献的热噪声和散粒噪声。(1)采用大的 LO 摆幅(在避免 Q_2 和 Q_3 饱和的情况下);(2)降低 C_p ，也就是减小 $Q_1 - Q_3$ 的尺寸，因而产生更高的基极电阻噪声；(3)降低 Q_2 和 Q_3 的基极电阻，但是会导致 C_p 的增大；(4)减小 Q_2 和 Q_3 的集电极电流。最后一个方案需要进一步解释。因为 Q_2 和 Q_3 出现在信号电流路径上，它们的散粒噪声电流 $\overline{I_n^2} = 2qI_C$ 直接破坏了信号，因此当 I_C 减小的时候这个噪声电流也减小。与此不同， Q_1 的散粒噪声的影响可以被看作是以输入端作参考的电压， $\overline{V_n^2} = 2kT/g_m = 2kTV_T/I_C$ ，如果 I_C 增加它会减小（只要基极散粒噪声的效应可以忽略）。换言之，当 Q_1 的集电极电流减小的时候，它会贡献更多的噪声而 Q_2 和 Q_3 的贡献就会减少。

上面提到的折衷选择要求仔细选择器件的尺寸和偏置电流以使得整体的噪声因子达到最小。为了让 Q_1 和 $Q_2 - Q_3$ 的集电极电流值互相独立，可以增加一个电流源，如图 6.29 所示[15]。例如，如果 $I_S \approx 0.8I_{C1}$ ，这时 Q_2 和 Q_3 的集电极电流就减小了 5 倍，大大降低了 Q_2 和 Q_3 的散粒噪声电流。进一步说，对于给定的，可以允许的 R_{C1} 和 R_{C2} 两端的电压降，它们的电阻值也能增加 5 倍，因而也提高了电压转换增益。但是，这种方法还面临两个问题。第一，当 Q_2 和 Q_3 的集电极

电流减少时，朝它们的发射极看进去的阻抗增加，导致 Q_1 提供更多的 RF 电流通过 C_p 分流到地。第二，由 I_s 产生的噪声电流直接加到了 RF 信号上。由于这些原因，在混频器上增加 I_s 也许只能降低一小部分噪声因子。

图 6.29 增加电流源 I_s 以降低 Q_2 和 Q_3 的集电极电流

另外还有两个有关图 6.28(a)的发现也是很重要的。第一，由 LO 电路的输出电阻产生的热噪声也会增加噪声因子，因此（为了减小这个噪声）通常要求在 LO 输出缓冲器通常消耗大的功率。例如，如图 6.30，如果在 LO 端口前有一差分对，这一级产生的噪声比由 Q_2 和 Q_3 产生的要高的多。这种效应因为典型的振荡器具有高的噪声基底而变得更为严重。注意，对于 LO 噪声，双均衡混频器具有更好的抗干扰性。第二，如果混频器以单端形式输出，比如说，在结点 X 对地，这时在 Q_1 的集电极电流中的低频噪声不经频率变换就直接抵达输出。如 6.2.1 节中所解释的，这是因为在 IF 范围中的 I_{c1} 的噪声成分也出现在 V_x 中。由于低频噪声不被在 RF 信号路径中的寄生电容衰减，这种现象是十分严重的。

图 6.30 在混频器中加入 LO 输出噪声

如果集电极散粒噪声为沟道热噪声所取代，则上面大多数分析都可以应用到 MOS 电路实现。下面是双极型和 MOS 混频器有关噪声特性的一些不同之处。

在图 6.25 的 MOS 混频器中， M_2 和 M_3 通常需要大约 1V 的差分 LO 信号来驱动以完成完全地转换。用正弦 LO 信号时，与对应的双极型结构相比，在一个周期内两个晶体管同时导通的时间要长一些。这样就可能注入更多的噪声到输出端。但是，在给定偏置电流下，MOSFET 的沟道噪声电流通常要比双极型晶体管的集电极散粒噪声电流小几倍。因此，在两种拓扑结构中转换对整体噪声的贡献是近似相等的。但是，由于在 M_2 和 M_3 同时导通的时间内 RF 信号是作为共模信号的扰动，使得用 MOS 管实现的电路的转换增益要小一些。这些参数强烈

依赖于个别的设计和可得到的 LO 摆幅。

另外一个不同点就是，对于相比拟的器件尺寸和寄生电容，MOSFET 的栅电阻远远小于双极型晶体管的基极电阻。

定量分析 考虑如图 6.31(a)所示的单均衡混频器。根据定性分析，我们可以将一个电路分成三个部分：RF 部分，时变部分和 IF 部分。为了获得噪声因子的估计值，我们采用下面的过程分析：(1)对每一个噪声源，计算到 IF 输出的“转换增益”；(2)将每个噪声的幅度乘以相应的增益，再将作为结果的功率相加，从而得到在 IF 输出端的整体噪声；(3)用噪声除以整体的转换增益，得到以输入端作参考的噪声电压 $\overline{V_{ni}^2}$ ；(4)计算 $NF = \overline{V_{ni}^2} / (4kTR_s)$ 。

图 6.31 (a) 混频器分块，(b) 噪声在 RF 模块中的模型，(c) 对时变模块中的噪声建模。

让我们采用上面的过程来分析图 6.31(a)所示的电路，假定 RF 输入频率在 900MHz 左右，LO 输出频率在 50MHz 左右。如图 6.31(b)所示，使用 SPICE 中的小信号交流分析用三个电流源来描述 R_s ， M_1 和 I_s 的噪声： I_{RF} 为 RF 范围， I_{im} 为镜像频率， I_{IF} 为 50MHz 范围。由于信号路径上的寄生电容， I_{RF} 和 I_{im} 的贡献是不同的，应该分开建模。我们也用 $I_{n,RD}$ 来描述 R_D 的 IF 热噪声。

下一步，我们计算每个电流源到输出端的转换增益。用 $I_{RF} = I_1 \cos \omega_1 t$ ， $I_{im} = I_2 \cos \omega_2 t$ ， $I_{IF} = I_3 \cos \omega_3 t$ 代入计算，并将 LO 信号输入 M_2 和 M_3 中，我们进行时域模拟，然后用快速傅立叶变换(FFT)来计算输出电压中 IF 的含量。为了区分每个噪声源对应的贡献，选择频率 ω_1 ， ω_2 和 ω_3 以使得它们在 IF 输出中对应的成分不会重合。 $I_{n,RD}$ 的贡献可以容易地包括在内。

转换对对噪声的贡献也是要考虑的。要包含这个噪声在内的困难是 M_2 和 M_3 的漏极电流是周期性改变的，使得对噪声的统计也是时间的周期性函数。作为一

个一阶近似，我们可以用一个与每个栅极串联的电压源来模拟噪声[如图 6.31(c)所示]，并根据 M_2 和 M_3 在平衡条件 ($I_{D2} = I_{D3}$) 下的跨导值来选择电压幅度。然后，再通过求得每个噪声源的转换增益来计算 IF 噪声（此时 LO 信号也加上）。另一种方法是噪声也可以看成是晶体管漏极电流的函数[16]。

通过以上的计算，在 IF 范围内的输出噪声可通过每个噪声源乘以各自的转换增益，再将噪声功率相加得到。在这个计算过程中，电路中不同点的阻抗值不同会在定义功率量时导致混淆。为了避免这个问题，输出端所有的噪声量都用电压形式表示。那么，整个均方输出噪声电压除以混频器的电压转换增益的平方就可以得到以输入端为参考的噪声值。注意，正如第三章所说，这个计算过程中用到的电压增益是从电压源而不是从结点 A 到输出端的电压增益。于是，以输入端作参考的噪声用 $4kTR_s$ 归一化后，就得到噪声因子。

其它用于计算混频器噪声的方法在参考文献[17]中有记载。

6.3 级联级再讨论

我们在级联级中有关噪声的分析导致了在第二章中的 Friis 方程。在实际过程中，如果这个方程被应用到一个阻抗值从一块到另一块变化的级联级中会出现很多混淆。在这一节中，我们研究两个例子来进一步巩固噪声，损耗，电压转换增益和功率转换增益的概念。

考虑如图 6.32 所示的电路，混频器的输入和输出阻抗分别等于 50Ω 和 500Ω 。如何计算电路的噪声因子呢？第二级的 NF 是要被第一级的电压增益或是被功率增益相除呢？在这个例子中噪声因子这个概念不但没有带来方便，而是造成了麻烦。

图 6.32 一个混频器和一个中频放大器的级联

为了得到整体的噪声因子，我们采用两种不同的方法：(1)避免对 A_{IF} 定义 NF，将 A_{IF} 简单地认为是混频器的一部分，再根据 6.2.4 节中有关计算噪声的

要点进行计算。从一个“模拟”的角度来看， A_{IF} 可以用输入噪声电压和输入噪声电流完全描述，而不需要计算它的噪声因子。(2)回到 Friis 公式的推导过程，特别注意 NF 的定义。

虽然方法(1)有助于 IC 设计，但是如果混频器和 IF 放大器是分立的电路并且仅由它们的噪声因子和增益来表征，这时就必须使用方法(2)。正是因为这个原因，我们依旧要找出 A_{IF} 的噪声因子是否是除以混频器的电压增益还是功率增益得到的。这里混淆产生的第一原因是 A_{IF} 的 NF 定义。图 6.32 中第二级的 NF 应在 $R_s = 500\Omega$ 基础上进行计算，但是厂商的数据表中通常给出的是源阻抗为 50Ω 时的 NF 值。如何将这两个量联系起来呢？如果放大器中以输入端为参考的噪声可以仅用一个电压源来模拟，也就是说，输入噪声电流可以忽略（图 6.33），这时两个不同的源阻抗对应的噪声因子分别可以表示成， $NF_A = 1 + \overline{V_n^2} / (4kTR_{SA})$ 和 $NF_B = 1 + \overline{V_n^2} / (4kTR_{SB})$ ，从而得到：

$$(NF_A - 1)R_{SA} = (NF_B - 1)R_{SB} \quad (6.36)$$

图 6.33 相对于信号源电阻 R_{SA} 及 R_{SB} 的噪声因子的计算

如果放大器的输入噪声电流不能忽略，则不能用式（6.36）表示两者的关系。如图 6.34 所示，在每种情况下总的以输入端为参考的噪声电压等于：

$$\overline{V_{nA,tot}^2} = \overline{V_n^2} + R_{SA}^2 \overline{I_n^2} \quad \text{和} \quad \overline{V_{nB,tot}^2} = \overline{V_n^2} + R_{SB}^2 \overline{I_n^2}$$

，其中 V_n 和 I_n 间的关联被忽略掉了。

图 6.34 考虑有限大小的输入噪声电流的噪声因子计算

因而，

$$(NF_A - 1)4kTR_{SA} - R_{SA}^2 \overline{I_n^2} = (NF_B - 1)4kTR_{SB} - R_{SB}^2 \overline{I_n^2} \quad (6.36)$$

上式表明如果 $\overline{I_n^2}$ 未知，一个 NF 值不能通过另一个 NF 值计算得到。对于工作在比 50MHz 更高频率的 IF 放大器， $\overline{I_n^2}$ 的贡献会是十分大的，因而式（6.36）不一

定能提供精确的变换。不过，我们依然忽略 $\overline{I_n^2}$ 来继续进行推导，这是因为这个电流的影响最终被放大器前面各级的总增益被抑止。

产生混淆的第二个原因是 Friis 方程中不同的级间阻抗值。回顾第二章在 Friis 方程中，仅当 NF_2 为对应 R_{out1} 时计算得到的噪声因子时， $NF_2 - 1$ 才能通过除以前级的有效功率增益得到。因而，在图 6.32 中如果 IF 放大器的噪声因子对应于源阻抗为 500Ω ，则 $NF_{IF} - 1$ 必须被除以混频器的功率转换增益。另一方面，如果只知道源阻抗为 50Ω 时的 IF 混频器的噪声因子，而且输入噪声电流可以忽略不计，那么 $NF_{IF} - 1$ 必须被除以混频器的电压转换增益的平方。

如图 6.35 所示，我们现在研究一种典型的外差式接收路径，并计算它的总噪声因子。假定端口 A，B，C 和 D 都是共轭匹配且阻抗等于 50Ω ，IF 滤波器看到的源和负载阻抗等于 500Ω 。并且 NF_6 对应于一个源阻抗 500Ω 。我们从最的一级开始分析。因为一个匹配的损耗电路提高了噪声因子一个量，这个量就等于它的损耗，端口 E 的噪声因子为 $NF_E = NF_6|_{dB} + L_5|_{dB} = 15\text{ dB} = 31.6$ 。因而，在端口 D 的 NF 为 $NF_D = NF_4 + (NF_E - 1)/A_{p4} = 15.85 + 9.68 \approx 25.53 = 14.1\text{ dB}$ ，在端口 C 的 NF 为 $NF_C = L_3|_{dB} + NF_D|_{dB} = 20.1\text{ dB} = 102.3$ 。

图 6.35 级联级的噪声因子的计算

由此得到，在 LNA 输入端口的等效 NF 可以表示为，
 $NF_B = NF_2 + (NF_C - 1)/A_{v2}^2 = 1.58 + 3.2 = 4.78 \approx 14.1\text{ dB}$ 。结果是，总的噪声因子等于 $NF_{tot} = L_1|_{dB} + NF_B|_{dB} \approx 8.79\text{ dB}$ 。注意 LNA 以后各级对总噪声因子的贡献很大。

“强度级别图表”是用来阐明在级联中每一级在信号，噪声和交调乘积中的作用的一个有力工具。图 6.36 为对应于图 6.35 中的级联的强度级别图表。注意除有每一级的参数外，还指出了在相邻级间界面上累积量的强度。

图 6.36 对应于图 6.35 所示的级联级的强度级别图

为了构筑一个这样的强度级别图表，有几点需要注意。第一，对累积的增益的计算是从第一级开始的，反之，对累积噪声因子和 IP_3 的计算是从最后一级开始的。第二，无源滤波器的 IP_3 值通常假定是一个很大的值，例如，等于 +100 dBm。第三，既然 IF 滤波器一定程度上抑止了相邻频道的干扰，IF 滤波器输入端的累积 IP_3 值要成比例地高于用简单关系式预测的值。这个简单关系式在第 2 章中对级联级的 IP_3 已推出。第四，因为第一个混频器之后的阻抗值不一定等于 50Ω ，随后级的 IP_3 应以伏特或 dBV 表示。同时，如果用伏特表示，每一级的 IP_3 值是被该级在有负载时的电压增益相除得到。这样，强度级别图表必须同时指出每一级的功率增益（为计算 NF ）和电压增益。

使用图表中的各级 IP_3 值，我们计算级联结构的累积 IP_3 。假定 IF 滤波器使干扰信号降低 30 dB，而且有用的频道衰减 5 dB。那么，E 点的等效 IP_3 是多少呢？如图 6.37 所示，因为输入到 IF 放大器的干扰信号比 E 点的要低 30 dB，则有

$$\frac{1}{IP_{3,E}^2} = \frac{1}{IP_{3,fil}^2} + \frac{M^2}{IP_{3,amp}^2} \quad (6.38)$$

图 6.37 邻近频道为 IF 滤波器所压制

其中， $M = 1/31.6 (= -30 \text{ dB})$ 。假设滤波器的线性度是很好的，

$$IP_{3,E} = \frac{IP_{3,amp}}{M} \quad (6.39)$$

$$= 22.1 V_{rms} \quad (6.40)$$

点 D 处的累积 IP_3 可以用下式表示：

$$\frac{1}{IP_{3,D}^2} = \frac{1}{IP_{3,mix}^2} + \frac{A_{v4}^2}{IP_{3,E}^2} \quad (6.38)$$

因而， $IP_{3,D} \approx 398 \text{ mV}_{rms} \approx +5 \text{ dBm}$ 。注意 $IP_{3,D}$ 受控于混频器的非线性度。类

似地, $IP_{3,C} \approx +11$ dBm, $IP_{3,B} \approx -12.6$ dBm 和 $IP_{3,A} \approx -10.6$ dBm。

参考文献

- [1] A. A. Abidi, "High-Frequency Noise Measurements on FETs with Small Dimensions," *IEEE Trans. Electron Devices*, Vol.33, pp. 1801-1805, November 1986.
- [2] R. G. Meyer and W. D. Mack, "A 1-GHz BiCMOS RF Front-End Integrated Circuit," *IEEE Journal of Solid-State Circuits*, Vol.29, pp. 350-355, March 1994.
- [3] J. R. Long and M. A. Copeland, "A 1.9 GHz Low-Voltage Silicon Bipolar Receiver Front-End for Wireless Personal Communication Systems," *IEEE Journal of Solid-State Circuits*, Vol. 30, pp. 1438-1448, December 1995.
- [4] B. Razavi, R. H. Yan, and K. F. Lee, "Impact of Distributed Gate Resistance on the Performance of MOS Devices," *IEEE Trans. Circuits and Systems -I*, Vol. 41, pp. 750-754, November 1994.
- [5] A. N. Karanicolas "A 2.7 V 900 MHz CMOS LNA and Mixer," *IEEE Journal of Solid-State Circuits*, Vol. 31, pp.1939-1944, December 1996.
- [6] D. K. Shaeffer and T. H. Lee, "A 1.5 V 1.5 GHz CMOS Low Noise Amplifier," *VLSI Circuits Symp. Dig. Tech. Papers*, pp.32-33, June 1996.
- [7] Triquint Semiconductor, "TQ9203, Low-Current RFIC Downconverter," in *Wireless Communication Products*, 1995.
- [8] A. R. Shahani, D. K. Shaeffer, and T. H. Lee, "A 12 mW Wide Dynamic Range CMOS Front End for Portable GPS Receivers," *ISSCC Dig. Tech. Papers*, pp. 368-369, February 1997.
- [9] Philips Semiconductor, *Philips RF/Wireless Communication Data Hand book*, 1996.
- [10] A. K. Wong, S. H. Lee, and M. G. Wong, "Current Combiner Enhances Active Mixer Performance," *Microwaves and RF*, pp. 156-165, March 1994.
- [11] R. G. Meyer, "Intermodulation in High-Frequency Bipolar Transistor Integrated-Circuit Mixers," *IEEE Journal of Solid-State Circuits*, Vol.21, pp. 534-537, August 1986.
- [12] J. C. Schmoock, "An Input Stage Transconductance Reduction Technique for High Slew Rate Operational Amplifiers," *IEEE Journal of Solid-State Circuits*, Vol.10, pp.407-411, December 1975.
- [13] B. Razavi, "A 1.5-V 900-MHz Downconversion Mixer," *ISSCC Dig. Tech. Papers*, pp.48-49, February 1996.
- [14] J. Durec, E. Main, and D. Lovelace, "Motorola's MOSAICV Silicon Bipolar Building Blocks Fill Gaps in High Performance Low Power Wireless Chip Sets," *Proc. Fourth Annual Wireless Symp.*, pp. 218-223, February 1996.
- [15] W. H. Sansen and R. G. Meyer, "Distortion in Bipolar Transistor Variable Gain Amplifiers," *IEEE of Solid-State Circuits*, Vol. 8, pp.275-282, August 1973.
- [16] B. Razavi, "A Study of Phase Noise in CMOS Oscillators," *IEEE Journal of Solid-State Circuits*, Vol. 31, pp.331-343, March 1996.

[17] C. D. Hull and R. G. Meyer, "A Systematic Approach to the Analysis of Noise in Mixers," *IEEE Trans. Circuits and Systems—I*, Vol. 40, pp. 909-919, December 1993

第 7 章 振荡器

在第五章关于射频接收发送器的学习中,我们已经认识到了无论在发射路径还是接收路径中,振荡器都被广泛应用。有趣的是,在多数系统中,混频器的输入端之一都是由一个周期信号驱动的,这样就需要一个振荡器。

本章讨论用于射频电路的振荡器设计,主要强调单片集成的实现。由一些一般性的问题讨论作为开始,我们考虑基本的振荡器拓扑结构,并且研究振荡器中的相位噪声。接着我们将描述双极型和 CMOS 振荡器以及目前产生正交和单边带信号的一些方法。

7.1 概述

振荡器产生周期性的信号。因此,它的电路必须有一个“自我维持”的机制,使得它自身的噪声可以增大,并且最终成为一个周期信号。

大多数的 RF 振荡器可以看作是一个反馈电路。考虑图 7.1(a)中简单的线性反馈系统,它的总传输函数为:

$$\frac{Y(s)}{X(s)} = \frac{H(s)}{1-H(s)} \quad (7.1)$$

当 s_0 满足 $H(s_0) = +1$ 时,一个“自我维持”的机制将启动,并且当 s_0 为纯虚数,也就是当 $H(s_0 = j\omega_0) = +1$ 的时候,震荡的幅度保持为一个常数。这样,作为一个稳定的振荡,在 ω_0 时必须同时满足两个条件: (1) 环路增益 $|H(j\omega_0)|$ 必须为 1; (2) 环路总的相移 $\angle H(j\omega_0)$ 必须等于 0 (如果直流反馈为负,则应为 180°)。

图 7.1 (a) 反馈振荡系统, (b) 在 (a) 上增加频率选择网络

上述的条件被称为 **Barkhausen** 准则,上面这些条件暗示了只要恰当的选择其环路增益和相移,任何一个反馈系统都会振荡。环形振荡器和相移振荡器就是例子。然而,在多数的射频振荡器中的环路中,通常都包括一个选频网络,例如 LC 槽路,来稳定振荡的频率。虽然这样一个选频网络也可以是 $H(s)$ 的一部分,在图 7.1(b)的概念性的说明中还是将它们分开了。我们将在 7.4 节中讨论频率的稳定性问题。这里提到的选频网络也称作一个“谐振器”。

我们应该指出,在某些情况下, **Barkausen** 准则只是一个必要条件,而非充分条件[2]。作为一个简单的例子,如果在频率为零时环路总的相移为 360° ,并且环路增益足够大,那么电路将锁定而非振荡。

上述振荡器的分析方法在微波理论中被称为“二端口”模型,这是由于反馈回路是环绕着一个二端口网络,比如说图 7.1(a)中的 $H(s)$,完成的。作为一个对照,“一端口模型”将

振荡器当作相互对接的两个一端口网络[图 7.2(a)]。为了理解这个模型，假设谐振器是一个简单的 LC 槽路，图 7.2(b)给出了这样一个槽路，包括寄生电阻。回忆一下第二章，对于一个窄的频率带，此电路可以转换为图 7.2(c)描述的一个并联回路。这个槽路自身不能无限地振荡，因为在每个周期中存储的能量会由 R_p 消耗掉。一端口模型的思想在于，一个有源网络产生一个恰好等于 $-R_p$ 的阻抗，因此由本征的，无损的谐振回路看到等效的并联电阻其值为无穷大。本质上，在每一个周期， R_p 所消耗的能量由有源电路补充，由此得到稳定的振荡。

图 7.2 (a) 振荡器的单端口模型，(b) LC 谐振器，(c) (b) 的等效电路

图 7.1(a)和 7.2(a)分别对应的模型在很多情况下是等效的，除非电路中包含本身就是一端口的有源电路，例如隧道二极管。射频应用中的振荡器通常属于反馈的范畴，但在可以适用的范围内，一端口模型可提供关于它们工作原理的额外的，深层次的信息。

振荡的频率通常由电路的属性决定，例如图 7.2(a)中的 LC 槽路的谐振频率。关于振荡的幅度又如何呢？自我维持效应使得电路噪声开始可以逐渐增大，但它需要另一种机制来限制其增长。例如，当幅度增加时，放大器出现饱和，使得环路增益在波形的峰值处降低[1]。从另外一种观点来看，为了启动振荡，电路的小信号环路增益必须稍微大于 1，但为了达到一个稳定的振幅，“平均”的环路增益必须回归到 1。

定义振幅的另一种机制是自动电平控制(ALC, automatic level control)。见图 7.3，这种技术利用峰值检测器（或整流器）来测量振幅，与参考值作为比较，并且用负反馈来调整振荡器的增益。这样，在启动的时候，振幅小而增益大，而在稳态，振幅近似等于 V_{REF} 。如果 V_{REF} 取的合适，ALC 的应用可以使得振荡器保持为线形，从而得到一个低失真的正弦输出。然而，由 ALC 带来的电路的复杂性的提高和可能的额外噪声阻碍了这项技术在射频领域的应广泛用。

图 7.3 自动电平控制

振荡器理想的输出波形是什么样的呢？回忆第 6 章中对于混频器的分析，如果本振波形表现为陡峭的变化，则由开关对产生的噪声达到最小。而且，当本振信号具有 50% 占空比时，直接馈送等效效应可以被最大的抑制。这样，在这种情况下，理想的波形应该是方波。然而在实际中，由于电路带宽的限制，尤其是在 LC 振荡器中，得到的波形接近于正弦波。从而，如图 7.4 所示，这个方波波形是由一个振幅很大的正弦波产生，在跨过零点附近这个正弦波可近似一个方波。

图 7.4 用一个大的正弦波来近似方波

在某些应用中，两个或者更多的振荡器的输出被混频从而产生新的频率，这时理想的波形应该是低失真的正弦信号。我们将在 7.9 节讨论这个问题。

多数的振荡器希望的 50% 的占空比需要用到差分拓扑结构，这是在学习不同的振荡器

结构时需要记住的一个要点。

7.2 基本的 LC 振荡器拓扑

多数分立器件实现的射频振荡器只采用了一个有源器件。这是基于两个原因：(1)最小化噪声；(2)降低成本。在集成电路中，第一点仍然成立，而第二点已经很少考虑了。作为理解设计原则时的第一步，本节介绍的基本的振荡器都是属于“单管”型的。

从图 7.1(a)的反馈模型，我们猜想一个单管的 LC 振荡器可能在双极型晶体管的集电极有一个 LC 槽路而反馈信号加在基极或者发射极(图 7.5)(在 FET 实现中同样的思路也适用)。我们假设，如图 7.6(a)所示，槽路由一个电感和一个电容并联而成。由于在谐振的时候，槽路的阻抗为实数，电流和电压的相差为 0。这样，为了使总的相移为 0，反馈信号必须回到三极管的发射极。这在很多振荡器拓扑结构中是一个基本的思想。

图 7.5 从 (a) 集电极到基极，(b) 集电极到发射极的反馈

从谐振槽路到发射极的连接需要考虑一个重要的问题，那就是从发射端看到的电阻性的负载 $1/g_m$ 。如果集电极电压直接加到发射极，这个电阻会大大降低槽路的 Q 值，使得环路增益降低到小于 1，从而阻止振荡。由于这个原因，发射极阻抗必须在与槽路并联之前被转换到一个更高的值[图 7.6(b)]。

图 7.6 (a) 从集电极到发射极的直接反馈，(b) 插入一个阻抗变换器，(c) 用一个具体的变压器

一个变换发射极阻抗的简单的方法是在槽路中采用一个变压器，如图 7.6(c)。作为一个匝数比为 n 无损的降压器，从槽路一端看到的并联电阻为 n^2/g_m 。

无源阻抗变换也可以通过应用容性或者感性的分压器（见第 2 章）来实现。见图 7.7(a)和 7.7(b)，由此得到的电路分别称作 Colpitts 和 Hartley 振荡器。从第 2 章的推导中，我们注意到里槽路中等效的并联电阻近似等于 $(1+C_1/C_2)^2/g_m$ [图 7.7(a)]和 $(1+L_2/L_1)^2/g_m$ [图 7.7(b)]，从而使得等效的 Q 值也有相应比例的提高。

图 7.7 (a) Colpitts 和 (b) Hartley 振荡器

我们现在讨论电感和电容值的选择和 Q_1 的属性。谐振频率 $\omega_r = \left(\sqrt{L_{eq}C_{eq}}\right)^{-1}$ ，其中 L_{eq} 和 C_{eq} 是图 7.7(a)和(b)槽路中等效的电感和电容值。从达到一个大的电压摆幅的角度来看，我们期望增大电感的值。为了理解为什么这样，首先注意到电感的 Q 值一般是远低于电容的，从而可以把等效的并联电阻写为 $R_p = (L_{eq}\omega_r)^2/R_s$ ，其中 R_s 为电感的等效串联电阻来计入电感的损耗。在典型的电感中， L_{eq} 与 R_s 成正比变化。也就是说，如果 L_{eq} 增大一个因

子 m ，则 R_p 也增大同样的倍数。由于槽路谐振时的阻抗等于 R_p ，则在给定偏置电流下电压的摆幅也增大同样的因子。如果电感是做在片外的，增大它的值同时可以减小键合线电感的影响，这在大于 1GHz 时是一个重要的问题。

然而，要用最大的可能的电感值，牵涉到有两个折衷考虑。其一是电感的自谐振频率降低了，会趋近应用的振荡频率。其二是槽路电容变得受器件寄生参数的限制，使得很难通过加入一个可变电容到振荡器中来调节振荡频率。我们会在 7.4 节讨论这个问题。

