

国外电子与通信教材系列

西蒙·赫金

通信系统 (第四版)

Communication Systems
Fourth Edition

[加] Simon Haykin 著
宋铁成 徐平平 徐智勇 等译
沈连丰 审校



电子工业出版社
Publishing House of Electronics Industry
<http://www.phei.com.cn>

著作
策划
出版
发行

通信系统(第四版)

Communication Systems, Fourth Edition

本书能够帮助学生夯实理论基础，畅游学科前沿！

本书已经过全面修订和内容更新，是现代通信系统理论和原理领域最完善的教材。

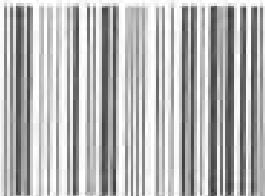
本书新特点：

- 演示通信理论主要内容的 MATLAB 实验
- 新兴数字技术的延伸，如数字用户线（DSL）、无载波幅度/相位调制（CAP）以及离散多音（DMT）等
- 数十个将理论与实际通信系统相结合的实例
- 通过精心组织，本书系统地指导学生学习从脉冲调制到宽带数据传输、从随机过程到差错控制编码等知识。全书以深入浅出的语言向学生传授了深奥的概念。

作者简介

Simon Haykin：IEEE 会士，毕业于英国伯明翰大学电子工程系，加拿大 McMaster 大学电子与计算机工程系教授、通信研究实验室主任。研究方向包括非线性动力学、神经网络和自适应滤波器及其应用。已出版了多本专著，其中包括《神经网络（第二版）》、《智能信号处理》、《自适应滤波器原理（第四版）》、《信号与系统（第二版）》等。

ISBN 7-5053-8254-3



9 787505 382541 >



责任编辑：马 岚

刘 静

封面设计：毛惠庚

本书贴有激光防伪标志。凡没有防伪标志者，属盗版图书

ISBN 7-5053-8254-3/TN · 1720 定价：69.00 元

国外电子与通信教材系列

通信系统

(第四版)

Communication Systems
Fourth Edition

[加] Simon Haykin 著

宋铁成 徐平平 徐智勇 等译

沈连丰 审校

电子工业出版社
Publishing House of Electronics Industry
北京 · BEIJING

内 容 简 介

本书对通信系统的基础理论和关键环节进行了深入分析，力图让学生在讨论中领会通信的精髓。全书首先给出通信系统的梗概及需要研究的关键技术，接着分章详细讨论了随机过程、连续波调制、脉冲调制、基带脉冲传输、信号空间分析、通带数据传输、扩频调制、多用户无线通信、信息论基础以及差错控制编码等。各章都附有大量习题，便于学生实践掌握。书中还给出了很有价值的附录，包括概率论、信号与系统简述、贝塞尔函数、汇合型超几何函数、密码学方面的介绍以及一些有用的表格等。全书强调通信理论的统计基础，并给出了用 MATLAB 模拟的 8 个计算机实验，这些实验几乎覆盖了各章节的主要内容，形成了独特的通信理论“软件实验室”。

本书可以作为高等院校通信类、信息类、电子类、计算机类等专业的研究生或高年级本科生的教材，也可供有关的科研和管理人员参考。

Simon Haykin: **Communication Systems, Fourth Edition.**

ISBN 0-471-17869-1

Copyright © 2001, John Wiley & Sons, Inc. All Rights Reserved.

AUTHORIZED TRANSLATION OF THE EDITION PUBLISHED BY JOHN WILEY & SONS NEW YORK, CHICHESTER, BRISBANE, SINGAPORE AND TORONTO. No part of this book may be reproduced in any form without the written permission of John Wiley & Sons, Inc.

Simplified Chinese translation edition Copyright © 2003 by John Wiley & Sons, Inc. and Publishing House of Electronics Industry.

本书中文简体字翻译版由电子工业出版社和 John Wiley & Sons 合作出版。未经出版者预先书面许可，不得以任何方式复制或抄袭本书的任何部分。

版权贸易合同登记号：图字：01-2002-5887

图书在版编目 (CIP) 数据

通信系统 (第四版) / (加) 赫金 (Haykin S.) 著；宋铁成等译. - 北京：电子工业出版社，2003.10
(国外电子与通信教材系列)

书名原文：Communication Systems, Fourth Edition

ISBN 7-5053-8254-3

I. 通... II. ①赫... ②宋... III. 通信系统 - 教材 IV. TN914

中国版本图书馆 CIP 数据核字 (2003) 第 086140 号

责任编辑：马 岚 刘 静

印 刷 者：北京兴华印刷厂

出版发行：电子工业出版社 <http://www.phei.com.cn>

北京市海淀区万寿路 173 信箱 邮编：100036

经 销：各地新华书店

开 本：787 × 1092 1/16 印张：46 字数：1178 千字

版 次：2003 年 10 月第 1 版 2003 年 10 月第 1 次印刷

定 价：69.00 元

凡购买电子工业出版社的图书，如有缺损问题，请向购买书店调换；若书店售缺，请与本社发行部联系。

联系电话：(010) 68279077。质量投诉请发邮件至 zts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

序

2001年7月间，电子工业出版社的领导同志邀请各高校十几位通信领域方面的老师，商量引进国外教材问题。与会同志对出版社提出的计划十分赞同，大家认为，这对我国通信事业、特别是对高等院校通信学科的教学工作会很有好处。

教材建设是高校教学建设的主要内容之一。编写、出版一本好的教材，意味着开设了一门好的课程，甚至可能预示着一个崭新学科的诞生。20世纪40年代MIT林肯实验室出版的一套28本雷达丛书，对近代电子学科、特别是对雷达技术的推动作用，就是一个很好的例子。

我国领导部门对教材建设一直非常重视。20世纪80年代，在原教委教材编审委员会的领导下，汇集了高等院校几百位富有教学经验的专家，编写、出版了一大批教材；很多院校还根据学校的特点和需要，陆续编写了大量的讲义和参考书。这些教材对高校的教学工作发挥了极好的作用。近年来，随着教学改革不断深入和科学技术的飞速进步，有的教材内容已比较陈旧、落后，难以适应教学的要求，特别是在电子学和通信技术发展神速、可以讲是日新月异的今天，如何适应这种情况，更是一个必须认真考虑的问题。解决这个问题，除了依靠高校的老师和专家撰写新的符合要求的教科书外，引进和出版一些国外优秀电子与通信教材，尤其是有选择地引进一批英文原版教材，是会有好处的。

一年多来，电子工业出版社为此做了很多工作。他们成立了一个“国外电子与通信教材系列”项目组，选派了富有经验的业务骨干负责有关工作，收集了230余种通信教材和参考书的详细资料，调来了100余种原版教材样书，依靠由20余位专家组成的出版委员会，从中精选了40多种，内容丰富，覆盖了电路理论与应用、信号与系统、数字信号处理、微电子、通信系统、电磁场与微波等方面，既可作为通信专业本科生和研究生的教学用书，也可作为有关专业人员的参考材料。此外，这批教材，有的翻译为中文，还有部分教材直接影印出版，以供教师用英语直接授课。希望这些教材的引进和出版对高校通信教学和教材改革能起一定作用。

在这里，我还要感谢参加工作的各位教授、专家、老师与参加翻译、编辑和出版的同志们。各位专家认真负责、严谨细致、不辞辛劳、不怕琐碎和精益求精的态度，充分体现了中国教育工作者和出版工作者的良好美德。

随着我国经济建设的发展和科学技术的不断进步，对高校教学工作会不断提出新的要求和希望。我想，无论如何，要做好引进国外教材的工作，一定要联系我国的实际。教材和学术专著不同，既要注意科学性、学术性，也要重视可读性，要深入浅出，便于读者自学；引进的教材要适应高校教学改革的需要，针对目前一些教材内容较为陈旧的问题，有目的地引进一些先进的和正在发展的交叉学科的参考书；要与国内出版的教材相配套，安排好出版英文原版教材和翻译教材的比例。我们努力使这套教材能尽量满足上述要求，希望它们能放在学生们的课桌上，发挥一定的作用。

最后，预祝“国外电子与通信教材系列”项目取得成功，为我国电子与通信教学和通信产业的发展培土施肥。也恳切希望读者能对这些书籍的不足之处、特别是翻译中存在的问题，提出意见和建议，以便再版时更正。

吴佑寿

中国工程院院士、清华大学教授
“国外电子与通信教材系列”出版委员会主任

出版说明

进入21世纪以来，我国信息产业在生产和科研方面都大大加快了发展速度，并已成为国民经济发展的支柱产业之一。但是，与世界上其他信息产业发达的国家相比，我国在技术开发、教育培训等方面都还存在着较大的差距。特别是在加入WTO后的今天，我国信息产业面临着国外竞争对手的严峻挑战。

作为我国信息产业的专业科技出版社，我们始终关注着全球电子信息技术的发展方向，始终把引进国外优秀电子与通信信息技术教材和专业书籍放在我们工作的重要位置上。在2000年至2001年间，我社先后从世界著名出版公司引进出版了40余种教材，形成了一套“国外计算机科学教材系列”，在全国高校以及科研部门中受到了欢迎和好评，得到了计算机领域的广大教师与科研工作者的充分肯定。

引进和出版一些国外优秀电子与通信教材，尤其是有选择地引进一批英文原版教材，将有助于我国信息产业培养具有国际竞争能力的技术人才，也将有助于我国国内在电子与通信教学工作中掌握和跟踪国际发展水平。根据国内信息产业的现状、教育部《关于“十五”期间普通高等教育教材建设与改革的意见》的指示精神以及高等院校老师们反映的各种意见，我们决定引进“国外电子与通信教材系列”，并随后开展了大量准备工作。此次引进的国外电子与通信教材均来自国际著名出版商，其中影印教材约占一半。教材内容涉及的学科方向包括电路理论与应用、信号与系统、数字信号处理、微电子、通信系统、电磁场与微波等，其中既有本科专业课程教材，也有研究生课程教材，以适应不同院系、不同专业、不同层次的师生对教材的需求，广大师生可自由选择和自由组合使用。我们还将与国外出版商一起，陆续推出一些教材的教学支持资料，为授课教师提供帮助。

此外，“国外电子与通信教材系列”的引进和出版工作得到了教育部高等教育司的大力支持和帮助，其中的部分引进教材已通过“教育部高等学校电子信息科学与工程类专业教学指导委员会”的审核，并得到教育部高等教育司的批准，纳入了“教育部高等教育司推荐——国外优秀信息科学与技术系列教学用书”。