晶体管 Q_1 是图 7.7 中振荡器中的主要噪声源，从而扮演着很重要的角色。基极电阻的热噪声和集电极的散粒噪声可以分别通过增加尺寸和减小偏置电流来减小到最小程度。然而，第一个方案会增加寄生电容，而第二个则会降低电压摆幅。这样，就需要一个折衷。偏置电流一般选择使得 X 点的摆幅达到最大，同时 Q_1 保持不达到饱和。

Colpitts 振荡器只采用了一个电感从而比 Hartley 振荡器要更常被使用。然而，这两种振荡器都有两个缺点。第一，为了使发射极阻抗对于槽路的影响可以忽略不计，电容器和电感器的比值需要很大。第二，这样的振荡器只能提供单端输出。

单端口模型 虽然我们已通过反馈模型得到了 Colpitts 和 Hartley 振荡器拓扑，单端口的表示在这里也是可能的，后者提供了对这种电路的更深层次的了解。考虑图 7.8(a) 中的电路，其中 M_1 偏置在饱和区。注意其中没有任何一个节点是接地的。小信号的输入阻抗 Z_{in}

是什么呢？从图 7.8(b) 可以看到，我们可以把 I_{D1} 和通过 C_2 的电流的总和写为

$$-\frac{I_x}{C_1 s} g_m + \left(-\frac{I_x}{C_1 s} + V_x \right) C_2 s = I_x \quad (7.2)$$

其中栅源电容已经被忽略掉了。这样

$$\frac{V_x}{I_x} = \frac{g_m}{C_1 C_2 s^2} + \frac{1}{C_1 s} + \frac{1}{C_2 s} \quad (7.3)$$

对于 $s = j\omega$ ，这个阻抗包括一个其值为 $-g_m / (C_1 C_2 \omega^2)$ 负的实部，表明如果电感连接在漏栅之间，此电路可以发生振荡[图 7.8(c)]。

图 7.8 (a) 单端口有源电路，(b) 输入阻抗的计算，(c) 电感器并联接入端口

由于上面的推导假设电路中没有到地的连接，我们可以通过分别把三个终端接地来得到三种不同的拓扑结构，见图 7.9。

图 7.9 通过将图 7.8(c) 的一个电路终端接地得到的振荡器拓扑（未给出偏置电路）

7.3 电压控制（压控）振荡器

大多数射频振荡器的频率都是必须可以调节的。例如，在第 5 章中学习的下变换和上变换模块的前端中需要在多条信道中选择一个，因为一个特定的接收发送器在不同时间被分配不同的载频。这样，本振的频率必须可以以精确的间隔进行调节（例如 IS-54 标准中的 30kHz）。如果振荡器的输出频率可以由电压控制，这样的电路称作一个压控振荡器。虽然电流控制（流控）振荡器也是可行的，它们之所以没有在射频领域被广泛采用是因为很难通过电流来控制高 Q 值的储能元件的数值。

一个振荡器的频率如何变化呢？在 LC 实现中，槽路电容(的一部分)可以由反偏的二极管（“变容器”）提供，所以结两端的直流电压控制了谐振的频率。图 7.10 显示了采用了变容二极管的两种槽路结构。

图 7.10 加到槽路中的变容二极管

在相位噪声和锁相环的研究中，我们需要一个 VCO 的数学模型。我们把一个理想的 VCO 定义为一个可以产生周期输出信号的电路，其频率是控制电压的线性函数， $V_{\text{cont}} : \omega_{\text{out}} = \omega_{FR} + K_{VCO} V_{\text{cont}}$ ，其中 ω_{FR} 是“自振”频率，而 K_{VCO} 是 VCO 的“增益”（用 $rad/s/V$ 表示）。 ω_{FR} 的存在只是简单地指出，对于实际的 V_{cont} 的范围， ω_{out} 不会趋近 0。

换句话说， V_{cont} 产生一个 ω_{FR} 附近的一个变化。由于相位是频率对时间的一个积分，正弦 VCO 的输出可以表示为：

$$y(t) = A \cos \left(\omega_{FR} t + K_{VCO} \int_{-\infty}^t V_{\text{cont}} dt \right) \quad (7.4)$$

例如，如果 V_{cont} 为一个常数 V_0 ，则 $y(t) = A \cos \left[(\omega_{FR} + K_{VCO} V_0) t + \phi_0 \right]$ ，其中 ϕ_0 表示相位的初始值。这样，与预期一样，如果 V_{cont} 是常数，则频率只是简单地偏移了 $K_{VCO} V_{\text{cont}}$ 。

(7.4)式暗示了一个 VCO 实际上是一个调制器，而且在 3.2 节中得到的结果在这里同样适用。特别的是，对于正弦调制 $v_{\text{cont}}(t) = V_m \cos \omega_m t$ ，我们有

$$y(t) = A \cos \left(\omega_{FR} t + \frac{K_{VCO}}{\omega_m} V_m \sin \omega_m t \right) \quad (7.5)$$

上式表明 VCO 有抑制控制输入端出现的高频成分的倾向。而且，如果 $|K_{VCO} V_m / \omega_m| \ll 1$ 弧度，则窄带调频近似成立(第 3 章)，且输出频谱包含一个主要成分 ω_{FR} 和位于 $\omega_{FR} \pm \omega_m$ 的两个边带。

7.4 相位噪声

如其它的模拟电路一样，振荡器容易受到噪声的影响。由振荡器自身的器件或者由外界输入的噪声都可能影响到输出信号的频率和幅度。在多数情况下，幅度的扰动可以忽略或者无关紧要的，只有随机的频率偏移需要考虑。后者也可以看作是周期长短的随机偏移，或者在时间轴上过零点时间相对于其理想值的偏移。

对于一个名义上的周期正弦信号，我们可以写出 $x(t) = A \cos[\omega_c t + \phi_n(t)]$ ，其中 $\phi_n(t)$ 是一个表示周期变化的很小的随机额外相位。函数 $\phi_n(t)$ 被称作“相位噪声”。注意，对于 $|\phi_n(t)| \ll 1 \text{ rad}$ ，我们有 $x(t) \approx A \cos \omega_c t - A \phi_n(t) \sin \omega_c t$ ；也就是说， $\phi_n(t)$ 的频谱被变换到 $\pm \omega_c$ 。

在射频应用中，相位噪声通常在频域中表征。对于一个理想的工作在 ω_c 的正弦振荡器，其频谱假定为一个冲击脉冲，虽然对于一个实际的振荡器，它的频谱在载频附近表现为一个“裙”状(图 7.11)。为了量化相位噪声，我们考虑相对 ω_c 偏移 $\Delta\omega$ 处的一个单位带宽，计算这个带宽内的噪声的功率，将结果除以载频(平均)能量。例如，如果载频功率是 -2 dBm，且在偏移 1-MHz 处，1-kHz 带宽内测得的噪声功率是 -70 dBm，则相位噪声表示为 $-70 \text{ dBm} + 2 \text{ dBm} - 30 \text{ dB} = -98 \text{ dBc/Hz}$ ，其中 dBc 表示“相对载波的 dB 数”

图 7.11 理想和实际振荡器的输出频谱

7.4.1 相位噪声在射频通讯中的影响

为了理解相位噪声在射频系统中的重要性，考虑一个图 7.12 描述的一般性的接收发送器。其中一个本地振荡器同时为接收和发射路径提供载频。如果本振输出包括相位噪声，则下变换和上变换信号同时被破坏。这可以由图 7.13 看出。

图 7.12 通用的接收发送器前端

从图 7.13(a)中的理想情况中，我们可以注意到需要的信号与一个冲击脉冲发生卷积，从而变换到一个较低（和一个较高）的频率，形状并未改变。然而在实际中，有用信号可能被伴随着一个位于相邻信道的很大的干扰信号，而且本振具有不可忽视的相位噪声（图 7.13(b)）。当两个信号与本振输出进行混频时，下变换后的频带包含有两个重叠的频谱，有用信号受到干扰信号的带尾的强噪声的干扰。这种效应称作“互易混频(reciprocal mixing)”。

如图 7.13(c)所示，相位噪声在发射路径上的影响稍微有点不同。假设一个无噪声的接收器试图探测在 ω_2 点的微弱信号，而一个临近的，很强的发射机在 ω_1 处产生一个具有很大相位噪声的信号。则有用信号会被这个发射机的相位噪声的带尾所破坏。

这里的重点在于 ω_1 和 ω_2 的差别可能只有几十 kHz 之小，而每一个这样的频率在 900MHz 或者 1.9GHz 的附近。因此，本振的输出频谱必须十分尖锐。例如，在 IS-54 标准中，偏离中心频率 60 kHz 处每单位带宽的相位噪声功率必须在 -115 dBc/Hz 左右。

图 7.13 (a) 一个理想振荡器的下变频，(b) 互易 (reciprocal) 混频，(c) 相位噪声在发射器中的影响

为了达到一些有代表性的相位噪声标准，我们考虑图 7.14 所示的例子。有用信道宽度为 30 kHz，其幅度比距离 60 kHz 处的无用信号小 60dB。那么在 60 kHz 频率处的无用信号的相位噪声应该多低，才能使得期望的信号的信噪比大于 15 dB 呢？由于干扰信号在期望信道上产生的总的噪声功率为：

$$P_{n,\text{tot}} = \int_{f_L}^{f_H} S_n(f) df \quad (7.6)$$

其中 $S_n(f)$ 表示无用信道的相位噪声分布， f_L 和 f_H 分别为期望信道的下限和上限。为简单起见，我们假设 $S_n(f)$ 在此带宽内是常数且等于 S_0 ，从而得到 $P_{n,\text{tot}} = S_0(f_H - f_L)$ 。所以，

$$SNR = \frac{P_{\text{sig}}}{S_0(f_H - f_L)} \quad (7.7)$$

从而得到 $10\log(S_0 / P_{\text{sig}}) = -15 \text{ dB} - 10\log(f_H - f_L)$ 。由于干扰信号功率 P_{int} 比有用信号功率 P_{sig} 高 60 dB，我们有 $10\log(S_0 / P_{\text{int}}) = -15 \text{ dB} - 10\log(f_H - f_L) - 60 \text{ dB}$ 。结果是，在 60 kHz 偏移点处的相位噪声应该不超过 -120 dBc/Hz。注意，虽然 -120 dBc/Hz 这个数字看起来很小，当它在 30-kHz 的带宽内做一个积分时，结果会增大到 -70 dBc。如果相位噪声在 f_L 和 f_H 之间变化很大，那么相位噪声应该在几个不同的频率偏移点给定，且要相应地进行 (7.6) 式中的积分。

图 7.14 由于相位噪声造成的信号质量变差的计算

本振的相位噪声还会破坏载频的相位中携带的信息。例如，如果一个混频器由一个具有噪声的本振驱动，且对 QPSK 波形进行的下变换，得到的星座图如图 7.15[4]。这样，接收器的误码率可能会有很大的增加。

图 7.15 相位噪声对 QPSK 信号星座的影响

7.4.2 振荡器的 Q 值

LC 振荡器的相位噪声依通常赖于它们的 Q 值。直观上来讲，我们可以预计谐振槽路的 Q 值越高，谐振峰就越尖锐，相位噪声的“裙摆”也就越低。在振荡器的背景下来考察三种

不同定义的 Q 值是有启发性的。

Q 值的基本的物理定义是 $2\pi \times$ (每周期存储的能量/消耗的能量)。对于一个 LC 谐振槽路，Q 值表征能量在电感和电容之间来回传输过程损失的多少。

谐振电路通常都表现出带通传输的功能。Q 值也可以定义为频响特性幅度的“陡峭程度”。更具体来讲，如图 7.16 所示，Q 值定义为谐振频率被双边-3-dB 带宽所除。对于简单的 LC 谐振回路，这两种定义得到同样的结果。

图 7.16 Q 值的一种定义

Q 值的另一种定义在振荡器中被证明更加有用，如图 7.17 所示。这里的电路被当作一个反馈系统，同时考察谐振时开环传输函数的相位 $\phi(\omega)$ 。Q 值然后定义为

$$Q = \frac{\omega_0}{2} \left| \frac{d\phi}{d\omega} \right| \quad (7.8)$$

这个 Q 值被称作“开环 Q 值”。如果我们考虑到对于稳定的振荡，环路总的相移应该为 0，那么这种定义有一种有趣的解释。现在假设振荡频率相对 ω_0 有一个很小的变化。则如果相位的斜率足够大，就会产生一个很大的相移。这样就违反了振荡的条件从而迫使频率回到 ω_0 。换句话说，开环 Q 值衡量的是闭环系统抗拒振荡频率变化的能力。

图 7.17 基于开路相位斜率的 Q 值定义

现在我们将(7.8)式应用到图 7.18(a)的 Colpitts 振荡器中，看看 Q 值与电路元件的关系。我们先在 M_1 的漏端断开环路，向槽路注入电流，然后计算 M_1 中的电流。根据 7.2 节的分析且假定槽路的 Q 值由经过阻抗变换后的 $1/g_{m1}$ 所限制，我们由图 7.18(b)的等效电路可以写出

$$\frac{V_x}{I_{in}} = Ls \left\| \frac{1}{C_{eq}s} \right\| R_p \quad (7.9)$$

其中 $C_{eq} = C_1 C_2 / (C_1 + C_2)$ 且 $R_p \approx (1 + C_1/C_2)^2 / g_{m1}$ 。如果 $(C_1 + C_2)s \gg g_m$ ，则加在 M_1 源极的电流近似等于 $V_x [C_2 / (C_1 + C_2)]$ ，且得到的漏电流

$I_{out} = -g_m V_x [C_2 / (C_1 + C_2)]$ 。从而开环传输函数为：

$$\frac{I_{out}}{I_{in}}(s) = -g_m \frac{C_2}{C_1 + C_2} \left(Ls \left\| \frac{1}{C_{eq}s} \right\| R_p \right) \quad (7.10)$$

圆括号中的项给出了相移

$$\phi(\omega) = \frac{\pi}{2} - \tan^{-1} \frac{L\omega}{R_p(1 - LC_{eq}\omega^2)} \quad (7.11)$$

因此，在 $\omega = \omega_0 = 1/\sqrt{LC_{eq}}$ 处，

$$\left. \frac{d\phi}{d\omega} \right|_{\omega=\omega_0} = -2C_{eq}R_p \quad (7.12)$$

这样， $Q = R_p C_{eq} \omega_0 = R_p / (L\omega_0)$ ，与前两种定义得到的结果一样。注意我们这里忽略了晶体管本身的相移。

图 7.18 (a) 开环 Colpitts 振荡器，(b) 漏极网络的等效电路

在这个例子中，我么假设了 R_p 仅由晶体管的输入电阻决定。实际中，电感和电容都有损耗，使得 Q 值进一步降低。总的 Q 值有时被称作是“加负载后的 Q 值”以便与槽路本身的 Q 值区分，且强调了晶体管以及后级缓冲器所引入的电阻性成分。

7.4.3 相位噪声机制

振荡器中的相位噪声已经被广泛地研究了 [5,6,7]。诸如非线性和电路参数的周期变化等效效应使得相位噪声的分析变得十分困难。虽然如此，对于简单的 LC 振荡器，线性近似被合理地利用，结果产生的对相位噪声估计的误差可以接受[6,7]。对振荡器更严格的分析目前仍然是研究领域的一个活跃课题。

根据噪声注入的途径，振荡器的相位噪声主要通过两种机制产生。回想一下一个 VCO 包括了一个（反馈）振荡信号路径和一个频率控制路径。如图 7.19 所示，出现在这两种路径上的噪声 $x(t)$ 会产生可区分的不同的效果。我们分别考查每一种情况。

图 7.19 在 (a) 信号路径，(b) 控制路径中的相位噪声

信号路径上的噪声 如图 7.19(a)所示的电路，考虑噪声注入到振荡信号路径的情况。 $y(t)$ 如何受 $x(t)$ 影响呢？我们可以把开环电路用线性传输函数 $H(s)$ 来表达，

$$\frac{Y(s)}{X(s)} = \frac{H(s)}{1 - H(s)} \quad (7.13)$$

在振荡频率的附近， $\omega = \omega_0 + \Delta\omega$ ，我们可以用泰勒展开式的前两项来近似表达

$H(j\omega)$ ，

$$H(j\omega) \approx H(j\omega_0) + \Delta\omega \frac{dH}{d\omega} \quad (7.14)$$

由于 $H(j\omega_0) = +1$ 且一般 $|\Delta\omega dH/d\omega| \ll 1$ ，(7.13)可以化简为：

$$\frac{Y}{X}(\omega_0 + \Delta\omega) \approx \frac{-1}{\Delta\omega \frac{dH}{d\omega}} \quad (7.15)$$

上式暗示了在振荡器输出端，位于 $\omega = \omega_0 + \Delta\omega$ 处的噪声成分被乘上了一个系数 $-(\Delta\omega dH/d\omega)^{-1}$ 。换句话说，如图 7.20 所示，噪声频谱被下式所整形[25]

$$\left| \frac{Y}{X} [j(\omega_0 + \Delta\omega)] \right|^2 = \frac{1}{(\Delta\omega)^2 \left| \frac{dH}{d\omega} \right|^2} \quad (7.16)$$

图 7.20 振荡器中的噪声成形

我们现在考虑 $|dH/d\omega|^2$ 这一项。用极坐标的形式表示 $H(\omega)$ ， $H(\omega) = |H| \exp(j\phi)$ ，我们有，

$$\frac{dH}{d\omega} = \left(\frac{d|H|}{d\omega} + j|H| \frac{d\phi}{d\omega} \right) \exp(j\phi) \quad (7.17)$$

从而，

$$\left| \frac{dH}{d\omega} \right|^2 = \left| \frac{d|H|}{d\omega} \right|^2 + \left| \frac{d\phi}{d\omega} \right|^2 |H|^2 \quad (7.18)$$

(7.18)式可以用来得到 Q 值的一个一般性定义[25]，但我们这里仅把我们的研究限制在简单的 LC 振荡器。我们可以得到两个结论。第一，对于一个 LC 振荡器，在谐振频率附近， $|d|H|/d\omega|^2$ 远小于 $|d\phi/d\omega|^2$ 。第二，在稳定振荡的情况下， $|H|$ 的值接近于 1。这样，

$|dH/d\omega|^2 \approx |d\phi/d\omega|^2$ ，从而(7.16)可以写为

$$\left| \frac{Y}{X} (j\omega) \right|^2 = \frac{\omega_0^2}{4\Delta\omega^2} \frac{1}{\frac{\omega_0}{2} \left(\frac{d\phi}{d\omega} \right)^2} \quad (7.19)$$

从(7.8)我们可以得到

$$\left| \frac{Y}{X} (j\omega) \right|^2 = \frac{1}{4Q^2} \left(\frac{\omega_0}{\Delta\omega} \right)^2 \quad (7.20)$$

此式被称作“Leeson 方程” [6]，它揭示了输出噪声与槽路 Q 值，中心频率以及偏移频率的关系。注意(7.20)同时包含了噪声在载频的幅度和相位上的影响。一般来说，相位噪声是(7.20)给出值的一半。

值得注意的是，总的相位噪声功率与载频的关系依赖于另外两个参数，那就是由器件产生的噪声，也就是 $X(j\omega)$ 的幅度，还有振荡的幅度 V_S 。例如，在图 7.7 的振荡器中，

由 Q_1 引入的噪声来源于其基极电阻和集电极电流 I_{C1} ，而其幅度由集电极电流摆幅和等效并联电阻 R_p 决定。由于噪声电流与 $\sqrt{I_{C1}}$ 成正比，而 V_S 与 $I_{C1}R_p$ 成正比，所以一般设计中都在不使晶体管出现饱和的情况下尽量增大 I_{C1} 。

从上面的讨论中，我们可以推出 LC 振荡器的一个有趣的特性。当槽路的 Q 值增加时，有三个方面的性能同时得到提高：(1) 噪声整形函数(7.20)变得更陡峭；(2) 功率耗散减小；(3) 由有源器件引入的噪声减小。为了解后两点，注意到槽路的等效并联电阻 $R_p \propto Q$ ，所以电压摆幅 $V_S \propto I_{C1} \cdot Q$ 。这样，高的 Q 值使得低的 I_{C1} 可以达到一样的电压摆幅，从而可以降低功耗。而且，当 I_{C1} 减小时，由晶体管注入的噪声电流也随之减小。基于这些原因，人们为了提高电感的 Q 值做出了很大的努力[10,11]。

我们对信号路径的噪声的分析一直是假设一个线性反馈系统的。如 7.1 节提到的，振荡器通常都有幅度限制，从而表现为非线性，进而会将噪声分量“折叠”。如 7.21 所示，当对幅度的奇次非线性导致了在 ω_n 频率下注入的噪声成分与载频进行交调时，这个折叠效应会起作用，从而产生一个新的频率分量 $2\omega_0 - \omega_n$ [25]。为了表示由于非线性引起的噪声折叠效应，Leeson 方程可以乘上一个因子[8,9]

$$\left| \frac{Y}{X}(j\omega) \right|^2 = \frac{A}{4Q^2} \left(\frac{\omega_0}{\Delta\omega} \right)^2 \quad (7.21)$$

其中 A 是实际的小信号环路增益。这个关系表明环路增益必须在保证可靠的启动的情况下，尽量接近于 1。A 的典型值在 2 到 3 之间。

图 7.21 振荡器中的噪声折叠

控制路径上的噪声 在上述讨论的机理中，注入到信号路径的噪声简单地与载频进行混频。作为一个对照，注入到控制路径的噪声通过改变振荡器的物理属性来影响频率。例如，如果用一个变容器用来调谐 VCO，则通过变容二极管两端的直流电压上的噪声会改变槽路电容，从而改变谐振频率。如果把此效应当作是一个模拟调频过程，可以认为它将控制路径的低频噪声成分变换到载频附近[图 7.22(a)]。

为了量化这个调频噪声机制，如图 7.22(b)所示，我们把每单位带宽的噪声表示为一个具有相同的平均功率的正弦波： $V_m \cos \omega_m t$ 。用 K_{VCO} 表示 VCO 的增益，且采用窄带调频近似，我们可以得到

$$v_{out}(t) \approx A_0 \cos \omega_0 t + \frac{A_0 V_m K_{VCO}}{2\omega_m} [\cos(\omega_0 + \omega_m)t - \cos(\omega_0 - \omega_m)t] \quad (7.22)$$

这样，在 $\omega_0 \pm \omega_m$ 处噪声功率相对载波功率等于 $(K_{VCO}/\omega_m)^2 V_m^2 / 4$ 。实际中， K_{VCO} 与

载频成正比，这是因为对于一个给定的控制电压范围，调谐范围相对于中心频率的百分比应该是个常数，这样才能补偿工艺和温度的变化带来的影响。注意，当 ω_m 减小时，这类噪声的影响会愈发显著，使得控制路径上的 $1/f$ 噪声尤其起有害作用。

图 7.22 (a) VCO 频率被控制线上的噪声所调制，(b) 用一个正弦波来近似噪声

值得一提的是，调频噪声机制不仅仅是由 VCO 的控制路径产生，而是由任何可能影响振荡频率的地方。一个有趣的例子是，在某些振荡器中采用了分立元件电感，其电感值可能会随机械振动而变化。另一个例子，图 7.7 中 Q_1 的集电极电流的噪声使得基极-发射极扩散电容发生变化，从而调制了谐振频率。

本节阐述的机制预测了当 $\Delta\omega$ 增加时，相位噪声不断会减小。然而实际中，在高的频率偏移下，环路不会对器件噪声进行整形，从而噪声会达到一个相对平坦的平台。

7.4.4 噪声—功率折衷关系

正如其它的模拟电路，振荡器表现出功耗与噪声的折衷中。直观上，我们注意到如果 N 个同样的振荡器的输出同相位进行叠加(图 7.23)，则总的载波功率会增加到 N^2 倍，而噪声功率只增加 N 倍（假设不同振荡器的噪声源是不相干的）。这样，相位噪声（相对于载波）减小一个因子 N ，其代价是功耗增加相应的比例。

图 7.23 N 个振荡器的输出电压的叠加

我们对相位噪声的研究表明，下列的参数必须考虑在内：中心频率 ω_0 ，功耗 P ，频偏 $\Delta\omega$ 。

这样，不同振荡器的相位噪声功率必须用 $(\omega_0/\Delta\omega)^2/P$ 进行归一化以便进行客观的比较。

7.4.5 分频与倍频对相位噪声的影响

由于频率和相位之间的关系为一个线性算子，将频率除以一个因子 N 是等价于将相位除以 N 。对于一个名义上的周期正弦信号， $x(t) = A \cos[\omega_c t + \phi_n(t)]$ ，其中 $\phi_n(t)$ 表示相位噪声，一个分频器简单地将总的相位除以 N

$$x_1/N = A \cos \left[\frac{\omega}{N} t + \frac{\phi_n(t)}{N} \right] \quad (7.23)$$

这里忽略了由分频器引入的相位噪声。上式表明，对于一个给定频偏，相位噪声的幅度除了一个 N 。对于窄带调频近似，相位噪声功率除以 N^2 。

有意思的是，上面这个趋势甚至当一个频分器没有被用到时一样被观察到。比如说，在图 7.7 中的 LC 振荡器，假若总的电感和电容被增加 N 倍，则振荡频率被减少同样的倍数。而 Q 值，噪声源，以及幅度相对保持常数。这样，从 (7.20)，在给定的偏移下的相位噪声功率降低了 N^2 倍。

类似的分析表明，倍频器使得相位噪声幅度提高同样的倍数。

7.4.6 振荡器的牵引与推动

在我们对相位噪声的研究中，我们已经策略性地假设了注入到信号路径的噪声幅度远小于载波幅度，从而得到一个振荡器的噪声整形函数。在噪声成分与载频很接近且幅度与载波可比拟的情况下，会出现一种有趣的现象（图 7.24）。当噪声幅度增加时，载波频率会向噪声频率偏移，最终“锁定”在噪声频率。这种效应被称作“注入牵引”，在文献[12,13]中有具体描述。

图 7.24 当噪声幅度增加时，振荡器的注入牵引

在一个接收发送器环境中，各种源都可能引入振荡器的牵引。例如如第 5 章中描述的，功放的输出可能耦合到本振。另一个注入牵引的例子出现在接收路径，当有用信号伴随着一个很强的干扰信号时（图 7.25）。如果干扰频率与本振频率很接近，通过混频器的耦合会将 ω_{LO} 拉到 ω_{int} 。这样，VCO 后必须有一个逆向隔离性能很好的缓冲级。然而如第 6 章所述，这样一个缓冲级的噪声会增加混频器的噪声系数。

图 7.25 由于大的干扰源的注入牵引

振荡器的另一种类型的牵引如图 7.26 的 GFSK 调制器所示，这种拓扑是用于 DECT 接收发送器中的[14]。这里，VCO 首先被置于一个反馈环路（第 8 章）中来稳定其输出频率。接着，VCO 的控制电压切换到基带信号，使得 VCO 成为一个调频器。但这里的开环连接使得 VCO 的频率对于负载阻抗的变化很敏感（“负载牵引”）。特别是，为了节省功耗，功率放大器会周期性的关闭和开启时，VCO 的中心频率会有相当大程度的变化，因此在 VCO 与 PA 之间需要很强的隔离级[14]。

图 7.26 由于阻抗变化的负载牵引

一般来说，射频振荡器对电源电压波动的抑制都很差。例如图 7.10，如果电源电压发生变化，变容器两端的反向电压同时发生变化，从改变振荡频率。在射频设计中，这称作“电源推动”。电源推动的一个重要的例子发生在便携式接收发送器的功放开启和关闭的过程中。由于电池的有限输出阻抗，电源电压可能会有几百毫伏的变化。

7.5 双极型与 CMOS LC 振荡器

射频系统对相位噪声严格的要求通常需要振荡器的拓扑结构中包含无源的谐振器。谐振器可由带宽非常窄的 SAW、晶体滤波器、传输线或者简单的 LC 振荡器构成。其相位噪

声则一般按以上次序逐渐升高。在许多应用中，Q 值大于 20 的 LC 振荡器被证明是性能足够的。然而回想一下 7.4.3 节所述，相对的相位噪声还依赖于噪声源和振荡幅度。为了使前者最小化，低噪声的振荡器在信号路径上通常只有几个有源器件。

正如其它的模拟电路一样，当电源电压按比例缩小时，VCO 的设计也变得更加困难。在低电源电压下，电压摆幅更小，从而使相对相位噪声更大。而且，变容二极管的反偏电压变化范围也更小，从而限制了调谐的范围。值得注意的是，在高频下，由于寄生电容的影响十分重要，变容二极管只能调节槽路电容的一部分。

LC 振荡器的有限的调谐范围既是一个优点又是一个障碍！一方面，由于 K_{VCO} 相对较小，频率对控制路径的噪声不敏感。另一方面，为了保证在工艺偏差的情况下频率仍然能在要求的频带内，需要其它的调节频率的方法（至少在组装过程中）。在许多射频系统中，可以通过在变容二极管两端并联一个小的用机械方法能校正的电容来达到，虽然其调节范围仍然需要足够宽以至于可以覆盖温度变化带来的偏差。然而，目前的人们朝向更高度集成的趋势要求通过达到足够的调谐范围来消除这样的调节。

7.5.1 负跨导 (Gm) 振荡器

如 7.2 节解释的，由晶体管的集电极（或漏极）到其基极（或源极）的反馈需要通过一个阻抗变换器，从而避免槽路过载。在图 7.7 的 Colpitts 和 Hartley 振荡器中，阻抗经过一个无源网络进行变换。现在考虑图 7.27(a) 的振荡器，其中一个有源的缓冲器 B_1 插在集电极与发射极之间，对槽路表现为一个高阻。如图 7.27(b) 所示，此缓冲器可以由一个射极（或源极）跟随器来实现。注意为了与 Q_2 基极在直流下同电位， Q_1 的基极接 V_{CC} 。（假定这两个三极管等同）这个电路也可以利用两个电感来实现差分结构[图 7.27(c)]。然而，如果电感做在片外，由于键合线带来的寄生电感使得这两个槽路之间的相位失配很难做到很小。

图 7.27 (a) 在 Colpitts 振荡器的反馈回路中加入有源缓冲器，(b) 用射极跟随器来实现缓冲器，(c) (b) 的差分实现。

图 7.27 (c) 的反馈振荡器也可以看作一个单端口实现。如图 7.28，计算从 Q_1 和 Q_2 集电极看进去的阻抗，我们注意到这里正反馈使得 $R_{in} = -2/g_m$ 。从而，如果 $|R_{in}|$ 小于或等于槽路等效并联电阻，则电路会发生振荡。这种结构被称作“负-Gm 振荡器”。

图 7.28 计算“交叉耦合对”输入阻抗的电路

除了会引入噪声，图 7.27(c) 的 Q_1 和 Q_2 在 X 和 Y 点的峰值摆幅接近 400 mV 时会发生深度的饱和。为了解决这个问题，可以在反馈路径上插入一个电容分压器(图 7.29)，以便容许更大的摆幅和在 X 和 Y 节点的共模电平比在 Q_1 和 Q_2 基极的共模电平更高。

图 7.29 振荡器负反馈回路中的电容分压

负-Gm 振荡器也可以用 MOS 工艺来实现。图 7.30 就是三个例子[15,16,17]。由于大摆幅的在这里不会造成问题，所以反馈路径上不需要电容分压。图 7.30 (c) 的电路只需要一个电感就实现了差分结构，但其代价是由于 I_1 和 I_2 使得噪声更高。

图 7.30 CMOS 振荡器，(a) 相对于 V_{DD} 的输出，(b) 相对于地的输出，(c) 只有一个电感器的差分对

调谐问题 为了在上述电路中控制振荡频率，槽路中应该加入变容二极管。图 7.31(a) 显示了一个例子，其中 D_1 和 D_2 的结电容由 V_{cont} 控制。在典型的 CMOS 工艺中， D_1 和 D_2 的实现如图 7.31(b)。为了理解这个电路的调谐范围的限制，我们假设一些典型值 $L_1 = L_2 = 10 \text{ nH}$ ， $(W/L)_1 = (W/L)_2 = 100 \mu\text{m}/0.6 \mu\text{m}$ ， $I_{DD} = 2 \text{ mA}$ ，且 $V_{DD} = 2.7 \text{ V}$ ，中心频率为 1.8 GHz。 M_1 和 M_2 的尺寸以及 I_{DD} 的选取是为了对 $Q \approx 4$ 时能达到足够的环路增益。

这个电路的调谐范围是多大呢？我们三个观察结论。第一，对于三层金属工艺的集成电感， L_1 和 L_2 的分布寄生电容分别为 5 pF 左右，可以看作是二个 0.25 pF 的电容分别接在电感两端与地之间。第二，由 M_1 和 M_2 以及典型的输出缓冲级在 X 和 Y 点引入的总电容分别约为 0.25 pF。第三，伴随二极管 D_1 和 D_2 有到 p-衬底之间的大约为 0.1 pF 的电容。

图 7.31 (a) 增加变容器以控制频率，(b) 用 CMOS 工艺实现的变容二极管

基于上述三点，我们可以计算出每个槽路总的固定电容约为 0.6 pF。对于 $f=1.8 \text{ GHz}$ 和 $L=10 \text{ nH}$ ，仅剩下 0.18 pF 的余量给变容二极管。那么这个电容变化的范围是多大呢？记住，为了降低相对相位噪声，漏电压摆幅较大。这样， V_X 和 V_Y 的最小值一般小于 0.5 V，说明 V_{cont} 应该低于 1.2 V 以避免 D_1 和 D_2 正偏。而且，如果振荡器工作在一个频率综合器内，由前一级决定的 V_{cont} 的最小值不会小于几百毫伏。 V_{cont} 这样的范围限制一般使得变容二极管的电容 C_j 变化大约 50%。

因为在这个例子中，每一个变容器的结电容只构成了总的槽路电容的 23%，在 C_j 上的 50% 变化相当于振荡频率的 10% 的调整。

调谐范围与相对相位噪声之间存在折衷关系。如果图 7.31(a) 中 X 和 Y 点的电压摆幅变小，则 V_{cont} 可以用更高的值，从而可以更大程度调节频率。而且，如果 L_1 和 L_2 降低，则总的槽路电容，因而变容二极管贡献的电容都能增加，使得调谐范围更大，其代价是电压摆幅降低（或者偏置电流提高）。

7.5.2 插值振荡器

直到目前为止，VCO 研究中的基本假设是频率只能由调节槽路电容（或电感）来改变。插值振荡器采用了多个谐振器，提供了另外一种调谐机制。

图 7.32 所示为一个具有两个传输函数 $H_1(s)$ 和 $H_2(s)$ 的反馈振荡系统， $H_1(s)$ 和 $H_2(s)$ 的输出分别乘上可变因子 α_1 和 α_2 ，然后相加。总的开环传输函数因此为 $H(s) = \alpha_1 H_1(s) + \alpha_2 H_2(s)$ 。为了使整个系统发生振荡，它必须等于 +1。在极端情况下， $\alpha_1 = 0$ 或 $\alpha_2 = 0$ ，振荡频率 ω_c 仅由 $H_1(s)$ 或 $H_2(s)$ 决定。对于介于极端情况之间的 α_1 和 α_2 ， ω_c 可以通过在其上界和下界中“插值”得到。显然，只有在 $H_1(s) - 1 = 0$ 的根有“秩序”地从一个极端向另一个极端移动的情况下，这样的插值才能实现。

图 7.32 插值振荡器

为了理解插值的原理和它的局限性，我们考虑图 7.33 所示的电路。这个电路是文献[18]中描述的振荡器的一个概念性描述。在此电路中， R_1 和 R_2 表示每个槽路的等效并联电阻。为了使系统发生振荡，开环传输函数必须为 1：

$$\alpha_1 \frac{G_{m1} R_1 L_1 s}{R_1 L_1 C_1 s^2 + L_1 s + R_1} + \alpha_2 \frac{G_{m2} R_2 L_2 s}{R_2 L_2 C_2 s^2 + L_2 s + R_2} - 1 = 0 \quad (7.24)$$

如果(7.24)在右半平面或虚轴上有唯一的一对共轭的根，则电路在一个单一的频率上发生振荡。

图 7.33 采用两个槽路的插值振荡器

如果两个槽路的谐振频率 ω_1 和 ω_2 距离很远，会发生什么事呢？图 7.34 绘出了这样一个振荡器的开环增益和相位[18]，显示了在两个频率处总相位为零且增益大于 1。因此，这个电路不止在一个频率点发生振荡。为了避免这种情况的发生，可以证明 ω_1 和 ω_2 的最大距离为 $(\omega_1 + \omega_2)/(2Q)$ ，其中假设两个槽路具有相同的 Q 值[18]。有趣的是，插值振荡器在相位噪声和调谐范围之间也存在着折衷。

图 7.34 插值振荡器的开环增益和相位

在插值振荡器中主要关心的一点在于信号路径上附加的有源器件，尤其是加法器带来的噪声。即便如此，电容调谐与插值的结合仍可以提供更宽的调谐范围。

7.6 单片集成电感

将电感与射频电路的其它部分集成在同一个硅片上,可以避免外部连线,从而解决了诸如电磁耦合以及压焊块和键合线的寄生参数。而且,将电感加到设计者的器件库中可以带来新的技术,使得射频设计中的折衷选择得到缓和。

片上集成电感中最重要的参数是它的 Q 值、自谐振频率、以及面积。这些参数都强烈依赖于版图和集成电路工艺。电感通常做成螺旋结构(图 7.35)。对于给定的电感值和金属线宽度,圆形电感具有更小的金属电阻。圆形可以用六边形或者八边形来近似[19]。

图 7.35 矩形螺旋电感

螺旋电感的值可以通过许多精确的[20]或者近似的[21,22]方法来得到。一个对于 5nH 到 50nH 电感误差小于 10%的经验公式见[17],且对于方形可以化简为下式,

$$L \approx 1.3 \times 10^{-7} \frac{A_m^{5/3}}{A_{tot}^{1/6} W^{1.75} (W + G)^{0.25}} \quad (7.25)$$

其中 A_m 为金属面积(图 7.35 的阴影部分面积), A_{tot} 为电感总面积(图 7.25 中约为 S^2), W 是线宽, G 是线间距离,所有的单位为米制。

电感 Q 值的预测相对更加困难。如图 7.36,片上电感的损耗来源于三个方面:金属线电阻,衬底容性耦合,衬底磁性耦合[23]。后二者由衬底的电阻率和电感面积决定,他们限制了电感在高频的 Q 值。注意在 CMOS 工艺的重掺杂衬底下,由于衬底电阻以与电感并联的形式出现,磁性耦合更加严重。

图 7.36 在单片电感中的损耗机制, (a) 电阻性, (b) 电容性, (c) 磁性

衬底的容性耦合效应可以通过在电感下方加入一个接地屏蔽得到减小(图 7.37)。由于一个连续的导电平面会增加磁性耦合,这个平面必须在垂直于电流的方向有规律的切断。在螺旋电感中,内圈的金属只贡献很小的电感,但受到所有损耗机制的影响。所以,去掉电感的内部 4—5 圈会稍微增加 Q 值。

图 7.37 粉碎的屏蔽板来减少电容性的损耗机制

对于一个给定的频率,电感的自谐振频率 f_{SR} 限制了电感值和电感面积的上限。螺旋电感的分布式结构使得 f_{SR} 的估算非常困难,通常需要电磁场求解器或者实际测量来确定。注意,即使工作频率低于 f_{SR} ,每个电感也会有相当大程度的对地的电容。

尽管目前已经进行了广泛的研究,在标准 CMOS 工艺中,电感的 Q 值仍然局限在 5 以下。片上电感的研究会不断地持续下去。

7.7 无谐振器的压控振荡器 (VCO)

振荡器的 Barkhausen 准则不需要电路中的谐振器也同样可以满足。如果开环电路在零相位频率下具有足够高的增益，振荡就会发生。同时，如果开环输入/输出特性具备磁滞效应，则即使小信号相移不足，电路也可以发生振荡。

第一种类型的一个例子就是环振。如图 7.38 所示，如果奇数个反相器构成的 M 级串成反馈的形式，则电路会发生振荡，其周期为 $2MT_d$ ，其中 T_d 是每级在扇出数为 1 的情况下的延迟。这个振荡可以看作是发生在总相移为零且闭环增益为 1 的频率下。不同的用在环振中的增益级的实现以及调频方法在文献[24]中可以发现。

图 7.38 环型振荡器

图 7.39 为采用磁滞特性的振荡器的例子。这种振荡器称作“松弛”振荡器。即使只有 C_1 一个储能元件，它也可以发生振荡，这是由于 M_1 和 M_2 构成的正反馈使得在每周期结束的时候发生急速的开关，而在大部分时间，反馈环是断开的。然而，在高频下，磁滞特性消失且只有在 X 和 Y 的极点能提供足够的相移时，振荡才会发生[25]。有趣的是，从交叉耦合对的 X-Y 端看进去的小信号阻抗近似等于 $-2/g_m - 1/C_1s$ 。也就是一个负电阻串连一个负电容。在高频下，前一项抵消了 R_1 和 R_2 的损耗，而后一项与 X 和 Y 点之间的电容发生谐振。

图 7.39 松弛振荡器

无谐振器的振荡器在射频设计中并不常用。这是由于它们的开环 Q 值只接近于 1，而且它们的信号路径包含了很多有源和无源的噪声元件。例如，在一个三级差分环振中，开环 Q 值近似等于 1.3，而 9 个晶体管（包括数个尾电流源）和 6 个负载电阻在载波上增加了噪声[25]。

7.8 正交信号产生

在第 5 章对接收发送器的学习中，我们注意到了无论是在接收路径还是发射路径，都需要正交的本振信号。在本节，我们介绍如何产生相差为 90° 的信号。

7.8.1 RC-CR 网络

一个简单的正交技术是利用 RC-CR 网络使信号分别移相 $\pm 45^\circ$ （图 7.40）。如 5.2.3 节所述， V_{out1} 与 V_{out2} 之间的相位差对所有的频率都是 90° ，但输出幅度只有在 $\omega = 1/(RC)$ 时才相等。所以，如果 RC 的绝对值随温度和工艺而变化，等幅的正交信号频率也随之发生变化。在本振路径，信息只存在在过零点时刻，而幅度可以通过“限幅”电路，例如差分对，

来达到相等。然而限幅措施在 GHz 的电路中变得很困难，除非通过几级电路的级联。然而在这样的情况下，在 I 和 Q 信号路径中限幅级链的相位（和幅度）的失配变得很严重。

图 7.40 应用 RC-CR 电路的正交网络

此外，限幅级会导致“AM-PM 转换”（调幅—调相转换）。一般来说，一个有限带宽的非线性电路会产生信号延时，其延迟量与输入信号的摆率（slew rate）有关[26]。为了理解这个效应，考虑图 7.41 的简单例子，其中电容在此表示有限带宽，忽略其它的寄生参数。我们考查电路的小信号和大信号相移。

图 7.41 具有有限带宽的限幅器

当频率为 ω 小信号正弦信号加在输入端，输出的差分电流也接近正弦信号，从而经历了相差 $|\theta_1| = \tan^{-1}(R_1 C_1 \omega)$ 。现在假设输入为一个大幅度的正弦信号，从而 M_1 和 M_2 在 V_{in} 的过零点处迅速发生转换。在这种情况下，差分输出电流近似为一个方波，其延时为 $R_1 C_1 \ln 2$ ，即相移为 $|\theta_2| = R_1 C_1 \omega \ln 2$ 。

从上述分析中可以得到，根据输入摆率的不同，相移从 $R_1 C_1 \omega \ln 2$ 到 $\tan^{-1}(R_1 C_1 \omega)$ 之间发生变化。如果输入信号为一个调幅信号，这种效应尤其值得注意。在限幅过程中，幅度变化扰乱了相移，我们称之为 AM-PM 转换。在正交信号产生这个前提下，关键的地方是如果限幅级无法提供足够的带宽，幅度的不同会转换到相位的不平衡。

在图 7.40 的电路中，电阻和电容的失配导致输出相差 ϕ 偏离 90° 。假设电阻相对失配为 α ，电容相对失配为 β ，我们可以将 ϕ 在 $\omega = 1/(RC)$ 表达为

$$\phi = \frac{\pi}{2} - \left[\tan^{-1} R(1+\alpha)C(1+\beta)\omega - \tan^{-1} RC\omega \right] \quad (7.26)$$

$$= \frac{\pi}{2} - \tan^{-1} \frac{RC\omega(1+\alpha)(1+\beta) - RC\omega}{1 + RC\omega(1+\alpha)(1+\beta)RC\omega} \quad (7.27)$$

$$\approx \frac{\pi}{2} - \tan^{-1} \frac{\alpha + \beta}{2} \quad (7.28)$$

$$\approx \frac{\pi}{2} - \frac{\alpha + \beta}{2} \quad (7.29)$$

例如，如果 $\alpha = \beta = 1\%$ ，则相位失衡为 $3.6/2\pi \approx 0.6^\circ$ 。

有意思的是，图 7.40 的相位与幅度的失衡与 V_{out1} 和 V_{out2} 到地之间的附加负载电容无关。这是由于这个电容只影响电路的极点，而相差由其上方路径的零点决定。然而，输出点之间的电容通路却会产生相位误差，因此在画版图的时候需要特别注意。模拟显示，如果这个电容为 C 的 2%，则会出现 1° 的相位失衡。负载电容的失配也会对相位误差产生影响。

图 7.40 所示正交电路中的另一个问题在于 v_{in} 的谐波成分。假设 $v_{in}(t) = A_1 \cos \omega t + A_n \cos n\omega t$ 。为了使 $v_{in}(t)$ 移相 90° ，我们用 $t - T/4$ 代替 t ，其中 $T = 2\pi/\omega$ 。这样， $v_{in}(t - T/4) = A_1 \cos(\omega t - \pi/2) + A_n \cos(n\omega t - n\pi/2)$ ，说明第 n 次谐波应该被移相 $T = n\pi/2$ 。而另一方面，在 RC-CR 网络中，对于所有的频率，相移都是 90° ，这样在 V_{out1} 和 V_{out2} 之间就存在着相位失衡。而且，谐波的幅度也经历了两路信号不同的增益，使得输出端出现幅度的失配。

为了量化二次谐波的效应，我们将输出表达为

$$\begin{aligned} v_{out1}(t) &= \frac{A_1}{\sqrt{R^2 C^2 \omega^2 + 1}} \cos[\omega t - \tan^{-1}(RC\omega)] \\ &= \frac{A_2}{\sqrt{4R^2 C^2 \omega^2 + 1}} \cos[2\omega t - \tan^{-1}(2RC\omega)] \end{aligned} \quad (7.30)$$

$$\begin{aligned} v_{out2}(t) &= \frac{A_1 RC\omega}{\sqrt{R^2 C^2 \omega^2 + 1}} \cos\left[\omega t - \tan^{-1}(RC\omega) + \frac{\pi}{2}\right] \\ &= \frac{2A_2 RC\omega}{\sqrt{4R^2 C^2 \omega^2 + 1}} \cos\left[2\omega t - \tan^{-1}(2RC\omega) + \frac{\pi}{2}\right] \end{aligned} \quad (7.31)$$

对于 $RC = \omega^{-1}$ ，我们有

$$v_{out1} = \frac{A_1}{\sqrt{2}} \cos\left(\omega t - \frac{\pi}{4}\right) + \frac{A_2}{\sqrt{5}} \cos(2\omega t - \tan^{-1} 2) \quad (7.32)$$

$$v_{out2} = \frac{A_1}{\sqrt{2}} \cos\left(\omega t + \frac{\pi}{4}\right) + \frac{2A_2}{\sqrt{5}} \cos\left(2\omega t - \tan^{-1} 2 + \frac{\pi}{2}\right) \quad (7.33)$$

模拟结果表明，要达到小于 1° 的相位误差，应该满足 $A_2 < 1.4\% A_1$ 。如果只存在三次谐波，则为达到 1° 的相位误差，其幅度应小于 1.2% 的基波幅度。如果本振谐波很大，应该在 RC-CR 网络前插入一个低通滤波器。

图 7.40 的 C 值通常选取使得负载电容上的衰减达到最小，例如 $C \approx 5C_L$ 。然而，大的 C 值，会使电路的输入阻抗 ($\approx 1/Cs$) 减小。

7.8.2 Havens 方法

如图 7.42(a)所示，Havens 方法首先将信号分离 90° 左右，得到 V_1 和 V_2 。接着对这两路

信号进行相加和相减，得到 V_{out1} 和 V_{out2} [27]。如图 7.42(b)所示，如果 V_1 和 V_2 具有相同的幅度，则 V_{out1} 和 V_{out2} 之间的角度为 90° 。这个结论可以从如下步骤看出：令 $v_1(t) = A \cos \omega t$ ， $v_2(t) = A \cos(\omega t + \theta)$ ，得到

$$v_1(t) + v_2(t) = 2A \cos \frac{\theta}{2} \cos \left(\omega t + \frac{\theta}{2} \right) \quad (7.34)$$

$$v_1(t) - v_2(t) = 2A \sin \frac{\theta}{2} \sin \left(\omega t + \frac{\theta}{2} \right) \quad (7.35)$$

图 7.42 (a) Havens 正交电路，(b) 操作的相量图

相移电路之后的各限幅级使得 v_1 和 v_2 的幅度相等。而且，如果 $\theta \neq 90^\circ$ ，加法器的输出具有不同的幅度，这些输出同样被加在限幅器的输入。在这两种情况下，AM-PM 转换的问题都很重要。

如果图 7.42(a)的两个加法器检测到的信号幅度出现失配，则输出会出现相位失衡。为了量化这个效应，我们假设 $v_2(t) = (A + \varepsilon) \cos(\omega t + \theta)$ （其中 ε 表示幅度失配），并且重新画出图 7.42(b)的相图[图 7.43(a)]。我们注意到 $\varepsilon \cos(\omega t + \theta)$ 的效应在于使 v_{out1} 逆时针旋转了 ϕ_1 而 v_{out2} 顺时针旋转了 ϕ_2 。借助于图 7.43(b)，我们可以得到

$$\tan \phi_1 = \frac{\varepsilon \sin(\theta/2)}{2A \cos(\theta/2) + \varepsilon \cos(\theta/2)} \quad (7.36)$$

$$\tan \phi_2 = \frac{\varepsilon \cos(\theta/2)}{2A \sin(\theta/2) + \varepsilon \sin(\theta/2)} \quad (7.37)$$

从而，对于 $\varepsilon \ll A$ ，我们有

$$\phi_1 + \phi_2 \approx \frac{\varepsilon}{2A} \left[\frac{\sin(\theta/2)}{\cos(\theta/2)} + \frac{\cos(\theta/2)}{\sin(\theta/2)} \right] \quad (7.38)$$

$$= \frac{\varepsilon}{A} \frac{1}{\sin \theta} \quad (7.39)$$

由于 θ 在 90° 附近，(7.39)表明 1% 的幅度失配导致 0.6° 的相位误差。

图 7.43 (a) 在 Havens 电路中幅度失配的效应，(b) (a) 的简化了的版本

上述分析是基于正弦信号的假设的。然而无论是 90° 移相级之后的限幅级，还是加法器的非线性，都会产生 $v_1(t)$ 和 $v_2(t)$ 的谐波。而且，输入信号本身也可能包含谐波。我们首先

考虑二次谐波效应。

假设 $v_1(t)$ 和 $v_2(t)$ 包含等幅的二次谐波，例如分别为 $a \cos 2\omega t$ 和 $a \cos(2\omega t + 2\theta)$ 。对这两个信号进行相加和相减同样得到两个频率为 2ω ，相差为 90° 的信号。然而，回忆 7.8.1 节所述，如果欲使 v_{out1} 和 v_{out2} 正交，则第 n 次谐波必须具有 $n\pi/2$ 的相差。这样，两个相差为 90° 的二次谐波分量的存在会导致两个输出之间的相位失衡。

奇次谐波的影响相对不是很大。假设在限幅之后， $v_1(t) = A \cos \omega t + a \cos(2k+1)\omega t$ 而 $v_2(t) = A \cos(\omega t + \theta) + a \cos[(2k+1)\omega t + (2k+1)\theta]$ ，其中 k 为一个整数。经过相加和相减，我们得到

$$v_{\text{out1}}(t) = 2A \cos \frac{\theta}{2} \cos \left(\omega t + \frac{\theta}{2} \right) + 2a \cos \frac{(2k+1)\theta}{2} \cos \left[(2k+1)\omega t + \frac{(2k+1)\theta}{2} \right] \quad (7.40)$$

$$v_{\text{out2}}(t) = 2A \sin \frac{\theta}{2} \sin \left(\omega t + \frac{\theta}{2} \right) + 2a \sin \frac{(2k+1)\theta}{2} \sin \left[(2k+1)\omega t + \frac{(2k+1)\theta}{2} \right] \quad (7.41)$$

为了 v_{out1} 和 v_{out2} 具有 $\pi/2$ 的相差，应有 $v_{\text{out2}}(t) = v_{\text{out1}}[t - \pi/(2\omega)]$ 。计算此项

$$v_{\text{out1}}\left(t - \frac{\pi}{2\omega}\right) = 2A \cos \frac{\theta}{2} \sin \left(\omega t + \frac{\theta}{2} \right) + 2a \cos \frac{(2k+1)\theta}{2} \sin \left[(2k+1)\omega t + \frac{(2k+1)\theta}{2} - k\pi \right] \quad (7.42)$$

我们注意到(7.41)和(7.42)的第二项有幅度和极性两方面的区别。这样，奇次谐波只引起幅度的失配，但不引起相位失配。

Haven 结构的另一个幅度失配来源于加法器两路输入信号之间的电容耦合。图 7.44 所示为一个例子，其中 v_1 和 v_2 先加到一个限幅级然后加到一个由两个差分对组成的加法器。晶体管的栅漏覆盖电容提供了 I 通道到 Q 通道的一条路径，实质上将近交信号的相位拉近。类似于图 7.40 的 v_{out1} 与 v_{out2} 之间的电容性馈送，如果限幅器的输出阻抗不够低，这种效应会使得加法器输入端的信号出现很大的相位失配。

图 7.44 加法器中的电容性路径

7.8.3 分频

产生频率为 ω_1 的正交周期信号的一个简单方法是采用一个主-从触发器对一个频率为 $2\omega_1$ 的信号进行二分频（图 7.45）。如果 V_{in} 占空比为 50%，则 V_{out1} 与 V_{out2} 的相差为 90° 。二分频器的不同电路实现会在第 8 章中介绍。

图 7.45 分频器用作正交发生器

这种方法的主要困难在于产生一个频率为 $2\omega_1$ 的信号并且对其分频会消耗大量的能量，或者由于工艺的限制不可能实现。另一个问题是由于输入信号的占空比偏离 50% 时引起的相位失衡。例如，如果 V_{in} 包含二次谐波，则类似所述， V_{out1} 与 V_{out2} 会出现类似于图 7.40 中的 RC-CR 网络描述的相位失衡。通过锁存器的信号路径上的失配也会带来相位误差。

7.9 单边带信号生成

接收发送器的结构通常需要对不同本振的输出进行相加或相减操作。例如 5.3.1 节所述的应用在图 5.49 的 GSM 系统的偏移本振技术将两路信号相乘，并且通过带通滤波器选择其和频或差频。然而如图 7.46(a) 所示，如果一个频率比另一个小很多，则滤波器需要有很陡峭的截止特性来抑制不想要的边带。

产生各种本振频率的另一个方法是采用单边带结构。如图 7.46(b) 所示，其思想基于等式 $\cos \omega_1 t \cos \omega_2 t \pm \sin \omega_1 t \sin \omega_2 t = \cos(\omega_1 \mp \omega_2)t$ 。它需要两路信号每一信号的正交相位。然而，这种拓扑结构受到两个非理想特性的限制：上下两条路径的失配以及混频器的非线性。由于相位和增益的失配带来的无用边带在 5.3 节有介绍。因此，我们只考虑信号路径的非线性。

图 7.46 (a) 简单混频，(b) 理想 SSB 混频

回忆第 6 章介绍的，混频器通常设计使得本振端口具有陡峭的开关特性。也就是说，本振端是如此强非线性的，以致于射频信号是被一个矩形波相乘。而且，即使有简并反馈，射频端口也表现为非线性。因此，图 7.46(b) 中， ω_1 和 ω_2 的谐波同时在每一个混频器的端口产生，使得相乘后出现各种交叉乘积项。

由于这个电路中的无用边带相对有用边带可以保持在约 -30 dB 以下，采用 SSB 产生器可以很大缓解图 7.46(a) 中带通滤波器的关断特性。更复杂的校正和线性化技术甚至可以免除对外部滤波器的需求。

参考文献

- [1] B. van der Pol, "The nonlinear Theory of Electric Oscillations," *Proc. IRE*, Vol.22, pp. 1051-1086, September 1934.
- [2] N.M. Nguyen and R.G. Meyer, "Start-up and Frequency Stability in High-Frequency Oscillators," *IEEE Journal of Solid-State Circuits*, Vol. 27, pp. 810-820, May 1992.
- [3] H.L. Kraus, C.W. Bostian, and F.H. Raab, *Solid State Radio Engineering*, New York: John Wiley, 1980
- [4] B. Law and M. Groh, "Identifying RF-Related Impairments in Full-Service Digital Networks," *Microwave J.*, pp. 88-94, March 1996.
- [5] E. Hafner, "The Effect of Noise in Oscillators," *Proc. IEEE*, Vol.54, pp.179-198, February 1966.
- [6] D. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," *Proc. IEEE*, Vol.54, pp. 329-330, February 1966.
- [7] W.P. Robins, *Phase Noise in Signal Sources*, London: Peter Peregrinus, 1982.
- [8] B. Jansen, K. Negus, and D. Lee, "Silicon Bipolar VCOs for 1.1 to 2.2 GHz with Fully-Integrated Tank and Tuning Circuits," *ISSCC Dig. Of Tech. Papers*, pp. 392-393, February 1997.
- [9] C.A.M. Boon, *Design of High-Performance Negative Feedback Oscillators*, Ph.D. Thesis, Delft University, The Netherlands, 1989.
- [10] N.M. Nguyen and R.G. Meyer, "Si IC-Compatible Inductors and LC Passive Filters," *IEEE Journal of Solid-State Circuits*, Vol. 25, pp. 1028-1031, August 1990.
- [11] R.B. Merrill et al., "Optimization of High Q Inductors for Multi-Level Metal CMOS," *Proc. IEDM*, pp. 38.7.1-38.7.4, December 1995.
- [12] K. Kurokawa, "Injection Locking of Microwave Solid-State Oscillators," *Proc. IEEE*, Vol. 61, pp. 1386-1410, October 1973.
- [13] R. Adler, "A Study of Locking Phenomena in Oscillators," *Proc. IRE*, Vol. 34, pp. 351-357, June 1946.
- [14] R. Mohindra, "Isolator for DECT Open-Loop Modulator," *RF Design*, pp. 30-42, January 1996.
- [15] B. Razavi, "A 1.8 GHz CMOS Voltage-Controlled Oscillator," *ISSCC Dig. Tech Papers*, pp. 388-389, February 1997.
- [16] B. Razavi, "A 900-MHz CMOS Direct-Conversion Receiver," *VLSI Circuits Symp. Dig. Of Tech. Papers*, June 1997.
- [17] J. Craninckx and M.S.J. Steyaert. "A 1.8 GHz CMOS Low Phase Noise Voltage-Controlled Oscillator with Prescaler," *IEEE Journal of Solid-State Circuits*, Vol. 30, pp. 1474-1482, December 1995.
- [18] N.M. Nguyen and R.G. Meyer, "A 1.8-GHz CMOS Monolithic LC Voltage-Controlled Oscillator," *IEEE Journal of Solid-State Circuits*, Vol. 27, pp. 444-450, March 1992.
- [19] J. Craninckx and M.S.J. Steyaert and Miyakawa, "A Fully Integrated Spiral LC CMOS VCO Set with Prescaler for GSM and DCS-1800 Systems," *Proc CICC*, pp. 403-406, May 1997.
- [20] H.M. Greenhouse, "Design of Planar Rectangular Microelectronic Inductors," *IEEE Trans. On Parts, Hybrids, and Packaging*, Vol. 10, pp. 101-109, June 1974.