为做好该系列教材的翻译工作，我们聘请了清华大学、北京大学、北京邮电大学、东南大学、西安交通大学、天津大学、西安电子科技大学、电子科技大学等著名高校的教授和骨干教师参与教材的翻译和审校工作。许多教授在国内电子与通信专业领域享有较高的声望，具有丰富的教学经验，他们的渊博学识从根本上保证了教材的翻译质量和专业学术方面的严格与准确。我们在此对他们的辛勤工作与贡献表示衷心的感谢。此外，对于编辑的选择，我们达到了专业对口；对于从英文原书中发现的错误，我们通过与作者联络、从网上下载勘误表等方式，逐一进行了修订；同时，我们对审校、排版、印制质量进行了严格把关。

今后，我们将进一步加强同各高校教师的密切关系，努力引进更多的国外优秀教材和教学参考书，为我国电子与通信教材达到世界先进水平而努力。由于我们对国内外电子与通信教育的发展仍存在一些认识上的不足，在选题、翻译、出版等方面的工作中还有许多需要改进的地方，恳请广大师生和读者提出批评及建议。

电子工业出版社

教材出版委员会

主任	吴佑寿	中国工程院院士、清华大学教授
副主任	林金桐	北京邮电大学校长、教授、博士生导师
	杨千里	总参通信部副部长、中国电子学会会士、副理事长 中国通信学会常务理事
委员	林孝康	清华大学教授、博士生导师、电子工程系副主任、通信与微波研究所所长 教育部电子信息科学与工程类专业教学指导委员会委员
	徐安士	北京大学教授、博士生导师、电子学系副主任 教育部电子信息与电气学科教学指导委员会委员
	樊昌信	西安电子科技大学教授、博士生导师 中国通信学会理事、IEEE会士
	程时昕	东南大学教授、博士生导师 移动通信国家重点实验室主任
	郁道银	天津大学副校长、教授、博士生导师 教育部电子信息科学与工程类专业教学指导委员会委员
	阮秋琦	北方交通大学教授、博士生导师 计算机与信息技术学院院长、信息科学研究所所长
	张晓林	北京航空航天大学教授、博士生导师、电子工程系主任 教育部电子信息科学与电气信息类基础课程教学指导委员会委员
	郑宝玉	南京邮电学院副院长、教授、博士生导师 教育部电子信息与电气学科教学指导委员会委员
	朱世华	西安交通大学教授、博士生导师、电子与信息工程学院院长 教育部电子信息科学与工程类专业教学指导委员会委员
	彭启琮	电子科技大学教授、博士生导师、通信与信息工程学院院长 教育部电子信息科学与电气信息类基础课程教学指导委员会委员
	徐重阳	华中科技大学教授、博士生导师、电子科学与技术系主任 教育部电子信息科学与工程类专业教学指导委员会委员
	毛军发	上海交通大学教授、博士生导师、电子信息学院副院长 教育部电子信息与电气学科教学指导委员会委员
	赵尔沅	北京邮电大学教授、教材建设委员会主任
	钟允若	原邮电科学研究院副院长、总工程师
	刘彩	中国通信学会副理事长、秘书长
	杜振民	电子工业出版社副社长

译 者 序

通信在国民经济中的地位是不言而喻的。近年来,我国高校中设立信息工程、通信工程、电子工程、电子与信息工程等专业的越来越多,而“通信原理”、“数字通信”、“通信系统”、“通信系统工程”等课程则成为这些专业的学位课或学科主干课。围绕这些课程,国内已经出版了不少教材和专著,加上近年引进的一些国外新作,无疑给开设这些课程的教师和学生以更多的选择。受电子工业出版社的委托,我们翻译了《通信系统(第四版)》。该书对通信系统的基础理论和关键环节进行了深入分析,着重强调了通信理论应用的统计基础,给出了用 MATLAB 模拟的 8 个计算机实验,形成了独特的通信系统“软件实验室”,力图让学生在讨论中领会通信理论的精髓。本书既是一本很有特色的教材,也是一本通信技术的专著。

全书以“背景与预览”开篇,首先给出了通信系统的梗概及需要研究的关键技术,接着详细介绍了随机过程,并着重强调其二阶统计特性,随后综合分析了模拟通信的连续波调制。第 3 章和第 4 章讨论了脉冲调制和基带脉冲传输,同时论述了基带形式的脉幅调制信号的传输、信道噪声的影响和数字通信系统中带限信道的性能等。第 5 章至第 7 章则分析了加性高斯白噪声信道下的信号空间,为学生建立起有限能量信号的几何描述的基础,研究了通带数据传输和扩频调制技术。第 8 章专门讨论多用户无线通信的一些关键技术,分析了多径信道和衰落信道的特性以及 TDMA 和 CDMA 等多址技术的性能,介绍了无线通信中的语音编码、自适应天线阵列等概念。第 9 章给出了信息论中的基本限制,它们被广泛应用于数据压缩和数据传输中。最后一章讨论差错控制编码,包括分组码和卷积码,还专门讨论了网格编码调制和 Turbo 码。本书的附录部分也很有价值,包括概率论、信号与系统描述、贝叶斯函数、汇合型超几何函数和密码学等方面的介绍,最后还给出了一些有用的图表。

本书的主要目的是为大学高年级学生或低年级研究生提供关于通信系统的现代分析方法,课程的内容可以根据学生的层次和教师的兴趣酌情增减。关于本书中所有习题的详细解答,以及其他教学用资料,都有相应的网络资源(详情见书后附带的“教学支持说明”),这也为授课老师提供了有力的帮助。全书内容深入浅出,概念清晰,理论分析严谨,逻辑性强,对 IT 领域的科研技术人员以及相关专业的教师和学生来说,也是一本很好的参考书。

本书的翻译,第 2 章至第 4 章由徐平平负责,第 5 章至第 8 章由宋铁成负责,其余由徐智勇负责。沈斐协助翻译了“背景与预览”以及第 1 章,于森、戚继群、董丽君、瞿慧、吴舟等分别协助翻译了第 2 章至第 11 章。全书先由瞿慧、董丽君负责初校,然后由徐智勇、宋铁成再校对,最后由沈连丰负责审校和统稿。本室博士后刘云以及博士生李文峰、傅永生、邱晓华、顾燕、谷金山、钱小勇、许国军、黄忠虎、胡静、鲍淑娣等做了许多具体工作。因此,本书的翻译出版是集体智慧的结晶,渗透着众人的汗水。在此,对所有为这本书的出版提供了帮助的人们表示诚挚的感谢!

译稿对原书中的错误和遗漏之处进行了更正和注释。由于译审者水平有限,加之时间仓促,译文中难免有不妥之处,敬请读者不吝指正。

译 者

2003 年 8 月

于东南大学移动通信国家重点实验室

前　　言

电子工程教育在过去的 20 年里发生了一些根本性的变化,预计这种情况仍将继续下去。现在的电子工程专业的本科教学计划包括下面两门基础课程:

- **信号与系统。**这门课讲述信号和系统的连续时间和离散时间形式,对傅里叶变换的不同形式、拉普拉斯变换和 z 变换进行了详细讲述。通常,这门课还包括对通信系统的初步论述。
- **概率论和随机过程。**这门课先使学生对离散和连续的随机变量有一个直观的理解,然后介绍了随机过程的概念及性质。

通常,在这两门预备课程之后会引入关于通信系统的高级课程。

本书的第四版就是在这种背景下并以此为主要目的写成的。简而言之,本书所提供的对通信系统的论述内容,适用于一到两个学期的本科高年级课程。本书的重点在于通信理论及其应用的统计学基础。

本书注重逻辑性并配有大量例题加以说明,其目的就是通过讨论帮助学生直观地掌握理论知识。除了“背景与预览”这一章,其他每章的结尾都附有大量的习题,这些习题不仅能够帮助学生测试他们对本章所涵盖内容的理解程度,同时也能促使他们拓展对这些内容的思考。每章均有注释与参考,为进一步阅读提供了建议。不影响连贯性而可以绕开的章节或者子章节由脚注加以标识。

本书的一个特色就是包含 8 个基于 MATLAB 的计算机实验。这组实验构成了“软件实验室”的基础,每一个实验均设计了用来扩充相关章节知识的内容。更为重要的是,这些实验开发了 MATLAB 独特的教育功能。

“背景与预览”这一章提供了预备内容和激发学习兴趣的材料,为在接下来的 10 章中全面论述通信系统做了铺垫。这些章节的内容是按以下顺序组织的:

第 1 章详细讲述了随机过程,重点是随机过程的部分性质(如二阶统计)。实际上,讨论仅局限于广义平稳过程。本章对随机过程的相关性和功率谱进行了详细论述。高斯过程和窄带噪声在通信系统的研究中起着重要作用,因此本章的后半部分对其进行了论述,论述将涉及到通信环境中常出现的瑞利分布和莱斯分布。

第 2 章总体讲述了连续波调制(如模拟通信)及其不同形式,简述如下:

(i) 幅度调制,可以是以下几种形式(根据已调载波频谱特性不同来区分)中的一种:

- 标准幅度调制
- 抑制载波的双边带调制
- 正交幅度调制
- 单边带调制
- 残留边带调制

(ii) 角度调制,可以假定是以下两种相关联形式中的一种:

- 相位调制

● 频率调制

本章还对这些波形的时域和频谱特性、调制和解调方法,以及信道噪声对性能的影响进行了讨论。

第3章讲述脉冲调制,同时也讨论了抽样、量化和编码过程,这是模拟信号数字化传输的基础。本章也可以看成是模拟通信向数字通信的过渡。特别对以下几种脉冲调制方式进行了讨论:

(i) 模拟脉冲调制(这里只将时间用离散形式表示)包括以下几种特殊形式:

- 脉冲幅度调制
- 脉冲宽度(持续时间)调制
- 脉冲位置调制

因为脉冲幅度调制是其他模拟或者数字调制方式的基础,所以对它的性质进行了详细讨论。

(ii) 数字脉冲调制(同时将时间和信号幅度用离散形式表示)包括以下几种特殊形式:

- 脉冲编码调制
- 增量调制
- 差分脉冲编码调制

在增量调制中,为简化系统的实现,其抽样率要远高于脉冲编码调制。相反地,在差分脉冲编码调制中,则通过采用一个预编码器来减小抽样率。这个预编码器利用了信息承载信号的相关特性。

(iii) MPEG1音频编码标准,包括一个作为编码器设计关键的心理声学模型。

第4章涵盖了基带脉冲传输。基带脉冲传输就是用基带形式的脉冲幅度调制信号来进行传输。主要讨论了以下两个要点:信道噪声和有限的信道带宽对数字通信系统性能的影响。假定信道噪声是加性白噪声,则可采用匹配滤波器使得噪声的影响最小。匹配滤波器是设计通信接收机的基础。由于是有限信道带宽,因此会出现码间干扰现象。为了克服这种信号相关干扰的不良影响,我们通常采用脉冲成形滤波器或相关编码器/译码器,本章对这两种方法都进行了讨论。本章还对用于用户和因特网服务提供商之间直接通信的数字用户线路进行了讨论。接着对为克服信道噪声和码间干扰组合影响而采用的最佳线性接收机进行了推导,进而介绍了自适应均衡。