- [21] F.E. Terman, *Radio Engineers' Handbook*, New York: McGraw-Hill, 1943.
- [22] F.W. Grover, *Inductance Calculations*, Princeton, NJ: Van Nostrand, 1946.
- [23] B. Razavi, "Challenges in the Design of Frequency Synthesizers for Wireless Applications," *Proc. CICC*, pp. 395-402, May 1997.
- [24] B. Razavi, ed. *Monolithic Phase-Locked Loops and Clock Recovery Circuits*, Piscataway, NJ: IEEE Press, 1996.
- [25] B. Razavi, "A Study of Phase Noise in CMOS Oscillators," *IEEE Journal of Solid-State Circuits*, Vol. 31, pp. 331-343, March 1996.
- [26] R.J van de Plassche, "An 8-Bit 100-MHz Full-Nyquist Analog-to-Digital Converter," *IEEE J. Solid-State Circuits*, Vol. SC-27, pp. 1334-1344, December 1988.
- [27] I.A. Koullias et al., "A 900 MHz Transceiver Chip Set for Dual-Mode Cellular Radio Mobile Terminals," *ISSCC Dig. of Tech. Papers*, pp.140-141, February 1993.

第 8 章 频率综合器

接收发送器中的振荡器通常是嵌入在频率合成器的环境中以取得精确的输出频率。由于频率合成器必须要满足非常严格的要求，所以它的设计仍然是射频系统中的最具挑战性的任务之一。

在本章中，我们描述了几种频率合成器的设计方法，重点强调了它们在低功耗，单片集成电路实现中的优缺点。我们首先学习锁相的概念并分析几种不同的锁相环（PLLs）¹。接下去，我们将展示几种频率合成器结构，包括整数 N ，分数 N 和直接数字频率合成器技术。最后，我们将处理分频这个问题。

如果需要更多的关于频率合成器的知识，请读者参考文献[1, 2, 3]。

8.1 概述

射频接收发送器中的振荡器的频率必须有很高的绝对精度。而且在许多情况下，这个频率还要以精确的小步长进行变化。重新回顾第四章中描述的无线标准，当中心频率在 900MHz 或者 1.9GHz 附近时信道间隔可以小到 30kHz。换一句话说，为了改变接收或者发射的信道，本机振荡器的频率可能需要仅仅改变 30kHz。同时，每个信道的下界和上界频率是精确定义的，只能容许不超过几百 Hz 的误差。因此输出频率的误差必须小于百万分之几。图 8.1 所示的通用接收发送器中给出了频率合成器的作用。

图 8.1 通用接收发送器结构

除了精度和信道间隔之外，频率合成器还有其他几个方面影响的收发器的性能：相位噪声，边带（“毛刺”）和锁定时间。就像第七章中所解释的，本地振荡器的相位噪声同时影响接收和发射路径。

尽管一个自由工作的振荡器通常不会产生边带，但是当它被嵌入在频率合成器里则有可能产生边带。如图 8.2 所示，不想要的边带的影响对于接收路径尤其带来麻烦。假设频率合成器的输出中包含一个频率为 ω_{LO} 的载波信号和一个频率为 ω_s 的边带信号，而接收到的信号伴

随着一个频率为 ω_{int} 的干扰信号。可以看到在下变频之后，出现了两个重要的分量：我们所希望接收的信道信号与载波的卷积和干扰信号与边带的卷积。如果 $\omega_f - \omega_s = \omega_0 - \omega_{LO} (= \omega_{IF})$ ，那么下变频的干扰信号落在希望的信道中。由于这个原因，

典型的系统要求边带比载波信号低大概 60dB。然而，如果 $\omega_{LO} - \omega_s$ 也就是 $\omega_0 - \omega_{int}$ 足够的大，则干扰信号出现在接收频带之外，因此将在一定程度上被前端天线收发转换开关或者带通滤波器所抑制。

¹ 8.2 节中的一部分是获得了许可的重印版本，来自于 *Monolithic Phase-Locked Loops and Clock recovery Circuit*, B. Razavi, IEEE Press, Piscataway, NJ, pp. 4-32, ©1996 IEEE.

图 8.2 接收器中频率合成器的边带的影响

频率合成器的锁定时间也是一个非常关键的参数。如图 8.3 所示，当一个数字的信道选择输入命令改变信道时，频率合成器需要一个有限的时间来稳定新的频率。在 8.3.1 小节中给出了更精确的定义，锁定时间表明了新的频率达到稳定有多快。这个参数在跳动快频扩展频谱的系统中特别重要。在典型的射频系统中要求的锁定时间从几十毫秒到几十微秒不等。

图 8.3 频率合成器稳定过程

8.2 锁相环

8.2.1 基本概念

压控振荡器动力学 从第七章我们看到，一个理想的压控振荡器的特性可以表征为 $\omega_{\text{out}} = \omega_{FR} + K_{VCO}v_{\text{cont}}$ 和 $y(t) = A_C \cos \left[\omega_{FR}t + K_{VCO} \int_{-\infty}^t v_{\text{cont}}(t) dt \right]$ 。在研究锁相环时，我们通常把压控振荡器当作一个线形时不变系统，其系统输入为控制电压，系统输出为载波的过量相位（第三章）。因为过量相位 $\phi_{\text{out}}(t) = K_{VCO} \int v_{\text{cont}} dt$ ，所以输入输出传输函数为

$$\frac{\Phi_{\text{out}}(s)}{V_{\text{cont}}} = \frac{K_{VCO}}{s} \quad (8.1)$$

压控振荡器中的积分导致一个有趣的特性：要改变输出相位，我们必须先改变频率并让积分发生²。举个例子：假设当 $t < t_0$ 时，压控振荡器以与参考信号相同的频率振荡但带有一个有限的相位误差（图 8.4）。为了减小误差，控制电压 v_{cont} 在 $t = t_0$ 这个时刻增加了 ΔV ，因此压控振荡器的频率增加了并使得输出信号相位的积累快过了参考信号。在 $t = t_1$ 时刻，当相位误差减小到零时 v_{cont} 恢复到它的初始值。现在这两个信号具有相同的频率和零相位差。注意在这个期间中也可以通过减小压控振荡器的频率来达到同样的目标。

图 8.4 压控振荡器与参考信号相位对准

上边的观察也产生了另一个有趣的结果：压控振荡器的输出相位不能仅凭当前的控制电压来确定，也就是说它依赖于 v_{cont} 的历史。由于这个原因，在时域范围内分析锁相环时，我们把压控振荡器的输出相位当作在 PLL 时域分析中的独立的初始条件（状态变量）。

鉴相器 一个理想的鉴相器(PD)产生一个直流信号，其值正比于两个周期输入信号的相位差

² 我们假设压控振荡器没有其它的输入来改变相位。

(图 8.5),

$$\overline{v_{out}} = K_{PD} \Delta\phi \quad (8.2)$$

这里 K_{PD} 是鉴相器的“增益”(单位是 V/rad) 而 $\Delta\phi$ 是输入信号的相位差。在实际情况中, 鉴相器的特性可能不是线形的, 对于较大的 $\Delta\phi$ 甚至不是单调的。而且, K_{PD} 也可能与输入信号的幅度或占空比有关。这些情况将在以后解释。

图 8.5 理想鉴相器的特性

图 8.6 给出了一个鉴相器输出的典型例子, 它的输出脉冲宽度等于两个输入信号相邻的越零时间的差值。因为这两个信号的频率不相等, 所以相位差显示出一个均值为零的“差拍”行为。

图 8.6 鉴相器的输入与输出波形图

一种常用的鉴相器是乘法器(也叫作混频器或者正弦鉴相器)。对于两个输入信号 $x_1(t) = A_1 \cos \omega_1 t$ 和 $x_2(t) = A_2 \cos(\omega_2 t + \Delta\phi)$, 乘法器的输出为

$$y(t) = \alpha A_1 \cos \omega_1 t \cdot A_2 \cos(\omega_2 t + \Delta\phi) \quad (8.3)$$

$$= \frac{\alpha A_1 A_2}{2} \cos[(\omega_1 + \omega_2)t + \Delta\phi] + \frac{\alpha A_1 A_2}{2} \cos[(\omega_1 - \omega_2)t - \Delta\phi] \quad (8.4)$$

这里 α 是一个比例常数。因此, 当 $\omega_1 = \omega_2$ 时, 相位/电压的特性由下式给出

$$\overline{y(t)} = \frac{\alpha A_1 A_2}{2} \cos \Delta\phi \quad (8.5)$$

如图 8.7 所画出的, 这个函数有着变化的斜率并且是非单调的。但是在 $\Delta\phi$ 位于 $\pi/2$ 的附近时, 它与 8.2 式类似,

$$\overline{y(t)} \approx \frac{\alpha A_1 A_2}{2} \left(\frac{\pi}{2} - \Delta\phi \right) \quad (8.6)$$

得到 $K_{PD} = -\alpha A_1 A_2 / 2$ 。需要注意, 当 $\omega_1 \neq \omega_2$ 时, 输出的平均值为零。

图 8.7 一个正弦鉴相器的特性

如果输入信号相当的大, 一个模拟的乘法器可以用一个 Gilbert 单元来实现, 像一个异或门那样工作。

8.2.2 基本的锁相环

简单回路 锁相环是一个反馈系统, 它工作在通常是周期信号的过量相位。这与通常那些关心电压和电流的幅度或者它们的变化率的反馈电路是截然不同的。图 8.8 给出了一个简单的

锁相环，由一个鉴相器，一个低通滤波器（LPF）和一个压控振荡器构成。鉴相器在反馈回路中起到一个“误差放大器”的作用，以此来最小化 $x(t)$ 和 $y(t)$ 之间的相位差 $\Delta\phi$ 。如果 $\Delta\phi$ 不随时间变化，则回路就被认为“锁定”了，其结果是输入信号与输出信号的频率是相等的。

图 8.8 基本的锁相环回路

在锁定的条件下，回路中的所有信号都达到了一个稳定的状态，而锁相环就像后面所描述的那样工作。鉴相器产生一个正比于 $\Delta\phi$ 的直流输出信号。低通滤波器抑制了鉴相器输出信号中的高频部分，而只允许直流分量来控制压控振荡器的频率。于是压控振荡器以与输入信号相同的频率进行振荡，两者之间存在一个相位差 $\Delta\phi$ 。因此，低通滤波器为压控振荡器产生了一个合适的控制电压。

来看一看锁相环中不同节点上的信号是有启发性的。图 8.9 中是一个典型的例子。输入信号与输出信号有着相同的频率但是一定的相位差，而鉴相器产生一个宽度等于输入输出信号之间越零时间差的脉冲信号。这些脉冲通过低通滤波器得到直流的电压来维持压控振荡器工作在所需的频率上。如同上边所描述的，电压本身并不直接决定输出相位。压控振荡器的输出相位可以看作是系统的初始条件，而与低通滤波器的初始条件无关。

现在让我们来定性研究当一个在时间 $t < t_0$ 时处于锁定状态的锁相环，在时刻 $t = t_0$ 时它经历了输入信号的一个正的频率阶跃变化时的响应情况（图 8.10）（为了显示方便，本图中的频率变化只有几个百分点）。我们注意到因为输入信号的频率 ω_{in} 瞬时地高于输出信号的频率 ω_{out} ，所以 $x(t)$ 的相位积累就要快于 $y(t)$ ，鉴相器产生了一个不断加宽的脉冲信号。每一个这样的脉冲在 LPF 的输出端产生了一个不断升高的直流电压，因此也就提高了压控振荡器的频率。当 ω_{in} 与 ω_{out} 之间的差别消失时，相位比较脉冲的宽度也减小，最终回复到一个略高于 $t = t_0$ 时刻之前的值。

图 8.9 锁相环中的波形图

上面的分析提供了锁相环具有“跟踪”能力的深层了解。如果输入信号的频率缓慢地变化，那么它的变化可以看作是一系列小窄步变化，在每一个小窄步变化中锁相环都如同图 8.10 中所示的一样工作。

图 8.10 锁相环对一个小的频率阶跃变化的响应

特别需要注意到，前面的例子中环路只有在两个条件得到满足时才锁定：(1) ω_{out} 已经变得与 ω_{in} 相同。(2) ϕ_{in} 与 ϕ_{out} 之间的差已经固定到了它合适的值[4]。如果在某个时间点上两个频率相等，而 $\Delta\phi$ 还没有为压控振荡器建立所需的控制电压，那么环路必须继续它的瞬态过

程，暂时地使得频率再次不相等。换一句话说，“频率获取”与“相位获取”必须都被满足。为了再次达到锁定，系统所有的初始条件包括压控振荡器的输出相位理所当然地都必须更新。

如果一个锁相环的输入信号存在一个固定的过量相位，也就是说输入信号是一个严格的周期函数，但输入输出相位误差 $\Delta\phi$ 随着时间变化，这种情况，我们称环路为“未锁定”。这是一种不希望得到的状态，因为输出信号未能跟踪输入信号或者两者之间的关系过于复杂而不能加以利用。例如，如果 ω_m 远离压控振荡器的自由工作频率时，环路也许永远不能锁定。

尽管，锁相环处于未锁定状态时的行为并不重要，而它是否以及如何进入锁定状态都是关键性的问题。

在更加仔细地研究锁相环之前，我们三个重要的观察结果。首先，因为锁相环是一个有“记忆”的系统，它的输出需要一个有限的时间来对输入信号的变化产生响应，因此要求对环路动力学过程有一个深刻的理解。第二，不同于其它许多的反馈系统，在锁相环中所感兴趣的量的量纲在环路中发生了变化：通过鉴相器将相位变为了电压（或者电流），仍然以电压的形式为低通滤波器所处理，又被压控振荡器变回成相位。第三，在锁定的情况下，输出信号的频率与输入信号 *严格地* 相等，而不管环路的增益的值为多少（尽管相位误差不一定为零）。这一特性非常重要，因为包扩频率合成器在内的许多应用都不能容忍输出信号与输入信号之间非常小（系统性的）的频率差别。注意如果用一个频率鉴别器来代替鉴相器，那么这个特性就消失了。

尽管锁相环是对相位进行操作，但是在很多情况下我们关心的参数却是频率。例如，在下列情况下我们经常需要知道环路的响应：(1) 输入信号的频率变化缓慢。(2) 输入信号的频率变化迅速。(3) 锁相环开启的时候，输出信号与输入信号的频率不同。所以在锁相环中，鉴相器在输入信号频率不等时的特性扮演了非常重要的角色。

锁定状态时的环路动力学 锁相环的瞬态响应通常是一种非线性现象，不可能很简单地得到公式。但是如同其它反馈系统的情况一样，在锁相环的设计中，可以利用线性近似来得到直觉与对折衷方案的理解。

图 8.11 给出了一个处于锁定状态时锁相环的线性模型以及每一模块的传输函数。这个模型是为相位来提供总的传输函数 $\Phi_{out}(s)/\Phi_{in}(s)$ ，因而鉴相器用一个减法器来代表。假设低通滤波器有一个电压传输函数 $G_{LPF}(s)$ 。这样锁相环的开环传输函数等于

$$H_O(s) = K_{PD} G_{LPF}(s) \frac{K_{VCO}}{s} \quad (8.7)$$

得到了如下的闭环传输函数：

$$H(s) = \frac{\Phi_{out}(s)}{\Phi_{in}(s)} \quad (8.8)$$

$$= \frac{K_{PD} K_{VCO} G_{LPF}(s)}{s + K_{PD} K_{VCO} G_{LPF}(s)} \quad (8.9)$$

图 8.11 锁相环的线性模型

最简单的低通滤波器形式如图 8.12 所示，有

$$G_{LPF}(s) = \frac{1}{1 + \frac{s}{\omega_{LPF}}} \quad (8.10)$$

这里 $\omega_{LPF} = 1/(RC)$ 。于是方程(8.9)可以简化为

$$H(s) = \frac{K_{PD}K_{VCO}}{\frac{s^2}{\omega_{LPF}^2} + s + K_{PD}K_{VCO}} \quad (8.11)$$

上式表明这是一个二阶系统，有两个极点，一个来自于压控振荡器，一个来自于低通滤波器。量 $K = K_{PD}K_{VCO}$ 称为“环路增益”，量纲是 rad/sec。

图 8.12 简单的低通滤波器

为了理解锁相环的动力学行为，我们将(8.11)式的分母化为控制理论中常用的形式 $s^2 + 2\zeta\omega_n s + \omega_n^2$ ，这里 ζ 是阻尼系数而 ω_n 是系统的固有频率³。因此

$$H(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (8.12)$$

这里

$$\omega_n = \sqrt{\omega_{LPF} K} \quad (8.13)$$

$$\zeta = \frac{1}{2} \sqrt{\frac{\omega_{LPF}}{K}} \quad (8.14)$$

注意 ω_n 是低通滤波器的-3dB 带宽与环路增益的几何平均值，在某种意义上代表了环路的增益带宽积。同时，阻尼系数与环路增益成反比关系，这是一个非常重要然而经常不受欢迎的折衷。

在一个设计良好的二阶系统中， ζ 通常大于 0.5 而最好是等于 $\sqrt{2}/2$ 从而得到最优的平坦响

应。因此， K 与 ω_{LPF} 不能各自独立的进行选择。例如，如果 $\zeta = \sqrt{2}/2$ ，那么 $K = \omega_{LPF}/2$ 。

式(8.12)所示的传输函数是一个低通滤波器。如果输入信号的过量相位缓慢地变化，那么输出信号的过量相位就会跟着变化；相反地，如果输入信号的过量相位迅速地变化，那么输出信号过量相位的变化就会很小。特别是，如果 $s \rightarrow 0$ ，我们会发现 $H(s) \rightarrow 1$ ，也就是说，输入信号一个静态的相位偏移会毫无改变地传输到输出信号中。这是因为对于相位量，当 $s \rightarrow 0$ 时，在压控振荡器中积分的存在使得开环增益趋近于无穷大。为了理解这一点，我们

³ 在一个简单的锁相环中， ω_n 与输入信号和输出信号的频率无关。

来研究一下“相位误差传输函数”，它在图 8.11 中定义为 $H_e(s) = \Phi_e(s)/\Phi_{in}(s)$ 的

$$H_e(s) = 1 - H(s) \quad (8.15)$$

$$= \frac{s^2 + 2\zeta\omega_n s}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (8.16)$$

当 $s \rightarrow 0$ 时下降为 0。

因为相位与频率是通过一个线性时不变运算关联在一起的，所以公式(8.12)和(8.16)所示的传输函数也可以用于输入信号和输出信号的过量频率。例如，式(8.12)表明如果输入信号的频率迅速地变化，那么输出信号频率的瞬时变化就会很小。

在式(8.12)的帮助下来再次进行我们前面对环路阶跃响应所做的分析（图 8.10）很有意义。

假设输入信号的过量频率等于 $\Delta\omega u(t)$ ，这里 $u(t)$ 是单位的阶跃函数（图 8.13）。输出信号

的过量频率就表现为一个典型的二阶系统的阶跃响应，最终稳定在比初始值高 $\Delta\omega$ rad/sec

的频率上。另一方面，输出的过量相位由下式给出

$$\Phi_{out}(s) = H(s)\Phi_{in}(s) \quad (8.17)$$

$$= \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \frac{\Delta\omega}{s^2} \quad (8.18)$$

这是一个二阶系统对斜坡输入的相应。更重要的是，相位误差为

$$\Phi_e(s) = H_e(s)\Phi_{in}(s) \quad (8.19)$$

$$= \frac{s^2 + 2\zeta\omega_n s}{s^2 + 2\zeta\omega_n s + \omega_n^2} \frac{\Delta\omega}{s^2} \quad (8.20)$$

最终的值为

$$\phi_e(t = \infty) = \lim_{s \rightarrow 0} s\phi_e(s) \quad (8.21)$$

$$= \Delta\omega \frac{2\zeta}{\omega_n} \quad (8.22)$$

$$\frac{\Delta\omega}{K} \quad (8.23)$$

因此，输入信号的一个静态频率变化体现在静态相位误差（图 8.13）上就是它除以一个系数 K 。这是十分合理的，因为为了让压控振荡器的频率变化 $\Delta\omega$ ，就必须让控制电压改变 $\Delta\omega/K_{VCO}$ 而鉴相器的输入改变 $\Delta\omega/(K_{VCO}K_{PD})$ 。

图 8.13 频率发生了一个阶越变化时锁相环的响应

在时域中知道阶跃响应的近似关系也是很有用处的。对于一个 $\zeta < 1$ 的二阶系统，阶跃响应由下式给出

$$y(t) = \left[1 - \frac{1}{\sqrt{1-\zeta^2}} \exp(-\zeta\omega_n t) \times \sin(\omega_n \sqrt{1-\zeta^2} t + \Psi) \right] u(t) \quad (8.24)$$

其中 $\Psi = \sin^{-1} \sqrt{1-\zeta^2}$ 。

跟踪与获取 在大多数的应用中，锁相环的性能中有两个方面十分重要。第一个是“跟踪”行为，也就环路可以跟随输入信号频率变化的程度。第二点就是“获取”特性，也就是环路如何从未锁定状态进入到完全的相位锁定。

关于这些特性的详细分析可以在[4, 5, 6]中找到，但是在这里我们应该提到，为了保证跟踪和获取的频率范围足够宽，大多数的锁相环在进行相位鉴别的同时还集成了频率比较。其思想就是当压控振荡器与输入信号的频率相差较大时，鉴频装置就主导了反馈回路的控制，使得压控振荡器的频率接近输入频率。当频率差下降到足够小时，用相位鉴别来取代它，实现最后的相位锁定。

8.2.3 电荷泵锁相环

相位/频率鉴别器 一个能够同时确定相位和频率差的电路是非常有用的，因为它可以极大地提高锁相环的捕获范围和锁定速度。

与乘法器和异或门不同，顺序相位/频率鉴别器(PFDs)产生的两个输出信号，它们并不是互补的。如图 8.14 所示，一个典型的 PFD 的工作过程如下。如果输入信号 A 的频率高于输入信号 B，那么 PFD 在 Q_A 端产生一个正的脉冲，而 Q_B 端保持为零。反过来，如果 $\omega_A < \omega_B$ ，

那么 Q_B 输出正的脉冲而 $Q_A = 0$ 。如果 $\omega_A = \omega_B$ ，那么 PFD 则在 Q_A 与 Q_B 中的任意一端产生一个宽度为两个输入信号相位差的脉冲信号(注意，原则上， Q_A 与 Q_B 不会同时为高电平)。

因此， $Q_A - Q_B$ 的平均值就代表了输入信号 A 与 B 之间的频率差或者相位差。输出信号 Q_A 与 Q_B 通常被称为“上升”和“下降”信号。

图 8.14 相位/频率鉴别器的响应 (a) 当 $\omega_A > \omega_B$ 时 (b) 当 A 滞后 B 时

为了在电路中实现上述的功能，我们要求至少要有三个逻辑状态存在： $Q_A = Q_B = 0$ ；

$Q_A = 0, Q_B = 1$ 和 $Q_A = 1, Q_B = 0$ 。同时，为了避免输入信号占空比对输出信号的影响，电路应该采用边沿触发的时序机。我们假定只有在 A 或者 B 的上升沿时电路的状态才会发生改变。为了简明起见，我们在以后的描述中省略了“上升”这个词。图 8.15 给出了操作过程的状态图。如果 PFD 处于“基态”， $Q_A = Q_B = 0$ ，那么 A 的一个沿将它改变到了状态 I，

这时 $Q_A = 1, Q_B = 0$ 。电路保持在这种状态，直到 B 信号发生了翻转，PFD 回到了状态 0。

从状态 0 到状态 II 的切换过程与此类似。

图 8.15 PFD 的状态图

这个状态图中非常重要的一点就是，如果 $\omega_A > \omega_B$ 那么在某段时间内，在 B 信号的两次翻转中间将会存在 A 信号的两次翻转。这一点就保证了即使 PFD 的初始状态是状态 II，最终电路也一定会离开这个状态而在状态 0 和状态 I 之间来回切换[7]。

图 8.16[7]中给出了一种可能的方法来实现上述 PFD。电路中包含了两个边沿触发，可重置的 D 触发器，其 D 输入端都被接在了逻辑值 1 上。信号 A 和 B 分别作为 D 触发器 DFF_A 和 DFF_B 的时钟输入。我们注意到，如果 $Q_A = Q_B = 0$ ，那么 A 信号的一个翻转导致 Q_A 变为高电平，A 信号接下来的翻转对 Q_A 不再起作用，而当 B 信号变为高电平时，电路中的与门重置了两个触发器。这样， Q_A 和 Q_B 在一段时间里同时处于高电平，而这段时间是由通过触发器的与门和重置路径的总延迟给定的。图 8.17 中显示了 PFD 的输入输出特性。

图 8.16 PFD 的实现

图 8.17 PFD 的特性

图 8.16 中的 D 触发器在双极型集成电路和 CMOS 集成电路中可能采用不同的拓扑结构来实现。在双极工艺中，带有一个附加的重置输入的，标准主从结构可以被采用。在 CMOS 工艺中，图 8.18[8]所示的简单电路就足够了。注意这里的 D 输入信号被“隐藏”了起来。还有其它的 PFD 实现方法在[9, 10]中得到了描述。

PFD 的输出信号可以用不同的方法来转换成直流信号。一种方法是用一个差分放大器来探测两个输出信号的差值，再把结果加到一个低通滤波器。另一种方法是让输出来驱动一个三态的“电荷泵”。

图 8.18 图 8.16 中每一个 D 触发器的实现

电荷泵 在迄今为止所考虑的低通滤波器中(图 8.12)，鉴相器输出的平均值是通过在每个相位比较周期中对电容进行充电然后再放电来实现的。另一方面，在电荷泵中，相位比较时刻之间的只有可忽略的放电，这导致了一些有趣的结果。