第5章讨论加性高斯白噪声信道的信号空间分析。尤其是建立了能量有限信号的几何表示基础,进而衍生出相干接收机,并证明了它和匹配滤波器是等效的,本章以差错概率及其近似计算的讨论作为结束。

第6章讨论通带数据传输,这里用正弦信号来简化数字调制信号在带通信道中的传输。本章建立在第5章提到的信号的几何表示基础之上。特别地,对于以下几种调制技术,估算了信道噪声对数字通信产生的影响:

1. 相移键控,这是和相位调制相对应的数字调制方式,即用载波的相位来表示一组给定的离散量。

2. 混合幅度/相位调制,包括正交幅度调制和无载波幅度/相位调制。
3. 频移键控,这是和频率调制相对应的数字调制方式,即用载波的频率来表示一组给定的离散量。
4. 多信道调制类,紧随离散多音之后,已经成为非对称数字用户线路的标准。

在数字通信系统中,定时非常重要,这意味着接收机必须与发射机同步,本书提及的接收机包括相干的和非相干的。相干接收机需要恢复载波相位和符号定时,而非相干接收机只需要恢复符号定时而不必恢复载波相位,采用这种策略是基于对随机载波相位恢复成本的考虑。本章的后半部分讨论了同步技术,尤其偏重于离散时间信号的同步。

第7章介绍扩频调制,这种调制方式与前面几章讨论的传统调制方式不同,它通过牺牲信道带宽来获得好的安全性并保护信号免受干扰。这里讨论了直接序列扩频和跳频两种扩频方式。

第8章论述多用户无线通信,多用户无线通信就是多个用户能接入一个普通的无线信道,这种通信信道常用于卫星和无线通信系统中,这里对这两种系统都进行了讨论。本章还介绍了链路预算分析,主要偏重于介绍相关的天线和传播的概念,以及噪声的计算。

第9章讲述信息论的一些基本原理,其中包括用于无失真数据压缩、数据压缩和数据传输的香农理论。这些理论为信源和通信信道界定了性能上限。形成这些理论的两个基本概念是:(1)信源的熵(类似于热力学中熵的定义);(2)信道容量。

第10章论述差错控制编码,围绕着为了在噪声信道中进行稳定传输而对数字数据进行编码的各种技术进行展开。讨论了以下四种差错控制编码:

1. 线性分组码,是一组线性无关的码字,每组码字均由信息比特和校验比特组成。校验比特用于差错控制。
 2. 循环码,是线性分组码的一种。
 3. 卷积码,对连续消息序列用串行方式进行卷积操作。
 4. Turbo码,用物理上可实现的方式,建立了一种接近香农信道容量的新的编码构成方法。
- 对上述码的编码和译码方法均进行了讨论。

本书还包括如下6个附录作为补充内容:

- 附录1是对概率论的复习。
- 附录2是对信号与系统的简述,回顾了傅里叶变换及其性质、各种带宽的定义、希尔伯特变换和窄带信号与系统的低通等价模型。
- 附录3介绍贝塞尔函数及其修正形式,贝塞尔函数用于研究频率调制、噪声环境下信号的非相干检测和符号定时同步。
- 附录4介绍汇合型超几何函数,用于噪声环境下幅度调制信号的包络检波。
- 附录5介绍密码学,这是通信安全的基础。
- 附录6包括12个有用的表格。

如前所述,本书的主要目的就是为高年级本科生提供适合一到两个学期教学的现代通信系统教材。课程内容的组织自然也由采用此书的学生背景和教师兴趣决定。本书的内容既有深度又有广度,能满足各种背景和兴趣的读者,这为本书内容的取舍提供了相当的灵活性。为了给授课教师提供一定的帮助,关于书中所有问题的详细解答,以及其他教学用资料,都有相

应的网络资源(详情见书后附带的“教学支持说明”)。

致谢

我要对以下同仁表示深切的谢意:Gregory J. Pottie 博士(加州大学),Santosh Venkatesh 博士(宾夕法尼亚大学),Stephen G. Wilson 博士(弗吉尼亚大学),Gordon Stüber 博士(佐治亚理工学院),Venugopal Veeraralli 博士(康奈尔大学),Granville E. Ott 博士(德州大学)。他们审阅了先前版本的草稿,对成书提供了大量的建议。第 2 章中讨论的噪声对包络检波的影响是基于 Santosh Venkatesh 博士为我们提供的课程笔记,对此我要表示谢意。我还要感谢 Gordon Stüber 博士同意我引用图 6.32。

我要感谢 Michael Moher 博士(渥太华通信研究中心),他审阅了先前版本草稿中的 5 章,并对 Turbo 码提出了很多建设性意见。同样还要感谢 Brendan Frey 博士(滑铁卢大学)在提炼 Turbo 码内容方面提供的极为宝贵的帮助,他还为讲述低密度奇偶校验码提供了绘制图 9.18 的软件,并允许我引用图 10.27 和图 10.33。我要感谢 David Conn 博士(McMaster 大学)审阅了“背景与预览”一章,并为如何改进此章内容的表述提供了建议。

我还要对下述人员致谢:Jean-Jacque Werner 博士(朗讯科技),James Mazo(朗讯科技),Andrew Viterbi 博士(高通),Radford Neal 博士(多伦多大学),Yitzhak (Irwin) Kalet 博士(Technion),Walter Chen 博士(摩托罗拉),John Cioffi 博士(斯坦福大学),Jon Mark 博士(滑铁卢大学)和 Robert Dony 博士(Guelph 大学)。他们为本书所选内容提出了很多宝贵的建议,谨此表示深切的谢意。同时也对 Donald Wunsch II 博士(密苏里大学)的校阅和建议表示感谢。

感谢我的研究生 Mathini Sellathurai(McMaster 大学),他完成了本书的计算机实验部分;Hugh Pasika(McMaster 大学),他对“背景与预览”一章提出了很多有用的建议,并对本书的一些插图进行了计算。也非常感谢 Mathini Sellathurai 和 Alpesh Patel 校对了全书。

我尤其要感谢 Wiley 公司的编辑 Bill Zobrist 在本书整个写作过程中的大力支持和帮助;感谢 Wiley 的高级编辑 Monique Calello 在写作的各个阶段对本书质量不知疲倦的监督;感谢 Katherine Hepburn 为本书所做的广告宣传和市场推广;感谢 Karen Tongish 仔细编辑抄写了本书草稿;感谢 Katrina Avery 认真校对了每一页;感谢 Kristen Maus 组织了本书的索引。

同样,我还要向我的技术协调人 Lola Brooks 致谢,感谢她为输入本书手稿所付出的不知疲倦的努力。我也很高兴在这里感谢 McMaster 大学图书馆的助理馆员 Brigitte Maier 和资料管理员 Regina Bendig 多次为我查找参考书目中的文献资料提供帮助。

Simon Haykin

目 录

背景与预览	1
第 1 章 随机过程	25
1.1 简介	25
1.2 随机过程的数学定义	25
1.3 平稳过程	27
1.4 均值、相关函数和协方差函数	28
1.5 遍历过程	34
1.6 随机过程通过一个线性时不变滤波器	35
1.7 功率谱密度	36
1.8 高斯过程	45
1.9 噪声	48
1.10 窄带噪声	53
1.11 基于同相和正交分量的窄带噪声表示法	54
1.12 基于包络和相位分量的窄带噪声表示法	56
1.13 正弦信号加窄带噪声	58
1.14 计算机实验:平衰落信道	60
1.15 总结与讨论	63
注释与参考	65
习题	66
第 2 章 连续波调制	75
2.1 引言	75
2.2 幅度调制	76
2.3 线性调制方案	79
2.4 频率搬移	87
2.5 频分复用	89
2.6 角度调制	91
2.7 频率调制	92
2.8 FM 系统中的非线性影响	107
2.9 超外差接收机	108
2.10 连续波调制系统中的噪声	110
2.11 相干检测线性接收机中的噪声	112
2.12 包络检波 AM 接收机的噪声	114
2.13 FM 接收机中的噪声	121
2.14 计算机实验:锁相环	133
2.15 总结与讨论	138

注释与参考	140
习题	141
第3章 脉冲调制	157
3.1 引言	157
3.2 抽样过程	157
3.3 脉冲幅度调制	161
3.4 其他脉冲调制形式	164
3.5 带宽-噪声权衡	165
3.6 量化过程	166
3.7 脉冲编码调制	172
3.8 PCM 系统中的噪声	179
3.9 时分复用	181
3.10 数字复接器	183
3.11 PCM 的优点、局限性和改进	186
3.12 增量调制	187
3.13 线性预测	191
3.14 差分脉冲编码调制	194
3.15 自适应差分脉冲编码调制	197
3.16 计算机实验:自适应增量调制	199
3.17 MPEC 音频编码标准	200
3.18 总结与讨论	203
注释与参考	204
习题	205
第4章 基带脉冲传输	213
4.1 引言	213
4.2 匹配滤波器	213
4.3 噪声引起的差错率	218
4.4 码间干扰	223
4.5 无失真基带二进制传输的奈奎斯特准则	225
4.6 相关电平编码	231
4.7 基带 M 进制 PAM 传输	238
4.8 数字用户线	239
4.9 最佳线性接收机	244
4.10 自适应均衡	248
4.11 计算机实验:眼图	253
4.12 总结与讨论	256
注释与参考	257
习题	260

第 5 章 信号空间分析	268
5.1 引言	268
5.2 信号的几何表示	269
5.3 连续 AWGN 信道到矢量信道的变换	276
5.4 似然函数	279
5.5 噪声中信号的相干检测:最大似然译码	280
5.6 相干接收机	283
5.7 差错概率	285
5.8 总结与讨论	292
注释与参考	292
习题	293
第 6 章 通带数据传输	299
6.1 引言	299
6.2 通带传输模型	302
6.3 相干 PSK	303
6.4 幅度/相位混合调制方案	321
6.5 相干 FSK	331
6.6 未知相位信号的检测	351
6.7 非相干正交调制	355
6.8 非相干二进制 FSK	360
6.9 差分相移键控	361
6.10 单载波数字调制方式的比较	364
6.11 语音频带调制解调器	367
6.12 多信道调制	376
6.13 离散多音	384
6.14 同步	391
6.15 计算机实验:载波恢复与符号定时	400
6.16 总结与讨论	406
注释与参考	406
习题	409
第 7 章 扩频调制	420
7.1 引言	420
7.2 伪噪声序列	421
7.3 扩频的概念	427
7.4 相干二进制 PSK 的直接序列扩频	429
7.5 信号空间的维度和处理增益	432
7.6 差错概率	435
7.7 跳频扩频	437