一个三态的电荷泵与一个三态的相位/频率鉴别器(图 8.19)放在一起研究，理解效果最好。

电荷泵本身是由两个驱动一个电容的开关电流源来构成的。注意对于 Q_A 中一个宽度为 T 的脉冲， I_1 向电容 C_p 充了电量为 IT 的电荷。因此当 $\omega_A > \omega_B$ 或者 $\omega_A = \omega_B$ 但是 A 超前于 B 的时候，正的电荷在电容 C_p 上不断地积累起来，从而对 PFD 产生了一个无限大的直流增益。

类似地，如果脉冲出现在 Q_B 中，在每次相位比较时 I_2 从电容 C_p 上放走电荷，使得 V_{out} 向 $-\infty$ 变化。在第三种状态中 $Q_A = Q_B = 0$ ， V_{out} 保持为常数。因为稳态的增益为无穷大，所以定

义在一个比较瞬间时的 PFD/电荷泵的组合增益更具意义，该增益的值为 $IT/(2\pi C_p)$ 。

通过上边的研究我们得到了一个重要的结论，如果偏移与失配可以被忽略，那么采用这种方

式实现的锁相环锁定的时候 A 信号与 B 信号之间的静态相位差为零。因为即使一个极小的相位差也会导致电容 C_p 上的电荷无穷无尽地积累。

PFD/电荷泵组合表现出一些非理想的特性，例如“死区”，电荷共享与注入等等。这些效应在[5, 9]中进行了描述。

图 8.19 带有电荷泵的 PFD

电荷泵锁相环 电荷泵锁相环(CPPLLs)用 PFD (或者鉴相器) 与电荷泵组合在一起代替在图 8.20 所示的鉴相器与低通滤波器的组合一般结构。如前所述，相比异或逻辑/低通滤波器，PFD 与电荷泵的组合有两个重要的优点：(1) 捕获范围仅仅由压控振荡器的输出频率范围所决定；(2) 如果忽略了失配与偏差静态的相位误差为零。在这一节中，我们研究这一类锁相环的特性，并与传统的类型进行比较。

(图 8.20 电荷泵锁相环)

电荷泵为 PFD 的输入信号中的静态相位误差提供了一个无限的增益。从另一个角度来看，PFD/电荷泵结构对相位阶跃的响应是一个线性的斜坡。这表明电路的开环传输函数在原点处包含了一个极点。再存在由压控振荡器所带来的另一个这样的极点，电荷泵锁相环不可能保持稳定。事实上，用 K_{PFD}/s 来表示 PFD/电荷泵的传输函数，我们注意到锁相环的闭环传输函数为

$$H(s) = \frac{K_{PFD} K_{VCO}}{1 + \frac{K_{PFD} K_{VCO}}{s}} \quad (8.25)$$

$$= \frac{K_{PFD} K_{VCO}}{s^2 + K_{PFD} K_{VCO}} \quad (8.26)$$

上式显示在 $\omega = \pm j\sqrt{K_{PFD} K_{VCO}}$ 处存在两个极点。为了避免不稳定性，在开环传输函数中必须再引入一个零点。这与一个正弦鉴相器和一个简单的低通滤波器的情况是不同的，在那种情况中即使没有零点，系统在理论上也是稳定的。电荷泵锁相环中的稳定零点可以通过为电荷泵电容串联一个电阻来实现 (图 8.21)。

图 8.21 在电荷泵中的加入一零点

为了进行小信号分析，我们注意到电荷泵的开关操作和在相位比较瞬间之间缺乏电荷泄放通路使得锁相环成为一个离散时间系统。但是，如果环路的带宽远远小于输入信号的频率，我们可以假设在每个输入的周期内，锁相环的状态改变很小[11]。使用离散时间参数的“平均值”，我们可以把环路作为连续时间系统来加以研究[11]。

假设环路开始时存在相位误差 $\phi_{in} - \phi_{out} = \phi_e$ ，那么给电容充电的平均电流为 $I\phi_e/2\pi$ ，而压控振荡器上控制电压变化量的平均值为

$$V_{cont}(s) = \frac{I\phi_e}{2\pi} \left(R + \frac{1}{C_p s} \right) \quad (8.27)$$

注意到 $\Phi_{out}(s) = V_{cont}(s) K_{VCO}/s$ ，我们得到了下面的闭环传输函数

$$H(s) = \frac{\frac{I}{2\pi C_p}(RC_p s + 1)K_{VCO}}{s^2 + \frac{I}{2\pi}K_{VCO}Rs + \frac{I}{2\pi C_p}K_{VCO}} \quad (8.28)$$

因此，系统为在 $\omega_z = -1/(RC_p)$ 处的一个零点所表征，并且

$$\omega_n = \sqrt{\frac{I}{2\pi C_p}K_{VCO}} \quad (8.29)$$

$$\zeta = \frac{R}{2} \sqrt{\frac{IC_p}{2\pi}K_{VCO}} \quad (8.30)$$

注意 ω_n 与电阻 R 无关。

从(8.24)式我们得到，系统的放电时间常数等于 $(\zeta\omega_n/2)^{-1} = [RIK_{VCO}/(8\pi)]^{-1}$ ，为与电容 C_p 无关的量。

在许多应用中，我们希望得到最大的环路带宽，这个值通常与 ω_n 成正比。尽管对一个用正弦鉴相器的锁相环， ω_n 与 ζ 不可能被同时最大化，式(8.29)和(8.30)表明，在电荷泵锁相环中，如果增大 I 或者 K_{VCO} 的话， ω_n 与 ζ 可以同时增大。但是，当环路带宽与输入信号的频率可以比较时，前面用到的连续时间近似不再成立，需要进行离散时间分析。通过这种分析，Gardner 推导了一个稳定性限制[11]，它可以化简为

$$\omega_n^2 < \frac{\omega_m^2}{\pi(RC_p\omega_m + \pi)} \quad (8.31)$$

表明 ω_n 存在着一个上限。这个关系式同时指出电阻 R 不可以无限制地增大[10]。在典型的设计中，为了保证稳定性，环路带宽大致被设计为输入信号频率的十分之一。在单端电荷泵中，即使环路已经锁定，串联在电容上的电阻也会在控制电压中引入“波纹”[11]。因为在每个相位比较瞬间， S_1 和 S_2 都要打开， I_1 与 I_2 之间的失配会通过电阻 R 溢出，在输出中引起跳变。而且 S_1 和 S_2 交迭电容的失配也会造成一个纯信号到输出的馈送。在频率合成器中，这种对压控振荡器频率的调制效应是尤其应该避免的。为了抑制波纹，在电荷泵和地之间可以连接第二个电容。这种改变在锁相环中引入了第三个极点，需要更深入地研究稳定性的问题。Gardner 为这种系统的稳定性提供了标准[11]。

8.2.4 I型和II型锁相环

我们关于锁相环的学习揭示了鉴相器与环路滤波器的组合在系统动力学中起了重要的作用。特别是，应用了鉴频器和RC滤波器的锁相环与那些使用三态PFD和电荷泵的锁相环在稳定性方面所体现出来的特性是完全不同的。这两者之间最基本的区别就是，前者的开环传输函数中只有一个位于原点的极点，而后者有两个位于原点的极点。这两种拓扑结构分别被称为“I型”和“II型”。

为了理解这两种锁相环的稳定性特性，我们来分别分析当环路增益 $K = K_{PD}K_{VCO}$ 发生变化时它们的根轨迹。I型锁相环（没有零点）的开环传输函数是 $K/[(1+s/\omega_{LPF})s]$ 。因此当 $K=0$ 时，闭环传输函数的极点开始于 $s_1=0$ 和 $s_2=-\omega_{LPF}$ [图 8.22(a)]。随着 K 的增加， s_1 和 s_2 沿着实轴彼此靠近，汇合于 $K = \omega_{LPF}/4$ 处。当 K 更高时，极点变成了虚数的，向着 $\pm\infty$ 移动。因为 $\zeta = \cos\theta$ ，所以我们会注意到 K 的增加会导致系统变得不稳定。

图 8.22 根轨迹图 (a) I型 (b) II型锁相环

对于 II型的电荷泵锁相环，开环传输函数等于 $I_p[R+1/(C_p s)]K_{VCO}/2\pi s$ 。当 $I_p K_{VCO}$ 从零开始增加的时候，两个极点从原点分开，在左半平面沿着一个圆形轨迹运动，最后在增益足够高，也就是 $IK_{VCO} = 8\pi/(R^2 C_p)$ 和 $\zeta = 1$ 时重新汇合于实轴。因此增加环路增益使锁相环变得更加稳定。

就像前面提到过的，II型锁相环通常因为将串联RC网络并联了一个电容，而引入了第三个极点。这时的根轨迹图与稳定性问题在[4, 11]中有所描述。

8.2.5 锁相环中的噪声

因为锁相环是处理相位信号的，所以它很容易受到相位噪声和抖动(jitter)的影响。在本书的范围内，我们将相位噪声当作是过量相位的一个随机分量，在 $x(t) = A\cos[\omega_c t + \phi_n(t)]$ 中表示为 $\phi_n(t)$ 。为了简明起见，我们用术语噪声来代表相位噪声。

如果输入信号或者构成锁相环的单元电路包含噪声，那么输出信号中也会受到噪声的干扰。通常说来，环路的所有元件，包括鉴相器，低通滤波器，压控振荡器和分频器都会产生噪声[15]。我们的目的是了解一个给定噪声源的频谱在传播到输出端的时候形状是如何变化的。我们来考察两个重要的情况：(1) 输入信号包含噪声，(2) 压控振荡器引入噪声。在每种情况中，我们找到感兴趣的噪声源到达锁相环输出端的传输函数。在单片集成电路实现中，压控振荡器的相位噪声通常要比其它元件的要大得多。

8.2.6 输入端的相位噪声

考虑图 8.23 中的锁相环，输入信号和输出信号分别为 $x(t) = A \sin[\omega_c t + \phi_{in}(t)]$ 和 $y(t) = B \sin[\omega_c t + \phi_{out}(t)]$ 。一个典型的 II 型二阶锁相环的传输函数 $\Phi_{out}(s)/\Phi_{in}(s)$ 为

$$H(s) = \frac{\omega_n^2 (1 + s/\omega_z)}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (8.32)$$

如果输入（过量）相位 $\phi_{in}(t)$ 不随着时间变化，也就是说，锁相环的输入信号为纯正弦曲线，那么 $s = 0$ 而 $H(s) = 1$ 。现在，假设 $\phi_{in}(t)$ 随着时间非常缓慢地变化，使得式(8.32)的分子和分母仍然很接近 ω_n^2 。那么 $H(s)$ 仍然很接近单位值，表明输出信号的相位（频率）会跟着输入信号的相位（频率）变化，这是锁相环作为一种跟踪系统的基本属性。

如果 $\phi_{in}(t)$ 以不断增加的速率变化会发生什么呢？如图 8.23 所示的情况中，公式(8.32)表明输出信号的过量相位 $\phi_{out}(t)$ 会下降，最终接近零，得到 $y(t) = B \sin \omega_c t$ 。换一句话说，当输入信号的过量相位或者频率快速变化的时候，锁相环不能成功地跟踪输入信号。

图 8.23 锁相环中从输入到输出的噪声传输函数

总括地说，输入信号的相位噪声频谱在通过锁相环出现在输出端的时候，被系统的具有低通传输函数的特征改变形状。

8.2.7 压控振荡器的相位噪声

如图 8.24 所示，压控振荡器的相位噪声可以表示为一个附加的分量 ϕ_{VCO} 。假设 ϕ_{VCO} 与 ϕ_{in} 是无关的，那么我们将 ϕ_{in} 设置为零，计算从 ϕ_{VCO} 到 ϕ_{out} 的传输函数⁴。注意 $\phi_{in}(t) = 0$ 代表输入信号的过量相位为零，而不是输入信号本身为零，也就是说我们必须在输入端施加严格的周期信号。

对于 II 型的二阶锁相环，当 $\phi_{in} = 0$ 时，我们得到

$$\frac{\Phi_{out}(s)}{\Phi_{VCO}(s)} = \frac{s^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (8.33)$$

⁴ 叠加原理适用于非相关信号源的能量

正如所期望的，这个传输函数与公式(8.32)具有相同的极点，但是它同时还在原点包含了两个零点，这使得其特性表现为一个高通滤波器。

位于原点的零点意味着 ϕ_{VCO} 缓慢变化时 ϕ_{out} 很小。这是因为，在锁定状态，压控振荡器的相位变化被鉴相器转换为电压，施加在压控振荡器的控制输入端来向着相反的方向积累相位。因为不论是电荷泵还是压控振荡器对于缓慢变化的信号都有近乎无限的增益，所以负反馈抑制了输出相位的变化。

图 8.24 锁相环中从压控振荡器到输出的噪声传输函数

从另一个观点来看，如果 ϕ_{VCO} 变化缓慢，锁相环可以被化简为图 8.25(a)和(b)的样子。因为位于负反馈回路上的理想积分器在它的输入端产生了一个“虚拟”地，所以当 ϕ_{VCO} 缓慢变化时 $\phi_{out} \approx 0$ 。现在假设 ϕ_{VCO} 的变化率增大。那么 K_{PD}/s 和 K_{VCO}/s 的值减小，因此环路增益减小，允许虚拟地经历显著的变化。

图 8.25 带有压控振荡器噪声的锁相环简化模型

从公式(8.33)，我们发现当 $s \rightarrow \infty$ 的时候， $\phi_{out} \rightarrow \phi_{VCO}$ 。这符合我们的预期，因为反馈回路对于 ϕ_{VCO} 的快速变化来说本质上是开放的。

总括地说，压控振荡器的相位噪声传输到输出端的时候，经历了一个高通滤波传输函数。因此，增加锁相环的带宽可以降低压控振荡器相位噪声的贡献。

8.2.8 倍频

频率合成器经常要求锁相环的输出频率为输入频率的倍数，一个锁相环可以像反馈电路放大电压一样来“放大”频率。如图 8.26(a)所示，为了放大输入信号，输出信号在反馈之前经过分压器降低。在锁相环中所关心的输出量为频率，所以在反馈回路中必须引入一个分频器(比如说，数字计数器)[图 8.26(b)]。从另一个角度来说，当回路锁定的时候， $\omega_F = \omega_{in}$ ，因此 $\omega_{out} = M\omega_{in}$ 。除率 M 也称为“模数”。

图 8.26 信号“放大器” (a) 一个电压放大器 (b) 一个锁相环

图 8.26 中所描述的类比对研究 $\div M$ 电路对锁相环行为的影响是十分有用的。就像对带有反馈的放大器时，环路增益被除以了 M ，因此前面所有的静态分析和动力学分析的结果都可以直接使用，只要用 K/M 来代替 K 。

有趣的是去注意到，图 8.26(b)中所示的倍频器同时也放大了输入信号的相位噪声。例如，

锁相环-3db 带宽中相位噪声分量的幅值，被乘以了一个近似为 M 的系数。

8.3 射频频率综合器结构

一个射频频率合成器产生由 $f_{\text{out}} = f_0 + kf_{\text{ch}}$ 给出的频率的输出信号，这里 f_0 是频率范围的下限， k 是从 0 到最大信道数之间变化的整数而 f_{ch} 是频率步长（也就是信道宽度）。例如，在 IS-54 的接受频带中， $f_0 = 869\text{MHz}$ ， $k = 0, \dots, 833$ ， $f_{\text{ch}} = 30\text{kHz}$ 。如图 8.1 所示， k 由一个数字输入信号来选择。在一个移动单元中，这个选择发生在当基站指定某一频率通道来接收或者发送信号的时候。

f_0 和 f_{ch} 的高精确性要求使得频率合成器中经常用到锁相环。因为锁相环处于锁定状态时，其（平均）输出频率与输入频率完全等同。

8.3.1 整数-N 结构

我们设计频率合成器时可以从图 8.26 所示的倍频电路开始。为了产生 $f_{\text{out}} = f_0 + kf_{\text{ch}}$ 的输出频率，我们可以使用一个具有可变模数的分频器。如图 8.27 所示，这样的拓扑结构产生了一个输出频率 $f_{\text{out}} = Mf_{\text{REF}}$ ，其中 M 以 1 的间隔从 M_L 变化到 M_H 。如果 Mf_{REF} 等于 $f_0 + kf_{\text{ch}}$ ，那么对于第一个通道 ($k=0$)，有 $M_L f_{\text{REF}} = f_0$ 。而且，对于第二个通道， $(M_L + 1)f_{\text{REF}} = f_0 + f_{\text{ch}}$ ，意味着 $f_{\text{ch}} = f_{\text{REF}}$ 。于是 $f_{\text{out}} = M_L f_{\text{REF}} + kf_{\text{REF}}$ 。这里很重要的一点是输入的参考频率必须等于信道宽度。

图 8.27 整数-N 频率合成器

图 8.27 中所用到的分频器必须可以提供一个由 $M = M_L + k$ ， $k = 0, 1, \dots, N$ 给出的可变模数。

图 8.28 给出了这种电路的一个例子，“脉冲吞咽分频器”。这个分频器由一个“预分频器”，一个“程序计数器”和一个“脉冲吞咽计数器”构成。在这里我们简单地描述一下这个电路的操作。首先让我们作三个观察：(1) 预分频器根据模数控制线的逻辑状态来将输入信号进行 $N+1$ 或者 N 分频。(2) 程序计数器总是将预分频器的输出信号进行 P 倍分频。(3) 而脉冲吞咽分频器将预分频器的输出信号进行 S 倍分频，这里 S 是由数字输入控制，可以从 1 到最大信道数变化的整数。这个计数器还有一个重置输入。我们将证明 $f_{\text{out}} = (NP + S)f_{\text{in}}$ 。当电路从重置态开始的时候，预分频器进行 $N+1$ 分频。预分频器的输出分别用程序计数器和脉冲吞咽计数器进行分频，直到后者的计数器“满”了，也就是它已经计数了 S 个脉冲。在这时候，即经过了主输入信号的 $(N+1)S$ 个周期之后，脉冲吞咽计数器改变模数控制信

号的状态，使得预分频器装置对 f_{in} 进行 N 分频。注意在这个变化之前，程序计数器一共已经计了 S 个数。在模数改变以后，预分频器和程序计数器继续分频，直到后者计数满了。因为程序计数器已经记录了 S 个脉冲，所以它还需要记录它的输入信号的 $P - S$ 个周期，因此也就是主输入信号的 $(P - S)N$ 个周期才会溢出。因此输出信号经过输入信号的 $(N + 1)S + (P - S)N = PN + S$ 个周期完成了一个周期。当脉冲吞咽计数器重置之后，系统重复整个操作过程。

图 8.28 脉冲吞咽分频器

整数- N 分频器的结构很简单，这使得它在过去的几十年中非常流行。在射频系统中，频率合成器通常被分为三个芯片：压控振荡器，双模数预分频器和一个组合模块。这个组合模块包括程序计数器，脉冲吞咽计数器，PFD 和电荷泵。作为电路中高速工作的部分，压控振荡器和预分频器通常用双极型硅工艺或者砷化镓工艺来生产，而电路的其它部分则用 CMOS 工艺来生产。注意，通常会在压控振荡器与预换算器之间插入一个缓冲器，以隔离后者的开关噪声对前者的影响。

但是，整数- N 分频器也受到几个缺点的困扰。下面就说明其中的一些问题。

参考信号毛刺 在图 8.27 所示的结构中，压控振荡器受到输入参考频率的调制，结果在载波附近产生了旁波。为了理解这种作用，我们回到 8.2.3 小节中的电荷泵锁相环。回想一下，即使输入相位差为零，在每个相位比较瞬时内，输出信号 UP 和 DOWN 会产生一个窄脉冲（图 8.29）。在理想情况下，这两个脉冲是大小相等，形状相反的，而且 S_1 和 S_2 的栅漏交迭电容是相等的。所以脉冲到节点 X 的馈送是完全可以抵消的。在实际情况下，这些都是不现实的，因此在每次进行相位比较的瞬间压控振荡器的控制电压都会经历一个有限的瞬变过程。而且，如图 8.30 所示，当 S_1 和 S_2 关闭的时候，他们的电荷注入失配在节点 X 引起一个误差阶跃，这导致压控振荡器的频率发生偏移。结果就是在下次相位比较的瞬间，压控振荡器的输出相位与输入的参考信号的相位不同，需要电荷泵产生一个矫正脉冲电流。因此，即使是在锁定状态下， V_x 也会在两个值之间切换。

图 8.29 电荷泵工作所引起的压控振荡器控制电压的周期性扰动

前面的机理以在模拟的帮助下进行量化。对于一个实际的鉴相器和电荷泵设计，由馈送和电荷注入所引起的系统的误差波形可以通过时域模拟得到。为了估算这个波形的影响，让我们假设控制线干扰表现为宽度为 Δt ，高度为 ΔV 的窄矩形脉冲（图 8.31）。用 $g(t)$ 来代表这个周期波形，我们可以得到压控振荡器的输出表达式为

$$v_{out} = V_0 \cos \left[\omega_{FR} t + K_{VCO} \int g(t) dt + K_{VCO} \int V_1 dt \right] \quad (8.34)$$

图 8.30 电荷泵电荷注入失配的影响

$g(t)$ 的傅立叶级数展开由下式给出

$$g(t) = \frac{\Delta V \Delta t}{T_{REF}} + \sum_{n \neq 0} a_n \cos(n\omega_{REF}t + \theta_n) \quad (8.35)$$

这里第一项代表直流分量，可以被合并并在 V_1 中。因此

$$\begin{aligned} v_{out}(t) \approx & V_0 \cos\left(\omega_{FR} + K_{VCO} \frac{\Delta V \Delta t}{T_{REF}} + K_{VCO} V_1\right)t \\ & - K_{VCO} \left[V_0 \sum_{n \neq 0} \frac{a_n}{n\omega_{REF}} \sin(n\omega_{REF}t + \theta_n) \right] \\ & \sin\left(\omega_{FR} + K_{VCO} \frac{\Delta V \Delta t}{T_{REF}} + K_{VCO} V_1\right)t \end{aligned} \quad (8.36)$$

表明旁波位于关于载波的 $\pm\omega_{REF}$ ，等等频率。这些旁波被称为“参考信号毛刺”。注意旁波的幅度与 ω_{REF} 成反比例关系。

图 8.31 电荷泵馈送引起的压控振荡器调制的估算

参考信号窜扰是一个难于解决的问题，特别是当 ω_{REF} 很小以致 $\omega_C \pm \omega_{REF}$ 落在有用的带宽之中的时候。典型的频率合成器通过在环路滤波器中用一个大的电容来减小 ΔV 或者最小化 K_{VCO} 以降低调制系数。在较低的电源电压下，后一种方式是行不通的，因为为了保证系统在不同的工艺和不同的温度下都能锁定，调谐范围必须足够的宽。

另一个抑制参考信号窜扰的方法是在环路滤波器和压控振荡器之间加入一个陷波滤波器。如图 8.32 所示，这种方法在传输函数中产生一个位于 ω_{REF} （如果需要的话，也可以在更高的谐波处）的零传输系数，来使得振荡器控制电压上的干扰最小化。但是，采用这种方法存在两个缺点：(1) 滤波器可能产生大量的噪声，除非它用的是片外的分立电容和电感。(2) 滤波器改变了环路传输函数，有可能使取得稳定态的过程变得糟糕，甚至不再稳定。在用有源器件实现这种方法时，这样一个滤波器的 $1/f$ 的噪声也常常带来麻烦，因为这种噪声能上变频到载波频率的频率范围（第七章）。

环路带宽 图 8.27 所示的整数- N 结构要求参考信号的频率 f_{REF} 等于信道宽度，也就是在 IS-54 中为 30kHz，在 GSM 中为 200kHz。现在回想在 8.2.3 小节中提到，为了保证稳定性要限制 II 型锁相环的带宽大约为 $f_{REF}/10$ 。因此对于 $f_{REF} = 30kHz$ 来说，环路的稳定时间常数可能像一毫秒一样长。在定量的分析结果之前，我们需要再次回顾锁定时间的概念。

图 8.32 用陷波滤波器来抑制参考信号窜扰

在许多锁相环中，锁定时间通过输出信号与输入信号之间的相位差来定义：在锁定前的瞬态情况接近结束时，这个相位差降低到一个可以接受的值时，环路就被认为处于锁定状态了。这种观点对于数据传输中的重定时和边沿对准确实是相当合适的。

另一方面，对于射频频率合成器来说，在瞬态情况下的输出参数中我们更关心频率而不是相位。因为信道之间的间隔频率非常小，因此在环路锁定之前收发器中的本地频率只能与他们的理想值有百万分之几的误差。如果振荡器的控制电压随着时间变化，那么输出信号的频率也即时变化（图 8.33）。因此收发器中的上变频函数和下变频函数并不准确地工作在所要的信道里，直到本地频率已经稳定。如图 8.34 所示，在接收路径中，这种效应畸变了下变频信道的中心频率，使得邻近信道的信号破坏了想接收的信号，而在发送路径中频率的误差导致上变频信号泄漏到相邻的信道中去。

图 8.33 在频率合成器稳定过程中压控振荡器的频率变化
图 8.34 频率合成器的稳定过程对接收通道和发送通道的影响

在图 8.27 所示的锁相结构中，除法器模数的每次改变都会引发环路的瞬态过程。有趣的是， M 的一个小改变会得到与输入信号中的微小频率改变相同的瞬态过程。这可以在图 8.35 所示的反馈系统的帮助下得到证明。这个反馈系统中，在时刻 $t = 0$ ， A 改变了一个小量 ε 。那么在 $t = 0$ 之后的输出为

$$Y(s) = \frac{H(s)}{1 + (A + \varepsilon)H(s)} X(s) \quad (8.37)$$

$$\approx \frac{H(s)}{1 + AH(s)} \cdot \frac{1}{1 + \varepsilon/A} X(s) \quad (8.38)$$

$$\approx \frac{H(s)}{1 + AH(s)} \left(1 - \frac{\varepsilon}{A}\right) X(s) \quad (8.39)$$

这表明， A 的改变等效于 $X(s)$ 乘以 $(1 - \varepsilon/A)$ ，而传输函数保持不变。因为在频率合成器中，在 $t = 0$ 时刻之前， $x(t)$ 为一个常数，所以乘以 $(1 - \varepsilon/A)$ 表现为一个在 $t = 0$ 时刻从 $x(t)$ 到 $(1 - \varepsilon/A)x(t)$ 的阶越。

图 8.35 反馈系数发生微小变化的影响
图 8.36 频率合成器跳变的最坏情况

通过上面的分析我们得知，当除法器的模数发生变化时，环路就表现为输入信号的阶跃响应，需要一个有限的时间来稳定频率以达到与它的最终值相差可以接受的余量的值。如图 8.36 所示，在最坏的情况下，频率合成器的输出频率必须从第一个通道 $(NP + 1)f_{REF}$ 变到最后一个通道 $(NP + S)f_{REF}$ ，或者反之亦然。假设除法器的模数从 M 变到 $M + k$ ，这里 $k \ll M$ 。那么如何估算稳定时间呢？采用上面的符号，我们写下 $A = 1/M$ 和

$A + \varepsilon = 1/(M + k)$ 。这等于参考频率从 f_{REF} 改变到 $f_{REF}(1 + k/M)$ 。为环路的输入频率和输出频率重写公式(8.24)，我们得到

$$f_{out}(t) = Mf_{REF} + kf_{REF} \left[1 - \frac{1}{\sqrt{1-\zeta^2}} \exp(-\zeta\omega_n t) \sin\left(\omega_n \sqrt{1-\zeta^2} t + \sin^{-1} \sqrt{1-\zeta^2}\right) \right] u(t) \quad (8.40)$$

因为输出频率的最终结果为 $(M + k)f_{REF}$ ，我们必须计算 f_{out} 进入到 $(1 \pm \alpha)(M + k)f_{REF}$ 所需的时间，这里 α 为稳定的精确度。

$$(1 \pm \alpha)(M + k)f_{REF} = Mf_{REF} + kf_{REF} \left[1 - \frac{1}{\sqrt{1-\zeta^2}} \exp(-\zeta\omega_n t_s) \sin\left(\omega_n \sqrt{1-\zeta^2} t_s + \sin^{-1} \sqrt{1-\zeta^2}\right) \right] \quad (8.41)$$

稳定的充分条件是指数的包络范围衰减到一个很小的值

$$\pm \alpha(M + k) = \frac{k}{\sqrt{1-\zeta^2}} \exp(-\zeta\omega_n t_s) \quad (8.42)$$

得出

$$t_s \approx \frac{1}{\zeta\omega_n} \ln \frac{k}{M|\alpha|\sqrt{1-\zeta^2}} \quad (8.43)$$