7.8 计算机实验:最大长度码和 Gold 码	442
7.9 总结与讨论	445
注释与参考	445
习题	446
第 8 章 多用户无线通信	449
8.1 引言	449
8.2 多址技术	450
8.3 卫星通信	451
8.4 无线链路分析	453
8.5 无线通信	464
8.6 多径信道的统计特性	469
8.7 瑞利衰落信道上的二进制信号	475
8.8 TDMA 和 CDMA 无线通信系统	479
8.9 无线通信中的语音信源编码	482
8.10 无线通信中的自适应天线阵列	484
8.11 总结与讨论	489
注释与参考	490
习题	491
第 9 章 信息论基础	497
9.1 引言	497
9.2 不确定性、信息和熵	497
9.3 信源编码定理	502
9.4 无失真数据压缩	503
9.5 离散无记忆信道	509
9.6 互信息	511
9.7 信道容量	514
9.8 信道编码定理	516
9.9 连续信号的相对熵和互信息	520
9.10 信息容量定理	522
9.11 信息容量定理的含义	526
9.12 有色噪声信道的信息容量	532
9.13 率失真定理	536
9.14 数据压缩	538
9.15 总结与讨论	539
注释与参考	540
习题	542
第 10 章 差错控制编码	549
10.1 引言	549

10.2 离散无记忆信道	551
10.3 线性分组码	553
10.4 循环码	562
10.5 卷积码	574
10.6 卷积码的最大似然译码	579
10.7 网格编码调制	586
10.8 Turbo 码	591
10.9 计算机实验:Turbo 码的译码	598
10.10 低密度奇偶校验码	599
10.11 不规则码	606
10.12 总结与讨论	607
注释与参考	608
习题	610
附录 1 概率论	617
附录 2 信号与系统简述	627
附录 3 贝塞尔函数	645
附录 4 汇合型超几何函数	650
附录 5 密码学	652
附录 6 表格	669
术语表	676
参考文献	682
索引	697

背景与预览

这里列出的背景与预览材料将为以后各章的通信系统学习做铺垫。我们将着重介绍以下几方面：

- 通信过程
- 基本通信资源，即发射功率和信道带宽
- 信源
- 两种主要交换方式：电路交换和分组交换
- 用于传送发射机与接收机间信息承载信号的通信信道
- 通信系统的基础——调制过程
- 模拟与数字通信系统
- 香农信息容量定理
- 一个数字通信的问题

本章以通信发展史上的重要事件作为结束，以激励读者。

通信过程^[1]

如今，通信已经渗透到我们日常生活的各个角落，这可以很容易从许多方面发现。如我们手边的电话，起居室中的电视和收音机，办公室和家中可以接入因特网的计算机终端，以及报纸等，都可以迅速地把世界各地的消息传送给人们。通信还使我们得以获知远海上航行的船只、空中翱翔的飞机，甚至外太空火箭和卫星的状态。坐在车中可以通过无线电话与远方的办公室或家中通话。通信使得由传感器测得的天气情况可以随时送达给预报员。事实上，诸如上述有关通信应用的例子不胜枚举。

最基本的通信意味着通过一个系列过程从一点向另一点进行信息的传送。其描述如下：

1. 消息信号的产生，如语音、音乐、图像或计算机数据。
2. 通过一系列诸如电子的、可听到的或可看到的符号对消息信号进行描述，并精确到某一精度。
3. 用一种适于在物理媒介中传输的方式对这些符号进行编码。
4. 将编码后的符号发送到指定的目的地。
5. 对原始符号进行译码和重现。
6. 对原始消息信号的重新生成。系统的非理想性会造成信号质量的下降。

当然，还有许多人类思维不能实时介入的通信方式。例如，在两台或多台计算机之间进行的计算机通信中，人能做的只是启动计算机程序或指令，或者监控这些程序或指令的结果。

无论考察哪种通信过程，每一个通信系统都由三个基本部分组成，即发射机、信道和接收机，如图 1 所示。发射机位于空间中的某一点，接收机位于远离发射机的另一点，而信道是连

接二者的物理媒介。发射机的作用是将由信源产生的消息信号转变成适于信道传送的形式。但由于信道特性是非理想的,发射信号在信道中传播时会发生失真,且噪声和干扰信号(由其他信源产生)也会被加入到信道输出中,使得接收信号成为被破坏的发射信号。接收机的任务是对接收信号进行处理,将其再现为可被用户识别的原始消息信号的形式。

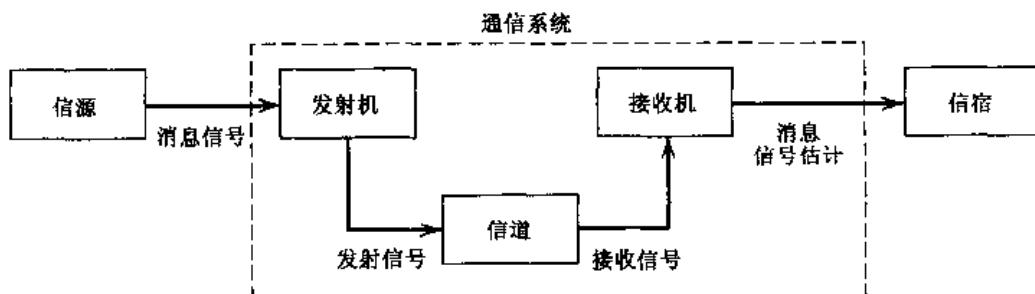


图1 通信系统的组成部分

以下是两种基本通信方式:

1. 广播方式:包括一个大功率的发射机和多个造价相对便宜的接收机。这里承载信息的信号流是单向的。
2. 点对点方式:这种通信方式发生在一个发射机和一个接收机之间的链路中。在这种情况下,承载信息的信号通常是双向传送的,这就需要链路的每一端都有一个发射机和一个接收机。

收音机和电视机是广播通信方式的典型例子。随处可见的电话机则是点对点通信的一种形式,点对点通信的另外一个例子是地球站与在相距很远的行星表面行走的机器人之间的链路。

以上这些不同的通信系统和在此未提及的一些通信系统都有一个共同特点:每个基本通信过程在本质上都具有统计特性。正是由于这个重要的原因,本书致力于通信系统统计基础的论述。同时我们阐述了不同通信方法研究的共同基本问题,从而得出对它们的比较性评价。

主要通信资源

在通信系统中,使用的两个主要资源是:发射功率和信道带宽。发射功率是指发射信号的平均功率。信道带宽是为传输信号分配的一个频带。一般系统的设计原则都是尽可能高效地使用这两个资源。在大部分通信信道中,一个资源可能比另一个更重要,因此我们可以将通信信道分为功率受限信道和带宽受限信道。电话线路是一个典型的带宽受限信道的例子,而空间通信链路和卫星信道则是典型的功率受限信道。

当消息信号的频谱向下延伸至零或很低的频率时,我们把信号的带宽定义为其上限频率,而高于上限频率的信号频谱可忽略,因此不用于传输信息。例如,一般语音的频谱远低于10 kHz,尽管大多数语音的平均功率集中在100~600 Hz的范围内,但从300~3100 Hz的频谱可以获得较好的清晰度。因此,我们发现对后一个频率范围有较好响应的电话电路能够提供相当令人满意的商用电话服务。

另外很重要的一点是,我们需要牢记在通信系统中会不可避免地出现噪声。噪声指的是会

干扰通信系统传输和处理消息信号的多余波形。噪声源可能在系统内部也可能在系统外部。

测量噪声影响的一个定量方法是引入信噪比(SNR)作为系统的参数。例如,我们可以定义接收机的输入信噪比为信号平均功率和噪声平均功率的比值,二者需要在同一点测量。通常用分贝(dB)来表述信噪比,dB 定义为功率比对数(以 10 为底)的 10 倍。例如,信噪比 10, 100 和 1000 分别对应于 10 dB, 20 dB 和 30 dB。

信源

电信环境主要由语音、音乐、图像和计算机数据这四个重要信源组成。信源可以由承载信息的信号来表征。信号定义为时间的单值函数,时间是其惟一的变量;在每个瞬时,函数有惟一的值。信号可以是一维的,如语音、音乐或计算机数据;可以是二维的,如图像;可以是三维的,如视频数据;也可以是四维的,如随时间变化的物体运动数据。下面我们详细阐述不同的信源。

1. 第一种信源是语音。语言是人类交流的基本工具。特别地,语言交流过程包括将信息从说者传送到听者的三个连续步骤:

- 产生:说话者脑中想好的消息是用语音信号表达的。语音信号由声道产生的声音(即压力波)组成,并根据语言规则组织。
- 传播:声波在空气中以 300 m/s 的传播速度到达听者耳中。
- 理解:传来的声音被收听者翻译为接收到的消息,从而完成了说者与听者之间传递信息的一系列过程。

语音的产生可以看成是一种滤波,由声源激励声道滤波器产生。声道由一条不均匀的代表性区域组成,它始于喉门(即声道的开口),中止于嘴唇。当声音沿着声道传播时,由于声道的频率选择性而形成声音频谱(即频率)。其作用与管风琴的风琴管里观察到的共振现象有些类似。这里必须强调的一点是,语音的功率谱(即长时间平均功率按频率的分布)在频率为零时接近于零,并在几百赫时达到最大值。但为了更合理地分析这个问题,我们必须注意到听觉机制对频率的敏感性。此外,频带是我们所考察的通信系统的重要方面,而通信过程的“本质”就是频带。例如先前提到的信道带宽为 300 ~ 3100 Hz 的适合于商业电话的通信系统。

2. 第二种信源是音乐,由诸如钢琴、小提琴和长笛等乐器产生。当我们按住钢琴键时,乐器演奏出的音符可以持续一小段时间;而当长笛演奏者拖长声音时,音符则可以维持很长一段时间。典型的音乐有两种形式:由一连串声音按时间顺序组成的旋律和由一系列同时发出的声音构成的和声。类似于语音信号,音乐信号也是双极性的。但音乐信号的频谱带宽要宽于语音信号频谱,可以延伸至大约 15 kHz。因此,传输音乐信号需要比语音信号更宽的信道带宽。

3. 第三种信源是图像,它依赖于人类视觉系统对它的感知。图像可以是动态的,如电视;也可以是静态的,如传真。首先讨论电视,运动的图像被转换为电子信号,以便于把它们从发射机传送到接收机。要完成这一转变,需要顺序扫描每张完整的图像,扫描过程在电视摄像机中完成。在黑白电视中,摄像机包含有能在光电阴极聚焦图像的光学