例如，考虑一个典型的频率合成器，其 $f_{REF} = 200\text{kHz}$ ， $Mf_{REF} = 900\text{MHz}$ ， $\zeta = \sqrt{2}/2$ ，

$k = 128$ ，对于 10ppm（百万分之一）的稳定精度，上面的公式得到 $t_s \approx 8.3/(\zeta\omega_n)$ 。也就是

说环路需要 8.3 个时间常数来稳定。在实际情况中，稳定时间要长于公式(8.43)所给出的值。回顾一下 8.2.3 节，大多数电荷泵锁相环需要在 RC 网络上再并联第二个电容来抑制压控振荡器控制线上扰动，尤其是在需要最小化参考信号毛刺的情况下。因此，这是一个三阶系统，通常需要更多的时间来稳定到更高的精密。由于这个原因，计算环路稳定时间需要进行更精确的模拟。

相位噪声 有限的环路带宽的另外一个缺点是在输出端更高的邻近相位噪声。如同 8.2.5 节中所解释的那样，只有位于环路带宽中的振荡器相位噪声会被反馈减小。例如，如果 GMS 频率合成器的环路带宽等于 20kHz，那么频率偏移超过了几千赫兹的相位噪声分量几乎得不到衰减。这在 MOS 实现方式中是一个很严重的问题，因为压控振荡器的上变频 $1/f$ 噪声在频率偏移超过几百千赫兹的时候相当明显。

典型的射频合成器中需要的大分频比率也加重了环路带宽中参考信号和鉴相器噪声的影响。例如，在 IS-54 系统中，为了用 30kHz 的参考信号产生 900MHz 的输出，大约需要 30000 的分频比率。结果就是在很小的频率的偏移情况下，参考信号和鉴相器的噪声底限增加了 $20 \log 30000 \approx 90\text{dB}$ [16]。

8.3.2 分数-N 结构

在整数-N 结构中，由于输入的参考频率必须等于信道宽度，所以环路增益受到了限制。而这又是因为输出频率只能改变 f_{REF} 的整数倍这个特性。另一方面，在“分数-N”频率合成器中，输出信号的频率可以为输入信号频率的分数倍，使得后者可以远远高于信道宽度。在研究分数-N 结构之前，我们来做一个观察。假设如图 8.37(a)所示，在一个频率为 f_1 的周期信号 $x(t)$ 中，每隔 T_p 秒移去一个脉冲信号。结果得到的波形 $y(t)$ 在每 T_p 秒中有 $f_1 T_p - 1$ 个脉冲，也就是说， $y(t)$ 拥有一个“平均”频率等于 $f_1 - 1/T_p$ 。这种方法可以以非常小的步长改变信号的平均频率。实际上， $y(t)$ 可以看作是 $x(t)$ 与一个周期为 T_p 的方波信号的乘积[图 8.37(b)]，带有位于频率 $f_1 \pm k/T_p$ 的旁波信号。然而，移去脉冲的想法在小步长的频率合成器中仍然是十分有用的。

图 8.37 (a) 在一个周期波形中周期性地移去脉冲。(b) 脉冲的移去可看作为乘以 $r(t)$

图 8.38 中显示了一个简单的分数-N 锁相环结构。除了 PFD，低通滤波器和压控振荡器，环路中还包括一个脉冲消除器，脉冲消除器是一个能够按照移除指令的要求来移除输入信号中的脉冲的电路模块。因为在锁定的情况下，提供给鉴相器的两个频率必须相等，所以脉冲移除器输出信号的平均频率必须等于 f_{REF} 。因此 $f_{out} = f_{REF} + 1/T_p$ ，这里 T_p 是移除信号被应用的周期。注意因为频率 $f_p = 1/T_p$ 可以通过 f_{REF} 的简单分频得到，所以 f_{out} 可以改变一个 f_{REF} 的分数倍的量。由石英晶振提供的 f_{REF} 通常被限制为几十兆赫兹。因此，如图 8.38(b)所示，分数-N 环路中需要在反馈中加入一个除法器来得到高的输出频率。

尽管初始的分数-N 频率合成器技术是基于脉冲移除概念的[17]，但是这种结构上现在是通过一种不同的原理来实现的。如图 8.39 所示，频率合成器中用一个双模的预变频器来代替了图 8.38(b)所示的脉冲移除器和除法器。如果预变频器将 A 个输出脉冲（来自于压控振荡器）进行 N 分频， B 个脉冲进行 $N+1$ 分频，那么等效的分频比率为 $(A+B)/[A/N + B/(N+1)]$ 。这个值可以通过选择合适的 A 和 B 来在 N 和 $N+1$ 之间以非常小的步长变化。这样得到的模数有时候记为 $N.f$ ，其中的点代表模数中整数部分 N 与分数部分 f 之间的十进制小数点。

图 8.39 使用双模除法器的分数-N 频率合成器

作为一个例子，我们来考虑图 8.40 所示的电路，这里 $f_{REF} = 1\text{MHz}$ 而 $N = 10$ 。我们假设预

变频器将参考信号的 9 个周期进行 10 倍分频而将参考信号的 1 个周期进行 11 倍分频。因此当输入信号产生 10 个脉冲的时候，输出脉冲的总数等于 $9 \times 10 + 11 = 101$ 。换一句话说，分频比例等于 10.1， $f_{out} = 10.1\text{MHz}$

图 8.40 分数-N 频率合成器例子

当 f_{REF} 的频率范围为几十兆赫兹的时候，一个分数-N 频率合成器的环路带宽可以高达几兆赫兹，这样就能够获得一个快速的锁定瞬态特性，同时抑制压控振荡器的邻近相位噪声。而且，较小的分频比率减弱了参考信号和鉴相器相位噪声的影响。

分数毛刺 分数-N 频率合成器具有一个非常严重的缺点：“分数毛刺齿”。为了理解这个问题，让我们看一下图 8.40 中鉴相器的输入和输出信号。假设压控振荡器与低通滤波器是断开的，而它的控制电压使得 $f_{out} = 10.1\text{MHz}$ 可以简化我们的分析。如图 8.41 所示，分频器信号的前九个周期的每一个周期都略短于参考信号的周期。结果，参考信号与反馈信号之间的相位差在 f_{REF} 的每一个周期中都不断增加，直到 11 倍分频发生的时候回复到零。因此，鉴相器产生一个不断增宽的脉冲信号，导致低通滤波器的输出产生一个锯齿波形。

图 8.41 在一个分数-N 频率合成器中非均匀的瞬时频率的影响

通过上面的观察，我们得到结论：如果压控振荡器的输出频率要等于 $(N + \alpha)f_{REF}$ （在图 8.40 中 $\alpha = 0.1$ ），那么低通滤波器的输出为一个周期为 $1/(\alpha f_{REF})$ 的周期性三角波。如果环路是闭合的，那么这个波形将会调制压控振荡器，在与中心频率相差 αf_{REF} ， $2\alpha f_{REF}$ 等等的频率上产生旁波信号。这样的旁波信号称为分数毛刺。从另一个角度来看，因为反馈信号不是一个严格的周期函数，而是包含有不可忽略的旁波信号，因此鉴相器的混频作用会将旁波信号变换到零频附近。有意思的是如果环路是闭合的，那么低频滤波器的输出信号将不再是线性的三角波形：当电压增加时，压控振荡器的周期减小，因此 f_{REF} 与 f_{out} 之间的相位差比线性的时间函数增加得更快。

分数毛刺的是一个严重的问题。因为当参考信号与反馈信号之间的相位差增加到一个不可忽略的数值时，低通滤波器的输出信号的幅值也相当的大，使得分数毛刺的幅值仅仅比载波信号低 20db 到 30db。因此，人们发明了很多种不同的方法来抑制毛刺[17, 2]。

回到图 8.41, 我们注意到每个参考信号周期中相位差的增加是相当确定的。如果 $f_{out} = (N + \alpha)f_{REF}$ ，那么反馈信号的周期等于 $1/[(N + \alpha)f_{REF}]$ ，因此相位增加为 $\alpha/[(N + \alpha)f_{REF}]$ （以秒为单位）。换一句话说，在每个相位比较的瞬间，电荷泵产生的电流脉冲宽度都会增加 $\alpha/[(N + \alpha)f_{REF}]$ 秒。因此，如果另一个具有相同宽度，相反符号的电流脉冲链注入到低通滤波器中，压控振荡器控制线上的干扰信号会变得最小。这被称为“分

数补偿”。

前面的分析可能会使人们得出这样的结论，即补偿电流可以简单地通过由同一个鉴相器所驱动的另一电荷泵产生。然而，理解分数补偿的最终目的是很重要的。事实上，读者可能会觉得奇怪：当注入到滤波器中的电流彼此精确抵消时，相位比较要达到什么目的呢？如果的确是如此，选择一个高的参考信号频率，亦即高的比较速率，将不会降低压控振荡器的邻近频率上的噪声。这样，补偿电流必须由一个稳定的低噪声的电流源提供，而不是环路中的鉴相器输出信号，以至于仅仅去消除电荷泵输出中的带确定性的分量。

分数补偿的主要缺点是由于失配造成的结果不精确。因为补偿电流的幅度与宽度与主电荷泵产生的电流有着一个一定大小的失配，所以在压控振荡器的控制线上残留有干扰信号。因为这个原因，也许需要用到外部的调节机制[18]。

另一个抑制分数毛刺的方法是随机地选择模数，也就是说分频比率的平均值仍为 $N + \alpha$ ，但是单个的分频系数仅仅发生在一个很短的时间内。这种技术实际上是将系统的分数毛刺转换为随机噪声（图 8.42）。这种想法可以更进一步发挥，来改变结果中噪声频谱的形状，使得它的大部分能量出现在偏移较大的频率范围上（图 8.43）。这样，在载波附近的噪声信号变得足够地小，而位于较大偏移的频率上的噪声可以在反馈信号通过鉴相器变为直流信号时被低通滤波器所抑制。

图 8.42 随机化模数控制信号以抑制毛刺

在上边的方法中所用到的噪声整形函数可以使用 $\Sigma - \Delta$ 调制来实现[19]。如图 8.44 所示，调制器产生一个二进制数据流（在双模分频器的情况中），这个流代表了伴随有量化噪声的确定的平均值。我们现在来作一个简单的分析以便更好地理解这个工作过程[19]。

图 8.43 模数控制信号中的噪声形状整形

假设分频器有两种模数 N 和 $N+1$ ，但必须提供一个等于 $N + \alpha$ 的模数。通过由 $\Sigma - \Delta$ 调制器产生的二进制模数控制信号 $b(t)$ ，分频器的即时分频比率可以写作 $N + b(t)$ ，这里 $b(t)$ 为 0 或者 1。因此 $x_F(t)$ 的即时频率为 $f_F(t) = f_{out} / [N + b(t)]$ 。比特流 $b(t)$ 可以被分解为一个等于 α 的平均值和量化噪声 $q(t)$ ，因此 $f_F(t) = f_{out} / [N + \alpha + q(t)]$ 。由此得到 $f_F(t)$ 的量化噪声等于

$$n_f(t) = f_F(t) - \frac{f_{out}}{N + \alpha} \quad (8.44)$$

$$= -\frac{f_{out}}{N + \alpha} \cdot \frac{q(t)}{N + \alpha + q(t)} \quad (8.45)$$

假设 $q(t)/(N + \alpha) \ll 1$ ， $\alpha \ll N$ ，我们得到噪声的功率谱密度为

$$S_{nf}(f) = \frac{f_{out}^2}{(N + \alpha)^2} \frac{|Q(f)|^2}{N^2} \quad (8.46)$$

这里 $Q(f)$ 是 $q(t)$ 的谱函数[19]。因此，在反馈信号频率上的量化噪声与 $Q(f)$ 具有相同的

谱形状。 $\Sigma-\Delta$ 调制器产生一个确定形状的 $Q(f)$ ，将噪声集中在较高的频率上。

[20, 21]中描述了另外一种分数- N 技术。

图 8.44 使用 $\Sigma-\Delta$ 调制器来改变噪声的形状

8.3.3 双环路结构

在整数- N 锁相环频率合成器中信道宽度与参考频率之间的关系可以通过使用两个或者更多环路来改变。在这节中，我们将研究两种双环路技术和他们设计之中的一些问题。

图 8.45 (a) 双环路频率合成器 (b) (a)的实现方式

产生精细频率步长的一种直接了当的方法是在一个固定的高频率上加上一个可变的低频信号。如图 8.45 所示，这种技术使用 PLL_1 来通过参考频率 f_{REF1} 产生一个频率为 f_c 的载波信号，使用 PLL_2 来产生频率增量和，其每一个频率增量等于 f_{REF2} 。改变 PLL_2 的分频比率也就可以在输出信号中得到精细的频率步长。注意 f_{REF1} 的频率可以是几十兆赫兹。两个频率的加和是通过一个单边带混频器实现的（第七章）。

这种结构相对于单环路整数- N 技术的主要优势是 VCO_1 的邻近中心相位噪声被较宽的环路带宽所抑制，因为 f_{REF1} 的频率比信道宽度要大得多。 VCO_2 的相位噪声可以比 VCO_1 低得多，因为前者是工作在一个低得多的频率上。通过第七章中对相位噪声的分析，在给定频率偏移上对于相同的功率损耗时， VCO_2 的相位噪声近似的等于 VCO_1 相位噪声的 $(Mf_{REF2})^2 / f_c^2$ 倍。

图 8.45 所示的这种结构的主要缺点是需要一个精确的单边带混频器。就像第七章中所解释的一样，这需要在 PLL_1 和 PLL_2 中同时产生精确的正交的相位，同时又要对相加的频率之一保持低的谐波失真。因此，很难保证失配和非线性所引起的旁波信号比载波信号低 60dB 到 70dB。

另一个问题是 VCO_2 的频率必须为一个很大的系数所改变：这个系数等于在图 8.45 实现中的通道数。结果就是压控振荡器必须有很宽的调谐范围。不仅如此，当信道（也就分频比率）发生改变时， PLL_2 的环路增益和因此阻尼系数变化很大。如果 f_2 的偏移很大，也就是

$f_2 = f_0 + kf_{REF2}$ ，这个缺点可以减轻，这里 f_0 是一个固定的频率。

另一种双环路结构如图 8.46(a)所示。与图 8.45 所示的频率合成器成对照，在这个电路中单

边混频器被放在了反馈回路中。因此如果混频器在 f_{out} 与 f_2 产生了差值，那么我们得到

$f_{out} = Mf_{REF1} + f_2$ 。这种结构一个令人感兴趣的特性是，如果在分频之后，由单边混频器产生的不希望有的旁波信号相对于有用旁波信号的频率偏移大于低通滤波器的带宽，那么它们会被抑制。这在图 8.46(b)中给予描述。

图 8.46 (a) 在反馈中带有单边混频器的双环路结构 (b) 旁波信号的抑制

8.3.4 直接数字频率综合

迄今为止我们对频率合成器技术的研究都假设在系统中应用相位锁定原理，来建立生成频率与参考频率之间的精确关系。直接数字频率合成器(DDS)是实现相同目标的另一种方法，相比相位锁定技术有着特定的优点和缺点。

直接数字频率合成器中的基本思想是，在数字域内产生一个信号，然后应用数字、模拟变换并进行滤波，最后在模拟域内重建波形。为了理解这种原理的工作方式，首先来考虑一个如图 8.47(a)所示的简单电路。一个计数器以单位值为步长在 0 到 N 之间计数，产生一个数字的斜坡波形。用计数器产生的每一个数字在只读存储器中选择一个对应的正弦波采样信号。这个结果接下来被转换成模拟信号并进行滤波抑制高频分量。

图 8.47 (a) 一个正弦波形的简单数字频率合成器，
(b) 通过减少采样点来增加输出信号的频率

现在假设必须保持固定的时钟频率 f_{CK} 而改变输出信号的频率。我们假设在每个正弦周期内计数器的采样点数减少，那么输出信号的频率将会变高，或者反之亦然[图 8.47(b)]。如果计数器的输出增量为可编程的步长 P ，这就是可能的。这样的计数器可以用一个累加器[图 8.48]来实现。这里用一个并行输入并行输出的 M 比特寄存器来驱动反馈环路中的加法器。在每一个时钟周期中，一个等于 P 的数值被加在了 Y_R 上，结果被保存在寄存器中，也就是说 $X_R(k) = Y_R(k-1) + P$ 。这种关系一直保持到寄存器溢出，在溢出时 P 的一部分作为新的 Y_R 的增量出现，也就是说 $X_R(k) = Y_R(k-1) + P$ 对 2^M 取模。

图 8.48 应用累加器的直接数字频率合成器

让我们考虑一个例子来更好地理解上边的工作过程。图 8.49 所示的是一个 $M = 3$ 的累加器和只读存储器的输出。如果增量 P 等于 1[图 8.49(a)]，那么寄存器的输出从 000 变到 111 时，一个完整的正弦波形从只读存储器中提取出来。换句话说，每个时钟周期中，输出信号的相位增加了 $2\pi/8$ 。现在如果 P 增加到 2[图 8.49(b)]，累加器在 110 之后就溢出了，只读存储器中的正弦波形中采样点每隔一个被读出一个，输出信号的相位变化的步长为 $2\pi/4$ 。当 $P = 3$ 时，累加器输出从 000 开始，在第一、第二和第三各个周期中分别在 110，111 和 101 处溢出。这样，如图 8.49(c)所示，三个正弦波形周期由八个等间隔采样生成。最后，对于 $P = 4$ ，四个正弦信号周期由奈奎斯特比率采样生成[图 8.49(d)]。

从上面的分析，我们得知图 8.48 中生成的正弦信号周期为

$$f_{\text{out}} = P \frac{f_{\text{CK}}}{2^M} \quad (8.47)$$

注意 $P/2^M$ 并不需要是整数。 f_{out} 的最小值和最大值分别为 $f_{\text{CK}}/2^M$ 和 $f_{\text{CK}}/2$ 。

图 8.49 3 比特直接数字频率合成器的波形

应用在只读存储器上的 M 比特的字为正弦波信号选择了一个幅值。只读存储器的输出信号的宽度应该是多少位的呢？因为只读存储器的输出被近似为振幅，所以比特的数量决定了重建的正弦信号中的“量化误差”。有趣的是注意到直接数字频率合成器中 f_{CK} 与 f_{out} 之间的

比值是一个其值为 $2^M/P$ 的有理数，这使得量化误差作为一个周期性的附加量出现，而不是

随机噪声。实际上，一个等于 $2^M/f_{\text{CK}}$ 的周期中容纳了 2^M 个时钟周期和 P 个输出周期。因此，对于奇数的 P ，量化噪声的周期为 $2^M/f_{\text{CK}}$ 而对于偶数的 P ，量化噪声的周期等于

$2^{M-j}/f_{\text{CK}}$ 。这里 j 代表 P 对应的 2 的指数。在输出频谱中，输出结果的误差波形和它的谐波表现为毛刺。可以得到，最坏情况下，这些毛刺的功率量与信号功率的关系近似等于

$2^{-2(k-1)}/3$ [22]。这里假设 $f_{\text{CK}} = 2f_{\text{out}}$ 。对于 $k = 12$ 比特时，锯齿大约比载波信号低 71db。

前面的讨论表明，在典型的射频应用中，只读存储器的输出信号至少要有 10 到 12 比特的分辨率，这给出了只读存储器一个维度的下限。只读存储器另一个维度的下限由它所产生的正弦信号所要求的相位步数所决定。公式 8.47 显示，在直接数字频率合成器中只要增加 M 的值就可以在输出频率上得到任意小的步长——这是直接数字频率合成器的一个重要特性。尽管可以选择一个宽度相对大的累加器，只读存储器可能变得大得难以置信。例如，如果累加器和只读存储器的输出字长分别为 16 比特和 10 比特时，那么只读存储器需要

$2^{16} \times 10 \approx 6.55 \times 10^5$ 个单元。由于这个原因，可能将累加器依然设计成输出字宽大以得到精细的频率步长，但是仅仅将这个词中有限个最高有效位施加到只读存储器。如果只读存储器的相位步长不像累加器那样小的话，一种“相位截断误差”会破坏输出的正弦波形。这种误差也是周期性的，最坏情况下这种毛刺的功率与载波信号的比为 $(2^{B-1} \pi)^2$ 。这里 B 是累加器输出信号与只读存储器输入信号之间的宽度差 [3]。

有很多不同的技术被设计来减小只读存储器的大小。例如，可以仅仅储存正弦波信号的四分之一周期，因为另外三个四分之一周期都可以垂直与水平翻转得到。[23] 中描述了其它一些只读存储器压缩技术。

相对于锁相环频率合成器，直接数字频率合成器提供了很多种优点：(1) 避免了使用模拟的压控振荡器，直接数字频率合成器能够得到很小的相位噪声——仅仅等于时钟信号的相位噪声。因为时钟信号的频率不需要改变，所以它可以通过一个使用晶体振荡器的宽带锁相环来产生。因此时钟信号的相位噪声可以相当的低。(2) 直接数字频率合成器提供精细的频率步长。增加累加器的字长可以相应地减小频率增量的值，虽然代价是增加了复杂度。(3) 直接数字频率合成器比锁相环有着快得多的信道转换速度，因为它不需要模拟的反馈环路。(4)

直接数字频率合成器可以在输出中提供连续相位的通道转换,这在某些调制方案中是一个非常重要的特性。(5) 直接数字频率合成器允许在数字域中对输出信号直接进行调制。

尽管有了上面提到的特性,直接数字频率合成器存在着几个缺点,这阻止了它在射频范围内的广泛使用。首要问题就是速度:根据奈奎斯特采样理论,时钟信号的频率必须至少是所需的输出信号频率的两倍,这意味着在 900-MHz 的系统中为 1.8GHz。实际上,为了降低对低通滤波器抑制作用的要求,典型的 f_{CK} 为 f_{out} 的三到四倍。在今天的超大规模集成电路技术中,很难在这样的速度下实现图 8.48 的操作过程,特别是在功率消耗有严格限制的情况下。即使数字部分可以达到相应的复杂度和功率消耗,数字模拟转换仍然是速度的瓶颈。DAC(数模转换)在稳定时间,谐波失真,伪响应和能量消耗之间需做的折衷阻碍了数字模拟转换电路在射频范围内的应用。

直接数字频率合成器的低相位噪声使得在如图 8.46 和图 8.46 所示的双环路结构中的低频生成器中使用它来代替低频锁相环非常的具有吸引力。尽管如此,在这种情况下仍然需要注意两个问题。首先,如果高频的压控振荡器与直接数字频率合成器电路位于同一个芯片中,那么由于累加器和只读存储器所造成的衬底噪声和电源噪声可能会严重恶化压控振荡器的输出信号。其次,在单边混频器中的相位失配与增益失配仍然会产生相对大的不希望有的旁波信号。

8.4 分频器

在对频率合成器与正交的生成器技术的研究中,我们已经看到了需要用到分频器,也称为预分频器(prescaler)。除了速度与功率的损耗,分频器的相位噪声也是需要注意的,因为它会使破坏频率合成器中的反馈信号。尽管分频器的相位噪声已经在一定程度得到了研究[24],但是在单片集成电路实现中需要做更多的工作来对它的影响进行分析。

在这一部分中,我们将会描述多种射频系统中常用的分频器拓扑。

8.4.1 二分频电路

二分频电路(DTCs)广泛地用于产生正交的输出。此外,因为它们可以得到比其它分频比率的分频器更高的速度,所以可以在锁相环中的压控振荡器后面连接一个二分频电路,将频率降低到小步长可编程分频器能够工作的范围内。如图 8.50 所示,当输出频率可与所采用的工艺技术的最高速度相比拟时,也就是说可编程分频器不能工作在 f_{out} 而二分频电路可以时,这种方法相当有用。尽管如此,这样做的缺点是 f_{REF} 必须变为原来的一半,以得到与无二分频电路的环路相同的频率步长。

图 8.50 在反馈中带有二分频电路的锁相环

如图 8.51(a)所示,一个二分频电路可以用两个锁存器构成的负反馈环路来实现。锁存器的实现依赖于可用的晶体管类型,但是由一个差分对和一个正反馈对构成的电流控制拓扑结构无论是在双极型工艺还是 CMOS 工艺[图 8.51(b)]中都可以得到很高的速度。适当地选择这种结构下的晶体管尺寸可以得到在千兆赫速率下的合理的速度功率折衷。

图 8.51(a)中的分频器结构仅仅当 CK 与 \overline{CK} 严格互补,并且两个锁存器完美地匹配的时候

才会在 X 和 Y 产生正交的相位。典型器件的不匹配性将会产生像 5° 那样大的相位失调。此外, 如果 CK 与 \overline{CK} 不是严格差分的, 那么还会产生附加的相位失调。一个典型的例子发生在用一个带有片外谐振回路的振荡器来驱动分频器时。在这种情况下, 振荡器通常输出一个单端输出或者一个严重失衡的差分信号 (图 8.52)。在高速情况下, 很难通过额外的差分对来将这样的输出信号进行平衡。因此加在分频器上的 CK 与 \overline{CK} 的相位仍然会表现出某些相位失配。

图 8.51 (a) 二分频电路 (b) 单个锁存器的实现方式

图 8.52 具有非严格差分输出的振荡器

MOS 电流转换的锁存器也可以用图 8.53 所示的结构实现。然而, 被时钟控制的晶体管的漏极电流不是确定地被定义, 导致不可预知的输出摆动。

图 8.53 另一种 MOS 锁存器技术

高速 CMOS 二分频电路也可以由动态锁存器构成。图 8.54 给出了两个例子。在图 8.54(a) 所示的电路中, 前两个 CMOS 反相器作为动态锁存器工作, 由 CK 与 \overline{CK} 控制。而第三个反相器为负反馈环路提供所需的整体反相。图 8.54(b) 是一个基于全单相时钟 (TSPC, true single-phase clocking) 结构的分频器 [25], 可以得到很高的速度。这两个电路的缺点就是缺少精确互补的或者正交的输出信号。其它的 CMOS 分频器技术在 [26, 27] 中有所描述。

图 8.54 动态 CMOS 分频器 (a) 使用反相器 (b) 使用 TSPC

一种由 Miller 首先提出的高速二分频方法在图 8.55 中给出 [28]。通过在反馈回路中使用一个混频器和一个低通滤波器, 电路按照下面的过程工作。在输入信号与输出信号相乘之后, 混频器产生 $f_{in} + f_{out}$ 和 $f_{in} - f_{out}$ 的频率分量。如果前者被低通滤波器所抑制, 而后者没有,

那么 $f_{in} - f_{out} = f_{out}$, 也就是 $f_{out} = f_{in}/2$ 。这个反馈环路的简单结构使得这种技术可以工

作在超过它所用的器件的 f_T 的一半的速度上 [29], 这通常是在所有的分频器结构中能够获得的最高的速率。然而, 普遍认为 Miller 分频器受到真实相位噪声的困扰。

图 8.55 Miller 分频器

8.4.2 双模分频器

绝大多数的锁相环频率合成器用到了双模或者多模的分频器。这种电路以一种对应于控制信号的模数对输入信号进行分频。

一种常用的双模分频器是 $2/3$ 分频电路。我们首先来考虑一个简单的 $\div 3$ 电路 [图 8.56(a)]。

这个分频器同时使用两个主从式 D 触发器和一个与门来生成三种状态: $Q_1\overline{Q_2} = 01, 10, 11$ 。

状态 $Q_1\overline{Q_2} = 00$ 不会发生 (除非是在电路启动时), 因为这需要 $\overline{Q_2}$ 和 G 的前一个状态分别为零和一, 而这是一个不可能的状态。

图 8.56 (a) 三分频电路 (b) $2/3$ 分频电路

为了将图 8.56(a)所示的拓扑变成 $\div 2/3$ 电路, 我们简单地在第一个触发器和与门之间插入一个或门来控制 Q_1 。如图 8.56(b)所示, 当 MC 为高电平时, 这个分频器被构成为一个 $\div 2$ 电路, 而当 MC 为低电平时, 则是 $\div 3$ 电路。

三分频电路通常说来比它的二分频对应电路要慢得多。例如图 8.56(b)所示的电路, 在 Q_2 必须发生变化的时钟沿之后, 必须有足够的时间来允许下一个时钟翻转到来之前 G_1 与输入级的 FF_2 的延时。这可以在图 8.57 中看到, 这里给出了电路一部分的晶体管实现。信号的延时几乎是 $\div 2$ 电路的两倍。而且, 在图 8.56(a)中 FF_2 的输出必须要同时驱动 G_1 和 FF_1 的输入电容。由于这些原因, $\div 3$ 电路的最高速率通常是 $\div 2$ 电路的一半。