器件,光电阴极由大量感光元件组成。产生于感光表面的充电图案被电子束扫描,从而产生一个瞬时变化的输出电流,这个输出电流和原始图片上的点与点之间的空间亮度变化方式相同。此输出电流称为视频信号。电视机中所用的扫描方式是一种称为光栅扫描的空间抽样方式,它将二维空间的图像亮度转变为一维空间的波形。这和我们阅读一张打印稿时的方法相似,从左到右一行一行地扫描。在北美的模拟电视中,一张图像被划分为 525 行以构成一帧。每帧被分解为两个交织场,每个场由 262.5 行组成。为了方便表达,我们称这两个场为 I 和 II。扫描过程如图 2 所示。场 I 用实线绘制,场 II 用虚线绘制。每个场的起始和结束位置也包括在图中。先扫描场 I,电视摄像机的扫描点以恒定的速率从左到右经过场中的每一行,并且测量扫描点中心的图像亮度,扫描点本身是形成图像本地空间平均的部分原因。当扫描到特定行结尾时,扫描点会快速地回扫(以水平方向)到场中下一行的开头。这种回扫称为水平回扫。这里描述的扫描过程一直持续到扫描完整个场。这时,扫描点快速地从场 I 的结束位置(垂直方向)移动到场 II 的起始位置。这种回扫称为垂直回扫。场 II 的扫描过程与场 I 的模式相同。扫描完每个场需要 $1/60$ s。相应地,扫描完一帧或整张图片需要 $1/30$ s。一帧 525 行,行扫描频率为 15.75 kHz。电视接收机的显像管每秒闪过 30 张静止的图片,因此人类的眼睛接收到的图像看上去是运动的。这一效果是由视觉暂留现象引起的。在水平和垂直回扫的间隔内,发射机产生的消隐脉冲使显像管不工作。此外,用消隐周期中发射的特殊脉冲来实现各种发射机和接收机上不同扫描操作之间的同步。所以,同步脉冲并不显示在重显图片上。电视图像的重显质量受以下两个基本要素限制:

- 光栅扫描可利用的行数,限定垂直方向上的图像分辨率。
- 发射视频信号所能利用的信道带宽,限定水平方向上的图像分辨率。

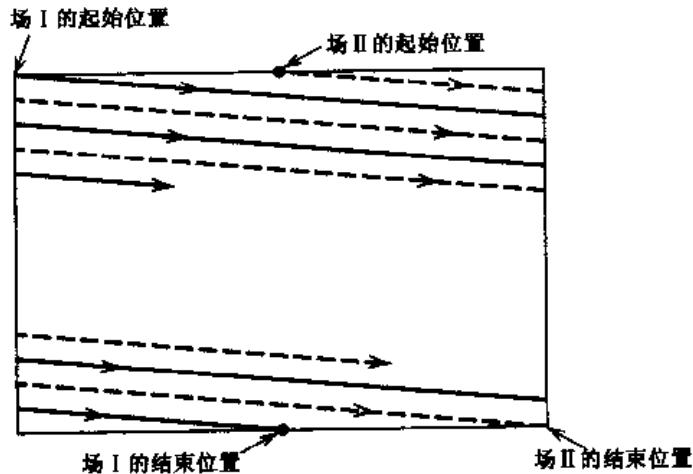


图 2 交织光栅扫描

在每个方向上,分辨率用沿相应方向上电视图像可被观察者分辨的黑白交替的最大线条数来表示。在 NTSC(国家电视系统委员会)系统即北美标准中,所用的参数为视频带宽4.2 MHz,频率可向下延伸至零。这一带宽比语音信号的带宽大几个数量级。然而

需要注意的是语音信号是双极性的,而视频(电视)信号只有正极(即是单极的)。

在彩色电视中,颜色的感知是依靠人眼中的三种色彩感应器(圆锥细胞)来实现的:红色、绿色和蓝色,其波长分别是 570 nm, 535 nm 和 445 nm。这三种颜色称为三原色,因为自然界中的任何其他颜色都可以近似地用三原色叠加混合而成。这一物理事实实际上就是商业电视广播发射彩色信号的基础。三原色分别用视频信号 $m_R(t)$ 、 $m_G(t)$ 和 $m_B(t)$ 表示。观测表明,为了节省带宽并产生能够在传统黑白(单色)电视接收机上观看的图片,三原色的传输可通过用三种信号对其进行惟一表示来实现,这三种信号是 $m_R(t)$ 、 $m_G(t)$ 和 $m_B(t)$ 信号的独立线性组合,如下所示:

- 亮度信号 $m_L(t)$, 当它被传统的单色电视接收机接收时,就会产生彩色图像的黑白版。
- 一对色差信号 $m_I(t)$ 和 $m_Q(t)$, 用于显示从灰色阴影中分离出的图片颜色。

亮度信号 $m_L(t)$ 被分配给整个 4.2 MHz 的带宽。实验表明:如果色差信号 $m_I(t)$ 和 $m_Q(t)$ 的标称带宽分别为 1.6 MHz 和 0.6 MHz, 则由于人类视觉上的某种特性,可以使色彩的再现达到令人满意的效果。

再来看传真机,这种机器的作用是通过通信信道(主要是通过电话信道)传送静止的图片。这种机器为从某处向另一处发送手写或打印稿件提供了极大的方便。传真机发送文章与发送图片类似。传真机产生信号的基本原理是通过扫描原始文件(图片)并用图像传感器将光信号转换为电信号。

4. 最后一种信源就是 PC(个人电脑)。PC 在我们的日常生活中起着很大的作用,可用来收发电子邮件、交换软件和共享资源。PC 发送的文本通常用 ASCII 码(美国标准信息交换码)进行编码,ASCII 码是第一种专门为计算机通信而开发的编码方案。每个 ASCII 码的字符由 7 个数据位(bit)组成,它们由 0 和 1 组成的惟一二进制形式构成。Bit 是英文“binary digit”的首字母缩写。所以 ASCII 码共可以表示 $2^7 = 128$ 个不同的字符。这些字符包括小写和大写字母、数字、特殊控制符和 @, \$ 和 % 等常用符号。有些特殊控制符如 BS(回退)和 CR(回车)是用于控制字符打印的。其他符号如 ENQ(查询)和 ETB(块传输结束)则用于通信。完整的 ASCII 字符码见表 A6.1。数据位的排列始于 7 个数据位中的最大比特 b_7 ,一直到最小比特 b_1 , 如图 3 所示。在数据位的结尾,有一个用于错误检测的附加比特 b_8 ,这个错误检测比特称为校验位。一个 8 比特的序列称为字节。奇校验就是设置校验位使得每个字节中 1 的总数为奇数,而偶校验则是设置校验位使得每个字节中 1 的总数为偶数。例如,假设通信双方约定使用偶校验,则数据位中 1 的总数为偶数时,校验位设置为 0,否则设置为 1。因此,如果字节中有一位接收错误,则违反了奇偶校验规则,检测出后将通过重新发送来纠正错误。PC 经常会通过它们的 RS-232 端口相互连接。当 ASCII 数据(实际上是所有字符数据)通过这些端口发送时,将起始位设置为 0,并存在一个或多个结束位,将结束位设置为 1 以提供如图 3 所示的字符帧结构。当传输过程空闲时,为保持电路的连接,发送一长串 1。在图 3 中,0 和 1 分别被设计为“低电平”和“高电平”。有时它们也可以分别表示为“空号”和“传号”。后一种术语是电报时代的产物。PC 中准备好的文本通常先进行保存,然后再以一次一个字符的速率通过通信信道(如电话信道)传输。这种数据传输

方式称为异步传输。而同步传输则正好相反，在一次很长的传输过程中通过信道发送整个编码字符序列。由同步和异步终端产生的编码字符用数据复接器捆绑在一起。由此形成的复用的数据流被提供给调制解调器(Modem)，然后才能通过信道传输。

总的来说，计算机产生的数据和电视信号都是宽带信号，其功率所占频率范围较宽。PC间数据通信的另一个重要特性就是突发性，它是指信息从一个终端传输到另一个终端是突发的，在两个突发之间则是空闲的。实际上计算机间各种形式的数据通信本质上都具有突发性，这与用于语音或交互式电视的数字传输网络相反。后者相对而言都是连续的。



图 3 RS-232 标准采用的发送异步串行数据的位格式

我们使用计算机的另一种方式是从在远端的服务提供商那里下载压缩过的文本、声音和视频数据。数据压缩为数据的有效保存和传输提供了一种实用的方法。一个数据压缩系统包括编码器和译码器，分别对输入数据流进行压缩和重构。数据压缩形式基本上有两种：

1. 无损压缩：通过去除数据中的冗余信息来实现。因为这种压缩是完全可逆的，即可以准确无误地再现原始数据，所以称为无损压缩，又可称为无失真数据压缩。
2. 有损压缩：包括人为控制的信息丢失，因此这种压缩不是完全可逆的。但是有损压缩可以达到比无损压缩更高的压缩率。

数字文本要求使用无损压缩。这里，我们采用 Lempel-Ziv 压缩算法，该算法具有自适应特性，并能够对经常出现的源符号组进行编码。它能将普通英语文本压缩大约 55%，粗略地说，这种算法是通过对字母进行编码来达到压缩目的的。Lempel-Ziv 压缩算法是一种熵编码，又称信源编码。我们将在第 9 章讨论信源编码。

在很多其他应用中，有损压缩则更为常用，因为它可以充分减小数据大小而不会明显改变图像或音频信号的再现质量。在这些应用中，这种形式的数据压缩是可接受的；在像因特网这样的高吞吐量的数据传输应用中，有损数据压缩则是必需的。但在其他一些应用，如临床医疗设施，医疗图像（例如数字 X 光透视片）的再现中，信号的再现质量则不能降低。

包含存储和传输的数字音频和视频应用已活跃在当今市场上，因此需要能实现不同厂商、不同设备之间互通性的标准压缩算法。这里，我们提供三种能满足不同需求的著名标准压缩算法：

- JPEG 图像编码标准^[2]用于压缩全彩色或灰度的自然图像、真实场景，它利用了已知的人类视觉系统的局限性。JPEG 代表“Joint Photographic Experts Group”。在输入编码器时，像素被分为 8×8 的块，并使用傅里叶变换的相关形式，离散余弦变换(DCT)^[3]。DCT 将像素的每一块分解为一组 64 个系数，以近似地达到下面两个相关的目的：
 1. 系数应尽可能不相关。

2. 输入信号的能量应尽可能地包含到最少数目的系数中。

编码器的下一步操作是量化,对64个DCT系数进行舍入操作。在JPEG中,量化是根据量化表完成的,量化表作为编码器输入由用户提供。量化表中的每一个元素是一个1~255的整数,用于限定DCT系数的步长,这样就允许用8比特码字来表示每个量化的DCT系数。从根本上说,量化的目的是丢弃那些感观无法辨别的信息。量化是多对一映射,因此丢失信息主要源于编码器。编码器进行的最后一步操作是霍夫曼编码,它是熵编码的一种,将在第9章中进行讨论。霍夫曼编码是根据已量化的DCT系数的统计特性对它们进行编码,从而无损地提供额外的数据压缩。在译码器中,数据重构的一系列过程与编码器中正好相反,即霍夫曼译码,根据量化表反量化,最后反DCT。