图 8.57 图 8.56(b)中分频器部分电路的晶体管实现

具有其他模数的双模分频器可以以 $\div 2/3$ 或者 $\div 3/4$ 电路为核心进行设计。例如, 一个 $\div 15/16$ 预分频器可以按照图 8.58 所示的方式实现。这里 FF_1 , FF_2 , G_1 和 G_2 构成了一个同步 $\div 3/4$ 电路, 当 MC 为高时, 将时钟信号进行四分频, 当 MC 与 MF 同时为低时, 将时钟信号进行三分频。由 FF_3 , FF_4 和 G_3 组成的异步部分将 FF_2 的输出进行四分频, 并在 $\overline{Q_3 Q_4} = 11$ 的时候将 MF 驱动为高电平。这样, 如果 MC 为高, 那么整个电路将输入信号进行 16 分频。如果 MC 为低, 那么电路将避免状态 0000。因为如果 $\overline{Q_3 Q_4} = 00$, 那么 $\div 3/4$ 电路只通过三种状态: 01, 10, 11。注意图 8.58 所示电路中信号的关键路径包括 G_1 和 G_2 , 这使得这个电路比图 8.56(b)所示的分频器还要慢。

图 8.58 15/16 分频电路

在图 8.58 所示电路中同时使用同步与异步部分的一个重要考虑是当电路进行 15 分频时潜在的竞争冒险情况。为了理解这个问题, 首先假设 FF_3 与 FF_4 在他们的时钟上升沿改变其输出状态。如果 MC 为低, 那么电路继续进行 16 分频, 也就是说 $Q_1 \overline{Q_2}$ 通过 01, 11, 10, 00 周期直到 $\overline{Q_3}$ 和 $\overline{Q_4}$ 同时为低。如图 8.59(a)所示, $Q_1 \overline{Q_2}$ 在 10 态之后将会跳过 00 态。因为从 $\overline{Q_3}$ 变低开始直到 $Q_1 \overline{Q_2}$ 跳过一个态的时间内, 已经过了三个 CK_{in} 周期, 所以通过 FF_3 和 G_3 的传播延时不需要少于 CK_{in} 的一个周期。

图 8.59 当 FF_3 和 FF_4 工作在时钟的 (a) 上升沿 (b) 下降沿时,

图 8.58 所示的电路的延时预算

现在考虑 FF_3 与 FF_4 在他们的时钟下降沿改变其输出状态的情况。然后, 如图 8.59(b)所示, 在 $\overline{Q_3Q_4}$ 下降到 00 之后, $\div 3/4$ 电路必须立即跳过 00 态, 这要求通过 FF_3 , FF_4 和 G_3 的延时小于 CK_{in} 的半个周期。这通常是很难实现的, 将会使设计变得复杂, 并需要更高的功率消耗。因此, 第一中选择是更为可取的。

参考文献

- [1]. W. L. Rhode, *Digital PLL Frequency synthesizer, Theory and Design*, Englewood Cliffs, NJ: Prentice Hall, 1983.
- [2]. W. F. Egan, *Frequency Synthesis by Phase Lock*, New York: John Wiley, 1981.
- [3]. J. A. Crawford, *Frequency Synthesizer Design Handbook*, Norwood, MA: Artech House, 1994.
- [4]. F. M. Gardner, *Phaselock Techniques*, 2nd ed., New York: John Wiley, 1979.
- [5]. B. Razavi, ed., *Monolithic Phase-Locked Loop and Clock Recovery Circuits*, Piscataway, NJ: IEEE Press, 1996.
- [6]. R. E. Best, *Phase-Locked Loop*, 2nd ed., New York: McGraw-Hill, 1993.
- [7]. C. A. Sharpe, "A 3-State Phase Detector Can Improve Your Next PLL Design," *EDN*, pp. 55-59, September 20, 1976.
- [8]. I. Shahriary et al., "GaAs Monolithic Phase/Frequency Discriminator," *IEEE GaAs IC Symp. Dig. Of Tech. Papers*, pp. 183-186, 1985.
- [9]. I. A. Young, J. K. Greason, and K. L. Wong. "A PLL Clock Generator with 5 to 110MHz of Lock Range for Microprocessors," *IEEE Journal of Solid-State Circuits*, Vol. SC-27, pp. 1599-1607, November 1992.
- [10]. D. K. Jeong et al., "Design of PLL-Based Clock Generation Circuits," *IEEE Journal of Solid-State Circuits*, Vol. SC-22, pp. 255-261, April 1987.
- [11]. F. M. Gardner, "Charge-Pump Phase-Locked Loops," *IEEE Trans. Comm.*, Vol. COM-28, pp. 1849-1858, November 1980.
- [12]. S. Barab and A. L. Bride, "Uniform Sampling Analysis of a Hybrid PLL with a Sample-and Hold Phase Detector," *IEEE Tran. Aerospace and Electronic Systems*, Vol. 11, pp. 210-216, March 1975.
- [13]. J. Crawford, "The Phase/Frequency Detector," *RF Design*, pp. 46-57, February 1985.
- [14]. J. Blake, "Design of Wideband Frequency Synthesizers," *RF Design*, pp. 26-34, May 1988.
- [15]. V. F. Kroupa, "Noise Properties of PLL Systems," *IEEE Trans. Comm.*, Vol. COM-30, pp. 2244-2252, October 1982.
- [16]. B. G. Goldberg, "Reviewing Various Techniques for Synthesizing Signals," *Microwaves & RF*, pp. 181-185, May 1996.
- [17]. J. Gibbs and R. Temple, "Frequency Domain Yield Its Data to Phase-Locked Synthesizer," *Electronics*, pp. 107-113, April 27, 1978.

- [18]. Philips Semiconductor, *Philips RF/Wireless Communications Data Handbook*, 1996.
- [19]. T. A. D. Riley, M. A. Copeland, and T. A. Kwasniewsky, "Sigma-Delta Modulation in Frequency Synthesis," *IEEE J. of Solid-State Circuits*
- [20]., Vol. 28, pp. 553-559, May 1993.
- [21]. B. Miller and R. J. Conley, "A Multiple modulator Fractional-N PLL Synthesizers," *Microwave J.*, pp. 124-134, September 1996
- [22]. J. Tierney, C. M. Rader, and B. Gold, "A Digital Frequency Synthesizer," *IEEE Trans. Audio and Electroacoustics*, Vol. 19, p. 48-57, March 1971.
- [23]. H. T. Nicolas and H. Samuelli, "A 150-MHz Direct Digital Frequency Synthesizer in 1.25-um CMOS with -19dBc Spurious Performance," *IEEE J. of Solid-State Circuits*, Vol. 26, pp. 1959-1969, December 1991.
- [24]. W. Egan, "Modeling Phase Noise in Frequency Dividers," *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 37, pp. 307-315, July 1990.
- [25]. J. Yuan and C. Svensson, "High-Speed CMOS Circuit Technique," *IEEE J. Solid-State Circuits*, Vol. 24, pp. 62-70, February 1989.
- [26]. B. Razavi, K. F. Lee, and R. H. Yan, "A 13.4-GHz CMOS Frequency Divider," *Dig. of International Solid-State Circuits Conference*, pp. 224-225, February 1994.
- [27]. P. Larsson and C. Svensson, "High-Speed Architecture for a Programmable Frequency Divider and a Dual-Modulus Prescaler," *IEEE J. of Solid-State Circuits*, Vol. 31, pp. 744-748, May 1996.
- [28]. R. L. Miller, "Fractional Frequency Generators Utilizing Regenerative Modulation," *Proc. IRE*. Vol. 27, pp. 446-457, July 1939.
- [29]. M. Kuriso *et al.*, "A Si Bipolar 28 GHz Dynamic Frequency Divider," *ISSCC Dig. Tech. Papers*, pp. 92-93, February 1992.

第9章 功率放大器

功率放大器是射频无线电接收发送器中消耗能量最多的电路模块。功率放大器的设计，尤其是用于线性、低电压工作，依然是一个难题，还没有很好的解决办法。实际上，功放的设计过程中需要大量反复试验——这也是为什么人们喜欢离散和混合实现的电路的一个原因。

这一章的目的是阐明功放设计中的问题与挑战，特别是在便携式应用方面。首先我们概述一下 PA 的设计要点和 PA 的工作类型，然后我们研究高效率的电路结构，例如 E 类和 F 类电路。其次，我们描述了大信号阻抗匹配的问题，并且分析了线性化的技术。最后，为了进一步巩固概念，我们列举了几个设计的实例。

9.1 概述

首先，我们考虑一个将 1W 能量传输给一个 $50\ \Omega$ 天线的发射机。天线上电压的峰峰值摆幅 V_{pp} 等于 20V，而通过负载的峰值电流等于 200mA。为了能够直接用共源（或者共射）电路来驱动负载，图 9.1(a)、(b)所示电路需要的电源电压要大于 V_{pp} 。但是，如果用一个大电感代替 9.1(b)中的电流源（称为“射频阻塞 (RFC)”），需要的电源电压可以降低至原来的一半，因为 V_x 可以在从约等于 0 到 $2V_{DD}$ 的范围内变化。实质上，RFC 近似为一个可以承受正电压和负电压的电流源。虽然 RFC 的使用可以降低电源电压，但是并不能缓解有源器件经受的“应力”，即 M_1 承受的最大漏源电压。另外，这个例子中如果 V_{DD} 超过了 10V，在便携式系统中就相当于要有很多的电池。

图 9.1 带有不同类型的负载连接的共源电路级

为了能在更低的电源电压下将 1W 的能量传输给 $50\ \Omega$ 天线，需要在功率放大器和负载之间插入一个匹配网络（如图 9.2(a)所示）。例如，如图 9.2(b)所示，X 节点处的峰峰值摆幅为 $5V_{pp}$ ，通过一个转换比为 1:4 的无损变压器，得到输出电压的峰峰值摆幅为 $20V_{pp}$ 。从另一个角度看来，匹配网络相当于将 R_L 的值减小，这样，功率放大器提供的有限电压摆幅仍然可以传输需要的输出功率。

图 9.2 (a) 作为电压放大器的匹配网络，(b) 使用变压器作为匹配网络

电压的摆幅需要放大，意味着功率放大器提供的电流必须同比例地增大。在图 9.2(b)的例子中，变压器原边的峰值电流超过了 800mA。当 $V_x \approx 0$ 时，由于流过 M_1 漏极的电流为通过 RFC 和变压器的电流之和，则这个电流的最大值为接近 1.6A。实际上，功率放大器的效率可能低于 40%，这就意味着有更大的峰值电流通过输出管。

通过输出器件和匹配网络的大电流是功率放大器的设计尤其是封装中的一个难题。如果通过输出管的峰值电流有几安培的话，那么 900MHz 时的摆率就在 10A/ns 左右。因此，即使是一个 10pH 的寄生电感也会使电压摆幅减少 100mV。另外，寄生电感还会引起电路的振荡和不稳定。类似地，存在于晶体管、RFC 或匹配网络中的几十毫欧的串联电阻也可能引起很大的损耗。由于这些原因，在其它模拟和射频电路中不重要的版图和封装的问题在功率放大器中就变得很重要了。

值得注意的是，与低噪声放大器中的输入级一样，功率放大器的输出级通常只有一个晶体管，这是因为如果在信号通路中有更多的有源器件的话，大的电流会引起更大的损耗。

表 9.1 列出了为便携式系统设计的 PA 的可以达到性能参数的范围。功率放大器的效率有两种定义方法。第一种是漏极（或者收集极）效率 η ，它等于负载上的功率（通常为一次谐波）与电源功率之比。另一种是功率附加效率 PAE，它是输入功率和输出功率之差与电源功率之比。如果功率放大器的功率增益比较大的话，那么 $\eta \approx \text{PAE}$ 。需要对效率和一些系统中线性特性进行补充的是，功率放大器输出中的毛刺和谐波必须满足无线标准和 FCC 限制。

如果收发器使用了在发射和接收的时间槽中无偏移的频分复用，那么功率放大器在接收频带的输出毛刺和噪声也是很关键的（见第 4 章）。这是因为双工机两个频带之间的隔离是有限的，这会在功率放大器的输出到低噪声放大器（LNA）的输入之间产生明显的窜扰[26]。由噪声系数和功率放大器的增益可以得出，为了使低噪声放大器输入端的噪声变得可以忽略，输出的热噪声必须降到-130dBm/Hz 以下。

表 9.1 典型功率放大器的性能指标

输出功率	+20 — 30dBm
效率	30 — 60%
IMD	-30dBc
电源电压	3.8V — 5.8V
增益	20 — 30dB
输出毛刺和谐波分量	-50 — -70dBc
功率控制	开/关 或者 1dB 步长
稳定因子	>1

在数字无线标准中，功率放大器的另一个特点是对于输出功率的控制。比如，在像 IS-54 和 GSM 之类的 TDMA 系统中，功率放大器周期性地打开和关闭以节约能源。而且，像在第 4 章中解释的那样，在 IS-95 中输出能量的变化步长必须要能达到 1dB。

9.1.1 线性和非线性功率放大器

在第 3 章中，我们解释了为什么功率放大器的线性度在某些调制（比如 $\pi/4$ -QPSK 调制）中变得很重要。问题的关键是频谱的再生和相邻频道中的能量。例如，在 IS-54 中，相邻频道中的总的积分能量必须至少比本频道中的能量低 26dB[1]。另外，动态系统中的非线性可能会导致幅度调制到相位调制的转化^[2]，这会破坏了载波的相位特性。

功率放大器的非线性特性通常用双频测试来表征。如图 9.3 所示，电路中所用的是两个幅度相等、频率有一点偏差的正弦信号，在输出端测量它们引起的交调（IM）积。对于相邻频道的干扰，三阶 IM 分量变得很重要；而对于另一种（度量）相邻频道能量，五阶分量是重要的。制造厂商通常会指定当输出的主要成分的功率比满功率低 6dB 时的，最大的 IM 积[3]。例如，一个功率为 1W 的功率放大器的表征是通过调整输入电平以使得输出端的每个信号的功率为+24dBm（在 50Ω 系统中）。

图 9.3 功率放大器的双频测试

但是，在某些应用中，对功率放大器的双频测试并不是一个接近实际的表征，因为如果信号是一个随机地调制的信号（例如 $\pi/4$ QPSK）的话，这种方法不能精确地表示相邻频道的干扰。因为这个原因，功率放大器的线性特性通常必须通过测量对调制波形的响应的频谱

再生来获得。尽管如此，双频测试法在设计开始阶段提供了一种粗略但行之有效的测量方法。

9.2 功率放大器的分类

功率放大器一般被分为以下几类：A类、B类、C类、D类、E类和F类等[4]。传统的功率放大器一般都假设输入和输出波形都为正弦波。而我们在9.3节中会看到，如果不做这种假设的话，功率放大器的性能将会变得更好。

在本节中，我们简要地介绍了A类到C类功率放大器，重点讲述了它们在便携式系统中的优缺点。

9.2.1 A类和B类功率放大器

A类放大器在整个的输入输出范围内都是线性工作的。在三种单管放大器中，共源（共射）结构的效率最高；在共栅（共基）电路中，输出电流等于流经输入网络的电流，这会引来比较大的损耗；而在共漏（共集电极）电路中，除了负载电流之外还要加额外的偏置电流，这样会使效率降低大约一半。

图9.4所示的分别是由双极型晶体管和场效应晶体管组成的A类放大器。我们如何定义它们是线性工作的呢？如果晶体管的工作点没有明显的变化，那么电路就可以认为是线性的。但是如果电路的输出电流达到几个安培的话，为了保证线性，应该加多大的偏置电流呢？这就是A类放大器的定义中不明确的地方。

图9.4 用（a）MOS器件，（b）双极型晶体管的A类电路级

图9.4(b)所示电路可以用电流源驱动，这样由于 β 值对偏置电流的弱依赖性，电路就能达到更好的线性特性^[4]。例如，变压器可以将信号源阻抗的值增大，而使之近似为一个电流源[4]。问题是大的基极-发射极电容会分流一部分输入电流，而由集电极电流变化引起的该电容（表达式为 $C_D = g_m \tau_F$ ）的变化会引起非线性。另一个问题是，当基极由阻抗很高的电流源驱动时，集电极-发射极击穿电压会降低。

为了计算A类放大器最大的漏极（集电极）效率，我们注意到：（1）如果图9.4中的漏极（集电极）电压为峰峰值约为 $2V_{DD}$ 的正弦电压，那么传送到匹配网络的功率就等于 $V_{DD}^2/(2R_{in})$ ；（2）、为了使 V_X 达到 $2V_{DD}$ ，RFC必须提供大小为 V_{DD}/R_{in} 的电流。由于RFC的电流是个相对不变的常量，从而电源提供的功率就等于 V_{DD}^2/R_{in} 。因此最大的效率就等于50%。

以上的计算有一个有趣的地方。在图9.4中，当 V_X 从 $2V_{DD}$ 变化到 V_{DD} 再到0时，晶体管的电流从0变化到 $2V_{DD}/R_{in}$ 。换句话说，只有当晶体管截止的时候 V_X 才能达到 $2V_{DD}$ ；也就是说，只有在非线性可以接受的情况下，最大的效率才能达到50%。这一点同A类放大器是线性电路的观点相违背，从而使得A类放大器的实际效率低于40%。A类放大器的导通角为360度，因为它在整个周期内都是导通的。值得注意的是，当射频输入信号为0的时候，输出级电路消耗的功率最大。

另一个值得关注的地方是小信号时A类放大器的效率。例如，在第4章中讲过的，在IS-95中，为了节约电池能量以及使与其他用户间的串扰降到最低，输出功率是可变的。在9.4(a)所示的电路中，如果晶体管的偏置电流为 V_{DD}/R_{in} ，不管传送到负载上的功率有多大，电源的功耗总是等于 V_{DD}^2/R_{in} 。效率因而减小为 $\eta = V_{out}^2/(2V_{DD}^2)$ 。因此，在小信号的时候为

了保存电池能量，晶体管的偏置电流也必须降低。

B 类放大器的效率比 A 类的高，它包括两个并联的部分，每一部分导通角为 180 度。我们熟悉的一个例子是通常用于低频功率放大器中的推挽电路[5]，如图 9.5 所示。当 V_{in} 越来越为正的时候， Q_2 几乎截止，输出电流由 Q_1 提供。同理，当 V_{in} 减小的时候， Q_2 导通而 Q_1 上的电流很小。我们可以直观地得出，B 类放大器的效率会比较高，因为电源电流总是通过负载，而在图 9.4 所示的 A 类电路中，电源电流在一个周期的部分时间内只通过晶体管。

图 9.5 推挽输出级

由于大部分的双极型晶体管和场效应晶体管工艺中缺乏高速的 p 型器件，因而使得图 9.5 中的推挽结构在射频范围内不能得以广泛应用。此外， Q_1 和 Q_2 的基极-发射极电压限制了输出电压的幅度，降低了可获得功率和效率。这可以用以下方法解决：用两个 n 型器件来产生差分输入，从而驱动一个差分到单端的转换器。图 9.6 所示就是一个例子[4]，其中 M_1 和 M_2 的漏极电流流过变压器 T_1 ，以此来驱动单端负载。

图 9.6 采用变压器的 B 类电路级

为了计算图 9.6 所示的 B 类放大器的最大效率，我们注意到 X 节点和 Y 节点的最大电压为 $2V_{DD}$ ，而每个节点与 V_{DD} 之间的等效电阻为 n^2R_L ，其中 n 为变压器的匝数比。因此， T_1 的输入功率为 $P_{in} = V_{DD}^2 / (2n^2R_L)$ 。而通过 V_{DD} 的平均电流为

$$I_{DD,avg} = \frac{2}{T} \int_0^{T/2} \frac{V_{DD}}{n^2R_L} \sin \omega t dt \quad (9.1)$$

其中，因子 2 代表 M_1 和 M_2 提供的两个半正弦电流。从上式可以得到 $I_{DD,avg} = 2V_{DD} / (\pi n^2R_L)$ ，从而平均功率为 $P_{supp} = 2V_{DD}^2 / (\pi n^2R_L)$ 。因此最大的效率等于 $\eta = P_{in} / P_{supp} = \pi / 4 \approx 79\%$ 。应该注意的是以上的推导假定 T_1 是没有损耗的。

与 A 类放大器类似，B 类放大器为了得到最大的效率，也需要晶体管的偏置电流有大的变化，因此也面临着非线性的严重问题。有意思的是，如图 9.5 所示的 B 类放大器通常用于高性能的音频功率放大器中，因为它们可以嵌入到高增益的反馈回路中使得失真降到最低。然而，在射频范围内很难达到同样的效果，因为很高的环路增益通常需要多级电路，导致严重的稳定性的问题。

图 9.6 所示的 B 类放大器的另一个缺点是它需要低损耗的高频变压器。通过 T_1 原边的大电流会引起显著的电阻损耗，从而降低整体的效率。另一个类似的但不那么严重的问题是在电路的输入端需要一个单端到差分的转化电路。

我们应该指出，B 类放大器也可以是指图 9.6 所示电路的半个电路，例如，一个只导通半个周期（导通角为 180 度）的共射（共源）电路。这样的电路是非线性的，但是，如果我们假定输出电压为正弦波的话，那么它的效率仍等于 $\pi / 4$ 。

9.2.2 C 类功率放大器

在上面研究的 B 类放大器中，我们注意到每个晶体管只在半个载波周期内导通。在 C 类放大器中，为了提高效率，输出管的导通时间小于半个周期。

图 9.4 中的 A 类放大器做一下修改就可以应用于图 9.7 中的 C 类放大器, 该电路加了偏置, 这样当 $V_{in} > |V_b| + V_{TH}$ 的时候, M_1 导通, 其中 V_b 是一个负电压。换句话说, 在每个周期内, 晶体管产生波形窄的电流来驱动输出。匹配电路中通常包括一些滤波器以滤除谐波分量。

图 9.7 C 类电路级

C 类放大器和单管 B 类放大器的区别在于导通角 θ 。当 θ 减小的时候, 晶体管的导通时间缩短, 因此消耗的能量也降低。但是, 同时负载上的功率也会随之降低。如果我们假定晶体管的电流是正弦波的一部分, 而输出电压是峰值为 V_{DD} 的正弦波的话, 那么效率就是 θ 的函数。从[4]中可以得到, 效率为

$$\eta = \frac{1}{4} \frac{\theta - \sin \theta}{\sin(\theta/2) - \theta/2 \cdot \cos(\theta/2)}, \quad (9.2)$$

当 $\theta = 360^\circ$ (A 类) 时, $\eta = 50\%$; 当 $\theta = 180^\circ$ (B 类) 时, $\eta = 79\%$; 当 $\theta = 0$ (C 类) 时, $\eta = 100\%$

C 类放大器的一个显著特征是它的最大效率能达到 100%。但是, 另一个需要考虑的问题是负载上的实际功率。从[4]中可以得到,

$$P_{out} \propto \frac{\theta - \sin \theta}{1 - \cos(\theta/2)}. \quad (9.3)$$

当导通角变为 0 的时候, P_{out} 的值降为 0。换句话说, C 类放大器只有当它输出其峰值功率的一小部分时才能显示出高的效率。因此, 实际的 C 类放大器不适用于便携式收发机, 因为在便携式收发机中全部输出功率时的效率才是最为关心的。

9.3 高效率功率放大器

A 类、B 类和 C 类放大器都假设输出管的电压和电流波形为正弦波 (或者是正弦波的一部分), 这样就限制了 A 类、B 类放大器的效率和 C 类放大器的输出功率。实际上, 存在于这些波形中的高次谐波分量可以用于提高放大器的性能。

A 类放大器 考虑图 9.4 所示的 A 类放大电路。我们在 9.2 节的分析中假定了漏极电流和电压为正弦波, 实际上大信号工作时会产生谐波分量。现在假定匹配网络的阻抗对基次谐波很小, 而对二次谐波很大。从而, 漏极电压的边缘变得比正弦波陡, 这样会提高效率。这是因为过渡时间变短, 在大电压工作时降低了在晶体管上有大电流的时间。

值得注意的是以上的修正并不需要增加在传输给负载的信号中的谐波分量。它只是利用不同谐波的阻抗不同的这一特点使得漏极电压接近于一个方波——虽然实际情况下为了给多于一个或两个的谐波提供指定的阻抗时, 匹配网络会变得很复杂。

举个例子, 考虑图 9.8(a)所示的 A 类电路, 其中 L_1 、 C_1 和 C_2 组成匹配网络, 它将 50Ω 的负载在 $f=850\text{MHz}$ 的时候转化为 $Z_1=9 \Omega + j0$, 而在 $2f=1.7\text{GHz}$ 的时候转化为 $Z_2=330 \Omega + j0$ ^[7]。图 9.8(b)为漏极电压的波形。该电路中负载上的功率为 2.9W, 效率为 73%, 三阶失真为 -25dBc^[7]。对别的谐波终止 (termination) 的考虑见[8]。

图 9.8 (a) 采用高次谐波终止的功放, (b) 漏电压波形

E 类放大器 E 类放大器是非线性放大器，当传输全功率时效率可以达到接近 100%，这是它相对于 C 类放大器的一个显著优点。在详细研究 E 类功率放大器之前，我们先回顾一下图 9.2(a)中的电路，如图 9.9。

图 9.9 开关输出电路级

假定电路中的晶体管工作在理想的开关状态，而不是依赖于电压的电流源。这种电路称为“开关功率放大器”，在以下条件下可以得到很高的效率：（1）当 M_1 上有电流时，它上面的压降很小；（2）当 M_1 上的压降为有限值时，它的电流为小电流；（3）开关状态之间的转换时间最小化[9]。由（1）和（3）我们可以推断出开关的导通电阻必须非常小，而且 M_1 栅极上的电压必须近似为矩形波。然而，即使以上两个条件都满足，如果 M_1 导通时， V_X 很高的话，那么条件（2）依然无法满足。此外，实际情况下，在高频时也很难得到陡峭的输入波形。

重要的是我们要认识到前面几节研究的功率放大器和图 9.9 中的开关电路之间的基本区别：前者的输出匹配网络是在假定了晶体管是作为一个电流源工作的条件下设计的，而对于后者，这种假设不是必须的。如果有源器件依然是作为电流源，为了使晶体管不进入可变电阻区（对于双极型器件是饱和区），必须要能够精确控制漏极（收集极）上的最小电压。这种要求使得效率成为电源电压、电路元件值和匹配网络的 Q 值的敏感函数。而且，即使所有的器件和波形都是理想的，晶体管上必须维持的最小压降仍然会使得效率是一个有限值。与之相反的是，在开关放大器中，最小的漏极（集电极）电压可以接近为 0（虽然这会使双极型器件工作在深饱和状态而产生别的不希望的效应[9]）。

E 类放大器通过设计适当的负载来解决输入输出转换时间有限的问题。如图 9.10 所示，E 类放大器由一个输出管 M_1 、一个接地电容 C_1 和一个串联网路 C_2 、 L_1 组成。RFC 在工作频率下有很高的阻抗， M_1 的漏极电容包括在 C_1 之中[9]。选择 C_1 、 C_2 、 L_1 和 R_L 的值使得 V_X 满足以下三个条件（如图 9.11）：（1）开关关断时， V_X 为低电平的时候足够长，从而使电流能够降为 0；（2）开关导通之前 V_X 的值刚好达到 0；（3）当开关导通时， dV_X/dt 也接近为 0。我们通过观察电路是否符合这些条件来检验电路的性能。

图 9.10 E 类电路级

图 9.11 在 E 类电路级中的电压，电流波形

第一个条件，它为 C_1 所保证，解决了晶体管的下降时间为有限值的问题。如果没有 C_1 ，当 V_{in} 降低的时候， V_X 的值会随之上升，从而引起 M_1 的功率的损失。

第二个条件保证了开关上的压降及通过开关的电流不会在导通点附近交迭，因此即使是在输入输出转化时间有限的情况下，也能使晶体管的功率损耗降到最低。

第三个条件降低了效率对第二个条件不满足的敏感度。也就是说，如果元件或电源电压的改变引起了电压和电流波形的交迭，效率只会稍微降低，因为 $dV_X/dt=0$ 意味着在关断点 V_X 的值不会发生很大的变化。

第二和第三个条件的实现比较不直接了当。在开关关断之后，负载网络的响应是一个带有 C_1 、 C_2 和 L_1 的初始值的阻尼二阶系统（如图 9.12）[9]。其时域响应与电路的 Q 值有关，图 9.12 中给出了欠阻尼、过阻尼和临界阻尼三种情况下的电压波形。在临界阻尼的情况下， V_X 接近为 0 而斜率也为 0。因此，如果开关在这个时候开始导通，那么第二和第三个条件就满足了。

E 类放大器在效率和输出谐波分量之间进行了折衷。为了降低谐波失真，输出网络的 Q 值必须比第二和第三种条件通常所要求的值要高。我们可以在负载电阻前加入滤波器，但是