- MPEG-1/视频编码标准^[4]主要是为将30帧/秒的视频信号压缩成1.5Mbps的比特流而设计的。MPEG代表“Motion Photographic Experts Group”。MPEG-1视频编码标准通过利用视频数据中固有的四种基本冗余,实现其设计要求:

1. 帧间(暂时的)冗余
2. 帧内像素冗余
3. 精神视觉冗余
4. 熵编码冗余

MPEG-1和JPEG的区别在于对帧间冗余的利用不同。原则上,典型视频序列中的相邻帧具有高相关度。高相关度的含义是,通常情况下,从一帧切换到另一帧时,视频信号变化不是很快,因此邻近帧之间的差异只有一个方差(即平均功率),它远小于视频信号自身的方差。因此,可大量减少帧间冗余以产生一个更为有效的压缩视频信号。这种减少是通过预测来实现的,即从邻近帧中估计每一帧。需要发送预测错误的结果,用于动态估值和补偿。这种预测在本质上是非线性的。JPEG减少像素冗余是通过综合运用DCT,量化和无损熵编码。综合的结果是全运动视频转变为1.5Mbps的计算机数据流,可以保存在光盘上,也可以同时集成文本和图片。最重要的是全运动视频和相关音频可以通过现有的计算机和电信网络传输,从而使在因特网上满足视频点播的需求成为可能。

- MPEG-1音频编码标准^[5]基于感觉编码,它是一个波形保存过程,即译码后的音频信号的幅度-时间波形与原始音频信号极其接近。一般情况下,编码过程围绕以下四个不同步骤进行:

1. 时-频映射:将输入音频信号分解到多个子带。
2. 心理声学建模:利用人类听觉系统的心理声学特性规则,同时作用于输入音频信号以计算某个门限。
3. 量化与编码:结合心理声学模型,完成时间频率映像的输出,以便将量化过程中产生的噪声保持在听觉器官感知不到的水平。
4. 帧打包:用于将量化后的音频抽样值格式化成可译码的比特流。

心理声学建模基于一种称为听觉屏蔽的感观现象。进一步讲,如果平均噪声功率低于屏蔽门限值(即刚刚能听到明显失真的门限值),则在指定的频率带宽内,人类的耳朵感觉不到量化噪声。对于一个给定的频带,屏蔽门限随带内频率变化而改变。最小屏蔽

门限是用于逐频带(band-by-band)基础上的心理声学模型。例如,在立体声节目的两个声道中使用MPEG-1标准的结果是每个原来速率为768 Kbps的数字音频信号被压缩至低达16 Kbps的速率。输入数据速率为768 Kbps,相当于抽样速率48 kHz,每个抽样值由16比特码字表示。因此,MPEG-1音频编码标准适用于将音频信号保存在廉价的媒体中,或者在保持听觉质量的同时通过带宽受限信道发射。

通信网^[6]

通信网(或简称为网络),如图4所示,由智能处理器(如微处理器)构成的多个路由器相互连接组成。这些处理器的主要用途是为数据通过网络选择路由,因而得名。每个路由器有一个或多个主机与其连接,主机是相互通信的设备。网络设计就是以有效的方式实现主机间的数据交换和资源共享,并提供支持新的应用和服务的框架。

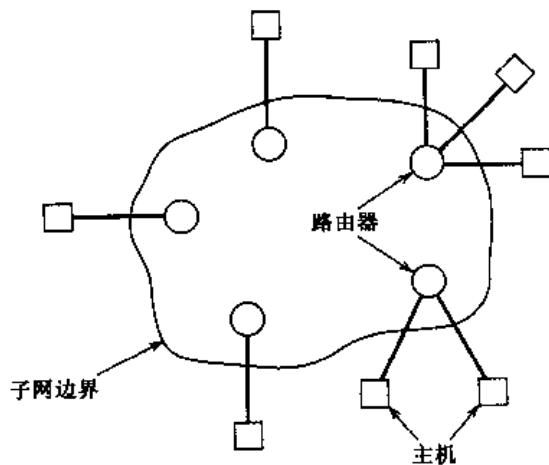


图4 通信网

电路交换用于在两个主机之间提供一种专门的通信通道和电路。电话网就是电路交换网络的一个例子。电路由从信源到信宿的一系列连接组成。例如,链路可能由许多时隙组成,这些时隙可以使多个用户能同时接入一个普通信道。此时,电路在传输过程中保持不间断。电路交换通常由具有网络组织功能的中央分级控制系统控制。要建立一个电路交换连接,需先获取一条可行的网络路径,再让两个要进行通信的主机独占地使用这个网络路径。特别地,一个呼叫请求信号必须能通过所有路径到达信宿,并在传输前予以确认。此时,网络对用户来说是有效而透明的。这表示连接建立后,分配给电路的带宽和资源本质上为两个主机所“拥有”,直到电路被断开。因此,只要分配的带宽适当就代表了资源的有效利用。尽管电话网络可以用于传输数据,但语音还是占据了网络通信的大部分。实际上,电路交换非常适合于传输语音信号,因为语音是按信号流的方式通信的,且交谈时间(平均大约2分钟)比请求建立线路连接的时间(大约0.1~0.5 s)长。此外,在大多数交谈中,信息流在连接建立的总时间中占的比例相当大,这一点使得电路交换线路更适合于语音通信。

在电路交换中,通过在不同时间段上使用基于固定分配的连接来共享一个通信链接。另一方面,在分组交换中,共享则建立在需求基础上。分组交换相对于电路交换的优势是,当连

接中需要发送信息时,连接可以更充分地利用。

分组交换网络的原理是“存储和转发”。确切地说,在分组交换网络中,任何大于指定大小的消息在发送之前都将被分成不超过指定大小的几段。这些小段通常被称为分组(或包)。原始消息在信宿端按分组被逐一重新组装。网络可以看成是网络资源(即信道带宽、缓存和交换机)的分配池,其资源动态地为所有需要进行通信的竞争主机所共享。相反,电路交换网络中,在通信的整个时间段内,资源被完全分配给一对主机。因此,分组交换更加适合于计算机通信环境,这种环境下主机间的数据突发随机地被交换了。但是,分组交换要求对用户的请求进行严格控制,否则网络会被严重滥用。

我们可以根据分层结构来看待一个网络,从而有序地进行数据网络(即主机全部由电脑和终端构成的网络)的设计,即将网络看成一个嵌套层的体系结构。层表示计算机系统中用来完成某项特殊功能的进程或装置。一般来讲,层的设计者会对其内部的细节和操作极其熟悉。但是从系统级来看,用户只把层看做用输入、输出和由输入输出之间的函数关系描述的“黑匣子”。在分层体系中,每一层将下一层视为一个或多个供上一层使用的具有指定功能的黑匣子。因此,数据网络中高度复杂的通信问题可分解为一系列便于操作的定义明确的连锁函数。正是由于这一系列原因,ISO(国际标准化组织)下属的委员会制定了OSI(开放系统互联)^[17]参考模型。开放这一术语表示任何符合这一参考模型及其相关标准的两个系统之间都能够实现互联。

在OSI参考模型中,通信和相关的连接功能由一系列层或等级组成,层间接口都有明确的定义,每一层都建立在前而层的基础上。特别地,每一层执行基本功能的一个相应子集,并依赖于下一层来执行额外的基本功能。此外,每一层为上一层提供明确的服务,并对上一层屏蔽这些服务完成的细节。在每一对层间都有接口,接口定义了下层提供给上层的服务。

OSI模型由7层组成,如图5所示。图中包括对模型中每个独立层的功能描述。系统A中的第k层根据一系列规则和约定与另一系统B中的第k层通信,这些规则和约定就相应地称为第k层协议,这里 $k=1,2,\dots,7$ 。(协议这一术语借用了它的通常用法,即描述人类普遍的社会行为。)包含不同系统相应层的实体称为对等进程。换言之,通信就是由两个不同系统中对等进程间通过协议进行通信来实现的,其中协议本身就是由一组进程的规则来定义的。两个对等进程间的物理通信仅存在于第1层。另一方面,第2层到第7层与它们的远方对等层之间是虚拟通信。不过这6层中的每一层都能通过层间接口与其邻近层(上面的或下面的)交换数据和控制信息。在图5中,物理通信用实线表示,虚拟通信则用虚线表示。

本书主要讨论OSI模型中的物理层。

因特网

任何对计算机网络的讨论都自然而然地会涉及到因特网。在因特网的例子中,通过采用对网络服务的抽象定义,基础网络技术与日常的应用被相互隔离开。进一步,我们可以说:

- 应用独立于组建网络的技术。
- 同样地,网络技术可以不断进步,但它不会影响应用。

因特网的结构如图6所示,其中有三个功能模块:主机、子网和路由器。主机构成网络节点,产生数据并发送。路由器构成用来跨接子网边界的中间节点。一个子网中主机间可以直接交换数据,相应的例子可以见图6中的子网1和子网3。

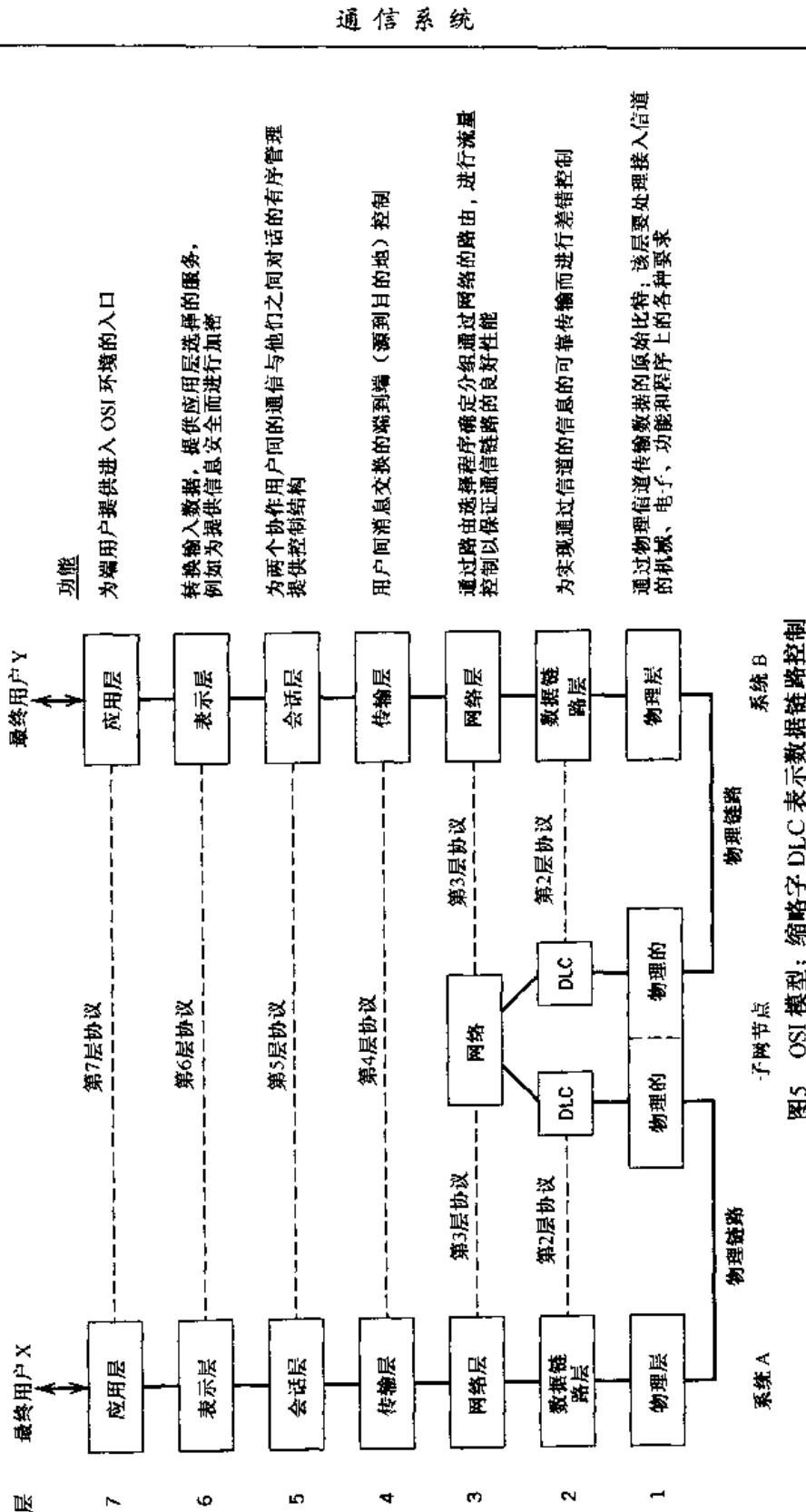


图5 OSI模型；缩略字 DLC 表示数据链路控制

像其他计算机网络一样,因特网有一套分层协议。特别地,主机和路由器间交换数据用IP(因特网协议)来完成,如图7所示。IP是一个在网络层(即OSI参考模型的第3层)通用协议。定义一个具有内置能力的点对点的以分组形式传输数据的寻址方案是一件容易的事。在穿过子网边界时,路由器判断如何对一个指定目的地址的分组进行路由选择。路由功能的实现是以路由表为基础的,路由表是通过特定协议的应用而发展起来的,应用这些协议,路由器之间可交换相关的信息。使用分层协议的结果是能提供尽可能好的(best effort)服务。即因特网能传送每一个数据分组,但是不能保证传输的时间,甚至不能确保分组是否能送达预期的接收者。

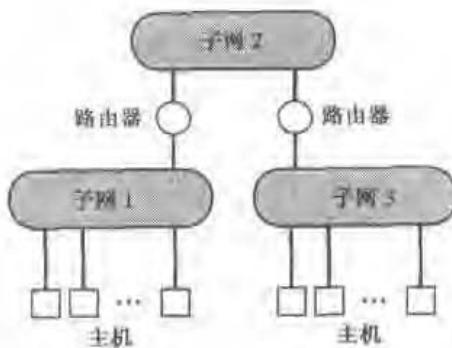


图6 子网的互联

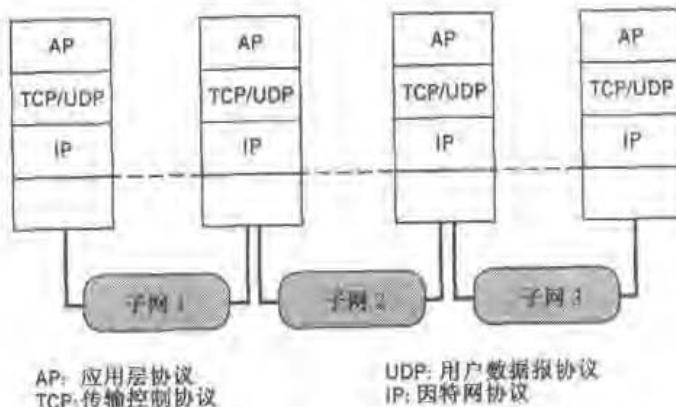


图7 因特网的网络结构

宽带网络

随着新业务(如视频点播、多媒体通信)需求的不断增长和关键技术(如光纤、数字交换)的不断应用,电话网正在向多用途宽带网发展。这种宽带网称为B-ISDN(宽带综合业务数字网)。使B-ISDN成为可能的重要技术是一种称为ATM(异步传输模式)的用户网络接口协议。ATM是一种高带宽、低时延的分组技术,用于交换和复用。这种技术的低时延特性能满足如视频之类的实时服务的需要,高带宽特性是视频点播所需要的。简而言之,ATM是一种技术(用户透明)又是一种服务(用户可见)。

顾名思义,ATM不是同步的(即依赖于一个主时钟)。它的基本思想是用一种小的、固定长度的分组——信元来传输数据信息。需要注意的是ATM的关键特性是,面向连接的服务是保

持顺序的。这就意味着重新组装信元时不要求发送到目标主机的数据流中的信元有优先顺序。在 B-ISDN 中所采用的信元交换技术相对于传统电话网中的电路交换技术来说是一个巨大的突破。

ATM 的主要目的是有效地分配本地网络资源(如带宽、缓存和处理能力),以保证每个连接的期望 QoS(服务质量)。QoS 用三个参数来衡量:

- 信元丢失率,定义为网络传输过程中,信元丢失数和发送到网络中的所有信元个数的比值。
- 信元时延,定义为一个特定连接的信元通过网络所需要的时间。
- 信元时延变化,定义为平均信元时延的弥散和抖动。

B-ISDN 中的 QoS 用来与因特网中的“best effort”服务相对应。

ATM 信元在产生后,还要组织成适于网络传输的结构。把 B-ISDN 信元放到光传输系统中就叫 SONET(同步光纤网)^[8],光纤将在下一节讨论。SONET 采用时分复用技术,光纤的实际带宽通过时间共享的办法分配给到来的各个不同的数据流,因此就需要同步。SONET 由一个精度为 $1/10^9$ 的主时钟控制,这样每一位数据就能在极精确的时间间隔内送到 SONET 线路上。不过 SONET 允许 ATM 信元无规律到达。

基本的 SONET 帧是 810 字节一组,每 $125 \mu\text{s}$ 发送一帧,总数据速率为 51.84 Mbps。每秒发送 8000 帧是为了精确匹配 8 kHz 抽样速率,8 kHz 抽样速率是语音信号在电话网上数字传输的标准抽样速率。对 51.84 Mbps 的基本数据速率进行同步字节内插后,可以产生一些数据速率的分级,汇总在表 1 中。

表 1 SONET 数据速率的分级

级 别*	数据速率(Mbps)
OC-1	51.84
OC-3	155.52
OC-9	466.56
OC-12	622.08
OC-18	933.08
OC-24	1244.16
OC-36	1866.24
OC-48	2488.32

* OC 代表光学载波(Optical Carrier)级别

通信信道

通过通信网传输信息是在网络中的物理层实现的,物理层就是信道。根据所用的传输方式的不同,通信信道基本可以分为两类:基于有线传播的信道和基于无线传播的信道。第一类信道包括电话线、同轴电缆和光纤,第二类信道包括无线广播信道、无线移动信道和卫星信道。下面分别介绍这两类共六种信道:

1. 如前所述,典型的电话网就是用电路交换方式建立一个端到端的临时通信链路。电话网的主要目的就是要确保在链路两端的交谈者之间的电话传输能替代面对面的交谈,

当然这种替代必须是可以接受的。在这种通信形式中,信源是说话人的语音,最终信宿则是听话人的耳朵。不过电话信道只能用于传输电信号,相应地就需要在系统收发两端各安装上一个合适的转换器。也就是说,放在说话人嘴边的话筒能把声波信息转换为电信号,而听话人耳边的动圈接收机则能把电信号转化为声音信号。现在这两种转换器的设计已经做得非常成熟,它们的频率响应范围可达 20~8000 Hz,此外,它们还能同时装入一个电话手柄,既能说又能听。电话信道是一个带宽受限信道,带宽的限制源于多个用户同时共享一个信道的需求。为了满足上述需求,转换器和信道特性必须符合基于主观测试的标准。这个主观测试是通过对具有代表性的男女说话人的电话信号的可理解程度和清晰度的测试来实现的。必须把语音信号(男或女)限制在 300~3100 Hz 频带范围内,此频带以外的频率部分对语音识别的贡献不大。这个频带可以看成能提供满意服务的电话信道频带的一个粗略准则,图 8 所列举的是一个典型的长途连接。图 8(a)所示为信道插入损耗和频率之间的关系图,插入损耗(以 dB 为单位)定义成 $10 \log_{10} (P_0/P_L)$,式中 P_L 表示信源通过信道传递给负载的功率, P_0 表示信源直接连在负载上所能传递的功率。图 8(b)是相应的包络时延(以 ms 为单位)和频率之间的关系。包络时延定义为角频率 $\omega = 2\pi f$ 的相位响应导数的负数。显然,图 8 清楚地描绘了一个电话信道的弥散性质。

电话信道用双绞线来传输信号。一对双绞线用两条实心的包有 PVC(polyvinylchloride)护层的铜导线绞合而成,一般每英尺绞合 2~12 次,特性阻抗为 90~110 Ω。通常把双绞线做成电缆,每根电缆由多组紧靠的双绞线组成。双绞线本质上易受电磁干扰(EMI)影响,但两条电线的绞合可以减轻这种影响。