滤波器会消耗一部分的功率[10]。

E 类放大器的另一个性质是当开关保持在关断状态时，峰值电压很高，接近为 $3.56V_{DD}-2.56V_S$ ，其中 V_S 为晶体管上的最小压降[9]。如果 $V_{DD}=3V$ 而 $V_S=200mV$ ，峰值电压就会超过 10V，这就需提高晶体管的击穿电压。

在设计 E 类放大器中需要的方程，读者可以参考[9]。

F 类放大器 将谐波应用于图 9.8 所示的 A 类放大器中的想法也可以应用到非线性放大器之中。如果在图 9.9 所示的一般的开关电路中，负载网络在二次或者三次谐波时有很高的阻值，那么开关上的电压波形的边缘会比正弦波陡峭，因此就降低了晶体管的功率损耗。这样的电路称为 F 类放大器[4, 11]。

图 9.13 给出了 F 类放大器的一个例子。由 L_1 和 C_1 组成的谐振回路在 $2f_m$ 或者 $3f_m$ 的时候发生谐振，从而在 X 节点提高二次或者三次谐波的幅度。在图 9.13 中我们可以看到，当三次谐波变强的时候，开关上的压降接近为方波。有趣的是，如果 M_1 的漏极电流是半个正弦波（例如经过半波整流过的正弦波）的话，那么三次谐波就不存在了。但是，实际上，因为晶体管的 I/V 特性并不完全满足平方关系，波形中依然存在着三阶失真。

如果晶体管的漏极电流为半波整流过的正弦波，可以证明 F 类放大器的最大效率在三次谐波阻抗最大时等于 88%，而在二次谐波阻抗最大时等于 85%[11]。

9.4 大信号阻抗匹配

功率放大器通常在输出管和负载之间有一个匹配网络。在 A 类放大器中，如果晶体管作为理想电流源，那么匹配网络就只是单纯地将负载电阻的值降低而不产生电抗分量。但是实际上，有源器件的输出阻抗是有限值，既有实部也有虚部，而且会随着输出电压和电流的变化而变化。因此必须将非线性的输出复阻抗与线性负载进行匹配。虽然在这里我们假设负载是电阻而且是个常量，但是实际上，天线的阻抗既可以包含电抗分量，也可以随收发机在外界的位置而变化。

在研究非线性阻抗匹配之前，让我们先考虑一个简单的例子，在这里晶体管的模型为一个理想的电流源和一个线性的电阻性的输出阻抗（如图 9.14(a)）。这里有一个明显的矛盾之处，即最大功率传输定理要求 $R_L=R_O$ ，而我们在 9.1 节中要求 R_L 有一个较小的值，比如 $R_L \ll R_O$ 。哪一个是正确的呢？如果 $R_L=R_O$ ，会有两个问题。首先，传输给 R_L 的功率等于 R_O 上消耗的功率，这会使 A 类放大器的最大效率从 50%降为 25%。换句话说，最大功率传输与最大效率不符合。其次，如果 R_L 变得与 R_O 一样大，那么除非有更高的电源电压，否则输出功率会很小。因此，在一般的功率放大器中， $R_L \ll R_O$ 。

图 9.14 用 (a) 线性电阻，(b) 线性电阻与电抗，(c) (b) 的简单的匹配网络，
的器件输出阻抗模型

下一步，如图 9.14(b)所示，我们假设晶体管输出阻抗包括线性的实部和虚部。我们注意到由于输出管的宽度一般为几毫米，因而 X_O 在高频时也相应的很小。因此匹配网络必须提供电抗部分来抵消 X_O 的影响。图 9.14(c)给出了一个简单的例子，其中当谐振在基波频率时， L_1 抵消 C_O 的作用，而 C_1 和 L_2 对 R_L 起阻抗变换作用。

现在来考虑一般的非线性输出复阻抗的情况。在输出电压范围的中点对阻抗作小信号近似可以得到匹配网络各元件的粗略的值，但是确定它们在大信号效率时的值需要很多尝试，也会有误差，尤其是当要考虑封装时的寄生效应的时候。实际情况下通常是采用一种叫做“拉

负载测量”的更为系统化的方法。

在一个拉负载测试中，对输出功率作为从晶体管看过去的复阻抗的函数测量和画图[12, 13]。由于复阻抗需要两个轴，因此画出来的图实际上是功率在复阻抗平面中的等值线，例如，Smith 圆图。图 9.15(a)建立了拉负载测试的概念。一个可变的精确校准的调谐器作为匹配网络，根据控制输入对晶体管提供了一个可变的复阻抗。借助于一个自动系统， Z_1 的实部和虚部缓慢变化，而功率表的读数始终保持为一个常数。结果是一个对应于该功率的等值线图（如图 9.15(b)）。实际上，由于 Z_1 改变时 Z_{in} 也改变，为了使从信号产生器看进去的阻抗保持恒定（通常等于 50Ω ），就需要在信号产生器和晶体管之间加入另一个调谐器。

图 9.5 (a) 拉负载测试，(b) Smith 圆图上的功率等值图

如果传输到输入电路的功率是个常数，那么当 Z_1 接近它的最佳值 Z_{opt} 时输出功率会增加。当 Z_1 的取值范围逐渐变窄时，等值线的间距逐渐变小；最后当 Z_1 取值为唯一的值 Z_{opt} 时，输出功率达到最大值 P_{max} 。换句话说，这种测试通过缩小 Z_1 的值的范围来得到最大的输出功率和相应的负载阻抗。应该注意的是功率等值线同时也指出了 P_{out} 对于在 Z_1 选择上误差的敏感度。

虽然拉负载这种方法需要利用电脑控制的带有极其精确和稳定的调谐器的测试装置，但是它仍然被广泛地应用于功率放大器的设计之中。但是，这种方法也有三个缺点。首先，不同尺寸器件的测量结果不能通用。其次，不同频率下得到的功率等值线和负载的值各不相同。第三，由于拉负载算法中不一定会提供高次谐波阻抗达到峰值的条件，因而它不能得出在多个谐波终断情况下的效率和输出功率。因此，用这种方法得到的数据来设计功率放大器还需要反复实验。

为了在仿真时避免进行复杂的测试，拉负载方法也可以通过电路仿真来实现[13]。与上述过程类似，负载阻抗作微小的变化，从而建立功率等值线，最后得到 Z_{opt} 的值。在这里，如果负载的变化不是自动控制的，那么这个过程也是冗长而且麻烦的。另外，这样的仿真需要晶体管的精确模型[14]，尤其是输出阻抗部分更要精确，而 SPICE 模型并不能给出在高频时和电压电流有大的变化时输出阻抗的精确模型。

9.5 线性化技术

在很多射频电路中需要使用线性功率放大器。在第 3 章说过，像 QPSK 和 $\pi/4$ QPSK 这类的频带宽度利用率高的调制技术需要使用线性功率放大器来使频谱再生达到最小。此外，同时处理多个频道的放大器需要是线性的以避免交叉调制。这种情况会发生在多载波系统中，比如基站发射机、有线电视发射机和“正交频分复用 (OFDM)”电路之中[15]。

目前，为便携式装置设计的大多数线性功率放大器中都使用了效率在 30%到 40%左右的 A 类输出电路。为了得到更高的效率，可以先使用非线性功率放大器再对电路进行线性化。理想情况下，这种方法可以降低整体的失真到可以接受的程度，而不会很明显的降低效率。

本节描述的线性化技术在某些场合已被用于复杂、昂贵的射频和微波系统中，但是还没有应用于低成本的便携式终端系统中。这是因为这种方法一般会使设计变得复杂，还需要进行多种调整，而且当器件特性随温度和输出功率变化时，这种方法会变得不那么有效。尽管如此，对这种技术及其局限性有基本的了解在功率放大器的设计中还是有用的，尤其是当集成电路工艺水平提高至能够提供比离散实现更好的可重复的，高复杂程度的应用。

很多线性化方法的一个重要的缺点是它们需要功率放大器的核心部分具有一定的线性

度，亦即，如果输出管工作为一个理想开关，这些方法就没有效果了。这个问题会在下面阐明。另外，每种技术只适用于某种类型的放大器。

9.5.1 前馈

非线性功率放大器产生的输出电压波形可以看成是一个输入信号线性复制和一个误差信号的和。前馈网络用于计算这个误差，对其进行适当的放大或缩小，再从输出波形中扣除掉[16]。图 9.16(a)给出了一个简单的例子，其中主功率放大器的输出为 V_M ，而 V_N 为 V_M 的 $1/A_V$ 。 V_N 减去输入信号，所得结果再放大 A_V 倍，再被 V_M 减去。如果 $V_M=A_V V_{in}+V_D$ ，其中 V_D 代表失真量，那么 $V_N=V_{in}+V_D/A_V$ ，由此得到 $V_P=V_D/A_V$ ， $V_Q=V_D$ ，因此 $V_{out}=A_V V_{in}$ 。实际上，电路中的两个放大器在高频时有很大的相位漂移，需要使用延时线，如图 9.16(b)，其中 Δ_1 用于补偿主功率放大器的相位漂移，而 Δ_2 用于补偿误差放大器的相位漂移。有时候称从 V_{in} 到第一个减法器的两个通路为“信号抵消回路”，而称从 M 和 P 节点到第二个减法器的两个通路为“误差抵消回路”。

图 9.16 (a) 简单的前馈拓扑，(b) 加入延迟元件

前馈结构与反馈结构相比其优点是：即使带宽是有限的，而且在每个电路模块都有大的相位漂移，它仍然是稳定的。这一点在射频和微波电路中显得尤为重要，因为在关心的频带附近极点和振荡是不可避免的，因而很难得到稳定的反馈。

但是前馈线性化也有一些缺点。首先，模拟延时器件需要使用无源器件（例如微带线），并且对 Δ_2 的功率损耗有很高的要求。其次，输出部分的减法器必须利用低损耗元件实现，例如高频变压器[17]。第三，线性化的量取决于每个减法器的输入信号的功率和相位匹配。可以知道[17, 18]，如果从 V_{in} 到第一个减法器的两个通路的相位失配为 $\Delta\phi$ ，相对增益失配为 $\Delta A/A$ ，那么 V_{out} 的 IM 积的幅度抑制量为

$$E = \sqrt{1 - 2\left(1 + \frac{\Delta A}{A}\right) \cos \Delta\phi + \left(1 + \frac{\Delta A}{A}\right)^2} \quad (9.4)$$

举个例子，如果 $\Delta A/A=5\%$ ， $\Delta\phi=5^\circ$ ，那么 $E=0.102$ ，也就是说前馈使 IM 积降低了大约 20dB。在误差修正回路中的相位和增益失配进一步降低了电路的性能。

前馈的概念可以应用于“嵌套环路”中^[19]，这种电路利用图 9.16(b)的拓扑作为起主要的功率放大器，并将其应用于另一个前馈系统中。当然，这样的电路要复杂得多。

9.5.2 反馈

虽然局部反馈与全局反馈技术广泛地应用于高速电路中，但并不能简单地用于射频功率放大器。这是因为，对于极为非线性的功率放大器，必须要有很高的环路增益。如果考虑了由于电路的极点和谐振引起的不稳定，这一点在高频时是很难做到的。封装上的寄生效应、大的瞬时电流以及由电容和磁耦合引起的不可避免的反馈使得功率放大器在其它频率时即使没有外加的反馈也很容易发生振荡。

如果大部分的环路增益是在低频下获得，那么与反馈有关的两个问题，即增益不足与过多的相位偏移，就可以得到缓解。在发射机中，这一点是可以做到的，因为经过功率放大器处理的波形实际上是通过基带信号或者 IF 信号进行上变频得到的。因此，如果对功率放大器的输出进行下变频，得到的信号就可以与负反馈回路中的原低频信号相比较。这个概念

在图 9.17(a)中说明,那里误差放大器 A_1 与功率放大器之间要求的频率转换可由通过两个混频器完成。换句话说,这个环路试图使 V_{PA} 成为 V_{in} 的复制,只不过是在不同的频率下。由于通过混频器和功率放大器在高频时的总的相位漂移通常都超过 180° ,因此为了确保稳定性,额外相位 θ 被叠加到其中一个 LO 信号中。

图 9.17 (a) 通过频率转换的反馈, (b) 笛卡尔反馈

在对 I 信号和 Q 信号都进行上变频的发射机中,功率放大器的输出信号在与输入信号相比较之前,必须将其分解为相位正交的分量(如图 9.17(b))。这种形式称为“笛卡尔反馈”^[20]。

笛卡尔反馈并不是便携式系统中的通用方法。由反馈解调器和误差放大器引入的额外的复杂度增加了功率损耗,和用离散元件实现的成本(但是在集成电路设计中这些复杂程度不造成太多的麻烦)。而且,在图 9.17 中必需的 θ 值随着温度、工艺参数和输出功率而变化,使得稳定性很难得到保证。

9.5.3 包络消除及恢复

正像在第 3 章提到过的,任意一个带通信号可以表示为 $v(t) = a(t) \cos[\omega_c t + \phi(t)]$,即可以通过包络 $a(t)$ 和相位 $\phi(t)$ 来表示。由此产生了以下方法:先将 $v(t)$ 分解为包络信号和调相信号,再分别进行放大,最后再进行混合。

图 9.18(a)所示电路称为“包络的去除和恢复(ERR)”^[21]。输入信号通过包络检测器(最简单的情况下为二极管检测器)和限幅电路,产生包络 $a(t)$ 和调相分量 $b(t) = b_0 \cos[\omega_c t + \phi(t)]$ 。如图 9.18(b)所示,这两个信号在功率放大器中进行放大和混合。如果晶体管 M_1 工作在开关状态,通过 RFC 的电流就是 $ka(t)$ 的函数,从而对 M_1 产生的信号的幅度进行调制。应该注意的是,如果 M_1 作为一个电流源,就不能进行包络的恢复,因为 $ka(t)$ 和 $I_{D1}(t)$ 的线性时不变组合不能进行调制。

图 9.18 (a) 包络消除与恢复, (b) 输出级的实现

ERR 与前馈和反馈相比,其主要优点是它不需要功率放大器核心是线性的,因而可以设计出效率最大的输出电路。但是,它也有其它缺点。

首先,图 9.18(a)中的两个通路的总相位偏移与增益间的失配必须降低到可以接受的水平,而这是很难的,因为两个通路的电路形式不同,工作的频率也相去甚远。其次,限幅电路中的有源部分(如差分对)在高频时会产生显著的调幅—调相转换(见第 7 章),破坏了图 9.18(a)中 $b(t)$ 的相位。第三,在输出电路中用 $ka(t)$ 作为电源电压降低了效率。比如说,考虑一下图 9.19(a)所示的电路,那里误差放大器 A_1 迫使 V_X 跟踪 $ka(t)$ 。由于功率放大器输出电路的总电流通过晶体管 M_2 , M_2 消耗了大量的功率。因此,一般都采用图 9.19(b)所示的开关调节器。其中,振荡器使 M_2 在开关状态间进行转换,而误差放大器决定了振荡器的占空系数,这样就确保了 V_X 的平均值等于 $ka(t)$ 。不过, M_2 的沟道宽度必须很大,这样 M_2 在导通状态消耗的功率就可忽略了。(9.1 节中曾指出几十毫欧的电阻也可能使效率降低很多。)同时, C_1 必须足够大,从而能在 M_2 关断时提供电源电流。

图 9.19 用 (a) 一个低频反馈放大器, (b) 脉冲宽度调制来对 PA 的输出进行调制

图 9.18(b)所示电路的另一个缺点是 M_1 的漏极电容会随 $ka(t)$ 的变化而变化。因而 M_1 产生的信号的相位会被包络分量所破坏。

9.5.4 采用非线性元件的线性放大 (LINC)

功率放大器系统中为了避免幅度发生变化的一个方法是“用非线性元件进行线性放大 (LINC, linear amplification with nonlinear components)” [22, 23], 如图 9.20 所示, 想法是带通信号 $v_{in}(t) = a(t)\cos[\omega_c t + \phi(t)]$ 可以表示成以下两个常幅度调相信号之和, 这两个信号分别为 $v_1(t) = 0.5V_0 \sin[\omega_c t + \phi(t) + \theta(t)]$ 和 $v_2(t) = -0.5V_0 \sin[\omega_c t + \phi(t) - \theta(t)]$, 其中 $\theta(t) = \sin^{-1}[a(t)/V_0]$ 。因此, 如果 $v_1(t)$ 和 $v_2(t)$ 由 $v_{in}(t)$ 产生, 通过非线性电路放大, 然后进行叠加, 那么输出信号就有与 $v_{in}(t)$ 相同的包络和相位。

图 9.20 采用非线性电路级的线性放大

由 $v_{in}(t)$ 产生 $v_1(t)$ 和 $v_2(t)$ 需要很复杂的电路, 主要是因为它们的相位必须通过对 $\theta(t)$ 调制得到, 而 $\theta(t)$ 本身是 $a(t)$ 的非线性函数。曾经提出过用非线性频率转换的反馈回路来实现 [22, 24], 但是类似于图 9.17 所示的环路的稳定性问题限制了它的应用。另一个方法 [25] 将 $v_1(t)$ 和 $v_2(t)$ 表示为

$$v_1(t) = v_I(t)\cos(\omega_c t + \phi) + v_Q(t)\sin(\omega_c t + \phi) \quad (9.5)$$

$$v_2(t) = -v_I(t)\cos(\omega_c t + \phi) + v_Q(t)\sin(\omega_c t + \phi) \quad (9.6)$$

其中 $v_I(t) = a(t)/2$ 而 $v_Q(t) = \sqrt{V_0^2 - a^2(t)}/2$ 。由于产生 $v_Q(t)$ 所需要的非线性运算可以在低频下进行 (可以用模拟方法, 或者是用供查表的 ROM), 因而这种方法可以简单地采用正交上变频来产生 $v_1(t)$ 和 $v_2(t)$ 。

除了复杂性以外, LINC 还必须面对另外两个问题。首先, 图 9.20 中的两个通路间的增益和相位失配产生了残留失真 [25]。其次, 输出加法器会引起很大的损耗, 因为它必须在两个功率放大器之间产生很好的隔离。虽然看起来加法器可以使用图 9.21 中的简单的形式, 但是我们注意到如果 M_1 和 M_2 作为开关工作, 那么两个调制信号会破坏彼此的相位。

图 9.21 两个 PA 输出的相加

9.6 设计实例

大多数的功率放大器都采用两级结构, 在输入、两级电路之间和输出都有匹配网络 (如图 9.22)。由于输出级的增益一般都低于 10dB, 因此在电路中加入了一个高增益的驱动器, 以此来降低最低输入电压。例如, 如果要求功率放大器必须提供 +30dBm (1W) 的功率, 图 9.22 中的电路可以提供 25 到 30dB 的增益, 因此允许输入从 0 到 +5dBm。最小输入的选择取决于前级电路—调制器或上变频器—的驱动能力, 但是由于功率放大器和调制器经常是独立设计的 (而且可能是不同厂商制造的), 因此一般选择接口电路的功率为 0 到 +5dBm。

图 9.22 典型的 PA 系统

图 9.22 中输入和输出匹配网络有几个用途： N_1 提供 $50\ \Omega$ 的输入阻抗，而 N_3 对输出级的电压幅度进行放大，以便将要求的功率传给 R_L 。如果将功率放大器设计成孤立电路，而通过外部电路实现与前级电路的连接，那么 $50\ \Omega$ 的输入阻抗就是必要的。而且，如果在调制器与功率放大器之间加入了一个无源的外部滤波器，以此来抑制频带外部的谐波和干扰，那么从滤波器看进去的阻抗必须等于标准值 $50\ \Omega$ ，以避免通频带的抖动和不稳定。

为了实用，图 9.22 中驱动电路与输出级间加入了匹配网络。由于设计是以对输出管进行拉负载测试为开始的，因而当输出级设计完成之后，为了得到最大效率而要求的电源阻抗也就决定了。因此，驱动器必须提供这样的一个输出阻抗，通常都需要有匹配网络。 N_2 的使用允许进行模块式的设计：首先设计输出级，其次是驱动器，最后是级间匹配，最终结果需要几次重复。如果没有 N_2 ，驱动器和输出级就必须作为独立的电路来处理，这会使设计过程变得复杂得多。

功率放大器依然主要靠 III-V 族工艺，尤其是 GaAs 结型场效应晶体管和近来的 GaAs 和 InP 异质结双极型晶体管（HBT）。由于 III-V 族器件有较高的截止频率和击穿电压乘积，很高的迁移率，因而大的单位宽度的电流驱动能力，并且在半绝缘衬底上能够制造出高质量的电感和电容，这些使得 III-V 族工艺在高效率功率的放大器的设计中成为首选。现在还不清楚主流的亚微米 CMOS 工艺能否有这样的取得相仿的性能，但是已经在研究它们在小范围内（如局域网）的应用了。

图 9.23 给出了用混合工艺研制出的一个 900MHz 的 MESFET 功率放大器[27]。电源电压为 5.6V，电路工作在 AB 类工作状态，输出功率大约为 +36dBm。设计的过程从输出级开始。通过拉负载测量，得到在 M_2 给定的尺寸下，最佳负载电阻等于 $4\ \Omega$ 。两单元输出匹配网络将 $50\ \Omega$ 的负载转换为这个数值；单元 A 产生 $R_{in1}=10\ \Omega$ 而单元 B 产生 $R_{in2}=4\ \Omega$ 。

图 9.23 在[27]中报告的功率放大器

输出级的输入电阻 R_{in3} 为 $2.7\ \Omega$ ，因此需要输入功率 +23dBm，为此 M_2 进入了深度三极管（即线性）工作区。级间匹配网络将 R_{in3} 转化为 $R_{in4}=53\ \Omega$ ，在电源电压为 5.6V，功率为 +23dBm 时，这个值为驱动电路提供了最佳的负载电阻[27]。在对 M_1 的输入阻抗进行大信号仿真的同时设计了输入匹配网络，得到了 $R_{in5}=50\ \Omega$ 。

虽然该设计是针对 5.6V 的电源电压，但是图 9.23 中的功率放大器在 3.3V 电源电压下的性能也很好：输出功率为 +31.5dBm，效率大约为 50%[27]。

图 9.24 给出了 900MHz 功率放大器的另一个例子[10]。这个放大器是用 $0.8\ \mu\text{m}$ GaAs 工艺做成的完整的单片集成电路，包括一个 F 类驱动电路和一个 E 类输出级。驱动电路由两个谐振回路串联而成，其中一个调到一次谐波，另一个在三次谐波谐振，从而在 X 节点产生近似的方波，因而缩短了 M_2 的开关转换时间。

图 9.24 文献[10]中报导的功率放大器

与图 9.10 中的 E 类电路不同的，图 9.24 的电路采用了一个 M_2 的在片负载电感，因而不需要外部电路元件，但是设计要复杂一些[28]。输出网络由基本的 E 类电路和一个并联一串联的匹配电路组成，为 M_2 提供了最佳的负载电阻。实际上，为了缓解在损耗与转换函数的 Q 值之间的折衷选择，这两部分电路被设计成一个单个滤波器。

电源电压为 2.5V 时, 功率放大器的输出功率为 250mW, PAE 大约为 50%。有趣的是, 输出匹配网络的功率损耗大约是输出管的功率损耗的 1.5 倍[10]。

图 9.25 给出了一个设计在 900MHz, AMPS/CDMA 双模系统下工作的 PA [29]。在 CDMA 模式下, 每级电路中的并联电阻性负反馈增加了稳定性, 并且使功率放大器线性化。在 AMPS 模式下, 输出级电路工作在静电流最小的情况下, 因而在输出功率为+31.5dBm 下, 效率达到了 55%。在 CDMA 模式下, 输出功率和效率分别降为+28dBm 和 35%。

图 9.23 文献[29]中报告的功率放大器

参考文献

- [1] J.S.Kenney and A.Leke, "Power Amplifier Spectral Regrowth for Digital Cellular and PCS Applications," *Microwave J.*, pp.74-92, October 1995.
- [2] S.A.Maas, *Nonlinear Microwave Circuits*, Norwood, MA :Artech House, 1988.
- [3] W. E.Sabin and E.Q.Schoenike, eds., *Single-Sideband Systems and Circuits*, 2nd ed., New York: McGraw-Hill, 1995.
- [4] H.L.Kraus, C.W.Bostian, and F.H.Raab, *Solid State Radio Engineering*, New York: John Wiley, 1980.
- [5] P.R.Gray and R.G.Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed., New York: John Wiley, 1993.
- [6] D.M.Snider, "A Theoretical Analysis and Experimental Confirmation of the Optimally Loaded and Overdriven RF Power Amplifier," *IEEE Trans. Electron Devices*, Vol.14, pp.851-857, December 1967.
- [7] S.C.Cripps, "High-Efficiency Power Amplifier Design," lecture given in Short Course: RF ICs for Wireless Communication, Portland, Oregon, June 1996.
- [8] J.Staudinger, "Multiharmonic Load Termination Effects on GaAs MESFET Power Amplifiers," *Microwave J.*, pp.60-77, April 1996.
- [9] N.O.Sokal and A.D.Sokal, "Class E-A New Class of High-Efficiency Tuned Single-Ended Switching Power Amplifiers," *IEEE Journal of Solid-State Circuits*, Vol.10, pp.168-176, June 1975.
- [10] T.Sowlati et al., "Low Voltage High Efficiency Class E GaAs Power Amplifiers for Wireless Communications," *IEEE Journal of Solid-State Circuits*, Vol.30, pp.1074-1080, October 1995.
- [11] F.H.Raab, "An Introduction to Class F Power Amplifiers," *RF Design*, pp.79-84, May 1996.

- [12] S.C.Cripps, "A Theory of the Prediction of GaAs FET Load-Pull Power Contours," *IEEE MTT-S Dig.*, pp.221-223, 1983.
- [13] L.A.Geis and L.P.Dunleavy, "Power Contour Plots Using Linear Simulators," *Microwave J.*, pp.60-70, June 1996.
- [14] J.Staudinger et al., "Considerations for Improving the Accuracy of Large Signal GaAs MESFET Models to Predict Power Amplifier Circuit Performance," *IEEE J. Solid-State Circuits*, Vol.29, pp.366-373, March 1994.
- [15] A.Lymer, "Digital-Modulation Scheme Processes RF Broadcast Signals," *Microwaves & RF*, pp.149-159, April 1994.
- [16] H. Seidel, "A Microwave Feedforward Experiment," *Bell System Technical J.*, Vol.50, pp.2879-2916, November 1971.
- [17] R.G.Meyer, R.Eschenbach, and W.M.Edgerley, "A Wideband Feedforward Amplifier," *IEEE Journal of Solid-State Circuits*, Vol.9, p.422-488, December 1974.
- [18] E.E.Eid, F.M.Ghannouchi, and F.Beauregard, "Optimal Feedforward Linearization System Design," *Microwave J.*, pp.78-86, November 1995.
- [19] D.P.Myer, "A Multicarrier Feedforward Amplifier Design," *Microwave J.*, pp.78-88, October 1994.
- [20] M.Johansson and T.Mattsson, "Transmitter Linearization Using Cartesian Feedback for Linear TDMA Modulation," *Proc. IEEE Veh. Tech. Conf.*, pp.439-444, May 1991.
- [21] L.R.Kahn, "Single-Sideband Transmission by Envelope Elimination and Restoration," *Proc. IRE*, Vol.40, pp.803-806, July 1952.
- [22] D.C.Cox, "Linear Amplification with Nonlinear Components," *IEEE Trans. Comm.*, Vol.22, pp.1942-1945, December 1974.
- [23] F.H.Raab, "Efficiency of Outphasing RF PA Systems," *IEEE Trans. Communication*, Vol.33, pp.1094-1099, October 1985.
- [24] D.C.Cox and R.P.Leek, "Component Signal Separation and Recombination for Linear Amplification with Nonlinear Components," *IEEE Trans. Comm.*, Vol.23, pp.1281-1287, November 1975.
- [25] F.J.Casadevall, "The LINC Transmitter," *RF Design*, pp.41-48, February 1990.

- [26] RF Micro Devices, "A High Efficiency HBT Analog Cellular Power Amplifier," *Microwave J.*, pp.168-172, January 1996.
- [27] E.Camargo and R.M.Steinberg, "A Compact High Power Amplifier for Handy Phones," *IEEE MTT-S Dig.*, pp.565-568, May 1994.
- [28] R.E.Zulinski and J.W.Steadman, "Class E Power Amplifiers and Frequency Multipliers with Finite DC-Feed Inductance," *IEEE Trans. Circuits and Systems*, Vol.34, pp.1074-1080, September 1987.
- [29] M.Cardullo et al., "Transmitter Chips for Use in a Dual-Mode AMPS/CDMA Chip Set," *Microwave J.*, pp.60-72, March 1996.