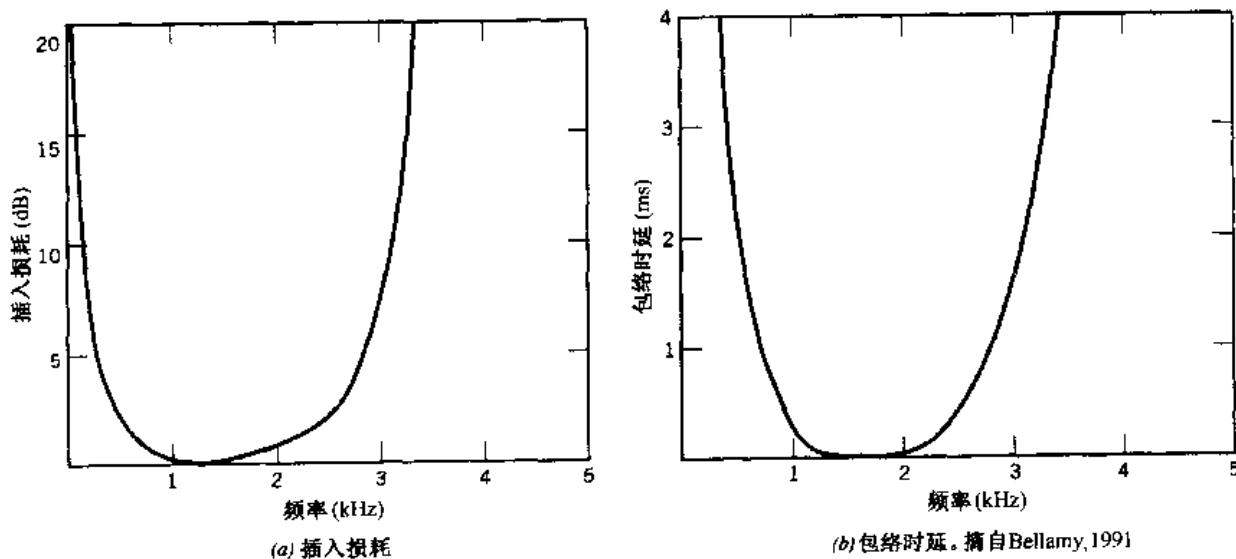


图 8 典型的电话连接的特性

2. 同轴电缆:同轴电缆由内层导体、外层导体和隔离内外层导体的绝缘隔离材料组成。内层导体一般用铜材料,它的外面包一层绝缘材料,外层导体由铜、镀锡的铜或镀铜的钢做成。普通同轴电缆的特性阻抗为 50 Ω 或 75 Ω。与双绞线相比,同轴电缆的抗电磁干扰性能更好。此外,由于带宽更高,同轴电缆所支持的数字数据传输速率要高于双

绞线。它的传输速率可达 20 Mbps,但一般标准用 10 Mbps。

双绞线的应用主要局限于点对点的服务,而同轴电缆采用高阻抗接头,因此可以作为多路接入介质使用。同轴电缆常见的应用是在办公室环境下作为局域网的传输介质。同轴电缆的另一个常见应用是有线电视,也叫 CATV(共用天线电视)系统。在这种应用中,同轴电缆将电视、音频和数字信号从中心端分配给用户。中心端是 CATV 系统的中央起始单元,信号的发送和处理都在这里进行。

3. 光纤:光纤是一种能把光信号从一端传到另一端的介质波导,这就如同双绞线和同轴电缆能传输电信号一样。光纤由内芯、包裹内芯的封套层和保护外套组成。^[9] 电磁场被限制在内芯中传播。内芯和封套层都用石英玻璃制成,保护外套用塑料制成。光纤的一些独有特性使其成为一种很有吸引力的传输媒介,尤其是它的以下一些特性:

- 巨大的潜在带宽,因为光载波频率可高达 2×10^{14} Hz,在如此高频率的载波下,带宽仅为载波的 10%,光系统的理想带宽便可达 2×10^{13} Hz,的确非常大。
- 低传输损耗,可低至 0.1 dB/km。
- 不受电磁干扰的影响,这是光纤可以看成介质波导的一个固有特性。
- 体积小、重量轻,光纤的直径与人的头发相当。
- 坚韧性,如有很高的拉伸强度,可以无损弯曲和绞合。

同样重要的是,光纤提供了一种潜在的廉价线路通信方式,因为光纤的材料是石英砂。石英砂不像金属导体,它不是稀有金属。光纤的这些特性促进了光波系统技术的显著进步,这些技术使长途通信发生了革命性的变化,并且这种变化还将继续下去。

4. 无线广播信道:无线广播信道用于传输广播和电视信号。承载有语音、音乐或图像信息的信号被调制到一个载频上,这个载频由发送基站决定。调制将在下一节介绍。传输过程的起点是天线,天线是原始调制信号和自由空间传播的电磁波之间的转换或匹配单元。设计天线的目的就是在所需的一个或几个方向上尽可能有效地激励电磁波。通常,天线安装在高塔上,这样可以无障碍地覆盖尽可能广的周边区域。由于衍射现象(一种波动的基本性质),无线电波会沿着地球表面曲线传播,这样就可能使无线电波传到视线以外,但是这样会产生比自由空间更大的衰减。

在接收端,用另一个天线捕捉发射的无线电波,以建立一条到发射机的通信链路。大多数无线电接收机采用超外差式。这种技术先把接收到的信号下变频到称为中频(IF)的合适频带,然后用相应的检波器恢复出原始的承载有信息的信号。

5. 移动无线电信道:通过其广播的能力,将移动性引入公用电话网以扩展其容量。移动无线电通常要覆盖无线电发射机或者接收机能够移动的地表环境,无论它们是否真的在移动。在多建筑物地区使用移动无线电时,对传播的主要影响是移动单元的天线远低于周围的建筑物。简而言之,就是没有了视距通信路径。更进一步讲,无线电传播是利用在周围建筑物的表面散射和衍射来实现的。最终的结果就是导致信号能量以多条传播路径到达接收天线。在移动无线电中,无线电波从不同方向和经不同时延到达的现象称为多径现象。事实上,有许多不同的传播路径,它们的电传播长度互不相同,这样就会导致接收到的信号是由从不同路径到达的信号组合而成。因而,接收信号的强度随着位置的不同以非常复杂的方式进行变化。所以,移动无线电信道从统计

本质上可以看成是一种线性时变信道。

6. 卫星信道：由于卫星信道提供对整个大陆或者大陆间的宽区域覆盖，从而为公众电信网增加了又一种宝贵的补充。此外，为那些未被陆地电缆和光纤通信覆盖的区域提供接入，也是卫星通信的一个显著特征。在几乎所有的卫星通信系统中，卫星都被安置在地球静止轨道上，此轨道必须满足下面两个要求：第一，必须是地球同步轨道，这里卫星的高度是 22 300 mile(1 mile = 1609.3 m)，地球同步卫星环绕地球一周需 24 小时（即与地球自转是同步的）；第二，卫星必须在赤道正上方飞行（即倾角为零）。从地球上看，在地球静止轨道上的卫星是静止在空中的。因而，卫星地球站不需要跟踪卫星，仅需把天线对准指向卫星的固定方向就可以了。这样整个卫星系统的设计就相对简化了。地球静止轨道上的通信卫星有以下的独特系统性能：

- 宽覆盖范围
- 稳定的传输链路
- 宽传播带宽

就提供的服务来说，卫星通信能提供固定的点对点的超长距离通信以及到遥远地区的链路和移动平台（如飞机、轮船）通信，或者说广播功能。通信卫星在把整个地球看成一个“地球村”的概念中扮演了重要的角色。在典型的卫星通信系统中，消息信号从地球站发出，通过上行链路到达卫星，然后在卫星上的转发器（即电子线路）内放大，接着从卫星转发，通过下行链路，到达另一个地球站。由于卫星定位在地球静止轨道上，对地面上卫星天线覆盖范围内的地球站来说，它是永远可见的。实际上，卫星起了一个功能强大的空中中继器的作用。最常用的卫星通信上行频带为 6 GHz，下行频带为 4 GHz。采用这个频带是因为以下的特性：

- 相对廉价的微波设备。
- 较低的降雨衰减：降雨是引起信号衰减的主要大气层原因。
- 可忽略的天空背景噪声：天空背景噪声（由银河系、太阳系和地球所发出的随机噪声）最低为 1 ~ 10 GHz。

通常，卫星专用 6/4-GHz 频带内的 500 MHz 带宽，这些带宽分配给卫星上的 12 个转发器，每个转发器大约占用 36 MHz 的卫星带宽，与一个通用无线信道相当。一个转发器至少可以传送一路彩色电视信号，1200 路语音信号或 50 Mbps 的数字数据。

总之，信道是通信系统工作的核心，它的性质决定了系统所能传送的信息容量和所能提供的服务质量。我们可以用不同的方法对通信信道进行分类：

- 信道可以分为线性的或非线性的。无线广播信道是线性的，而卫星信道是（但也不总是）非线性的。
- 信道可以分为时不变的或时变的。光纤是时不变信道，而移动无线电信道是典型的时变信道。
- 信道可以分为带宽受限的或功率受限的（即可用的发射功率是有限的），电话信道是带宽受限的，而光纤链路和卫星信道都是功率受限的。

现在我们对信源和通信信道都有了一些了解，可以返回到图 1 所示的通信系统框图中去了。

调制过程

通信系统将消息信号以一种可识别的形式从信源发送到信宿，而信源和用户在物理上是隔离开的。为了实现这个传递功能，发射机就需要把消息信号变为适于在信道中传输的形式，这个变换是通过一种称为调制的过程来实现的，即根据消息信号的不同来改变载波的一些参数。接收机则把信道传输过来的质量已经下降的信号再生为原始消息信号的形式。这个再生过程由解调过程来实现。解调是发射机调制过程的逆过程。不过，由于不可避免的噪声的出现和接收信号的畸变，接收机不能完全精确地再生原始消息信号，最终整个系统性能会受所采用的调制方式影响。明确地讲，我们发现一些调制方式受噪声和畸变的影响要比其他调制方式小一些。

调制分为连续波调制和脉冲调制。连续波调制(CW)用正弦信号作为载波。其中，载波幅度随消息信号变化的调制方式称为幅度调制(AM)；载波角度随消息信号变化的调制方式称为角度调制(Angle Modulation)。角度调制可以进一步分为频率调制(FM)和相位调制(PM)，在这两种角度调制方式中，载波的频率和相位分别随消息信号的变化而瞬时变化。

脉冲调制的载波由周期性矩形脉冲序列构成。脉冲调制可以是模拟调制，也可以是数字调制。在模拟调制中，载波的幅度、持续时间或者位置随模拟信号的抽样值进行变化，对应的调制方式分别称为脉冲幅度调制(PAM)、脉冲持续时间调制(PDM)和脉冲位置调制(PPM)。

数字脉冲调制的标准形式叫做脉冲编码调制(PCM)，PCM 中没有相应的连续波。PCM 是本质上来源于 PAM 但又更为重要的调制方式。每个已调脉冲的幅度(也就是原始消息信号的抽样值)被量化为一个最接近的离散幅度电平值，并编码为一组对应的二进制符号，二进制符号 0 和 1 用适于信道传输的脉冲信号形式来表示。无论如何，量化都会导致部分信息的丢失，使原始信息不能完全精确恢复。不过只要量化级(离散的幅度)足够多，就能使因量化而产生的失真不被人耳(传送语音信号时)或者人眼(传送二维图像时)识别出来。在所有的不同调制方式中，推荐采用 PCM 来传输模拟信号。其原因在于：

- 在规定时间间隔下，从噪声环境中再生发射信号时，系统具有鲁棒性。
- 操作的灵活性。
- 可以把多种信源集成为一种格式。
- 在信源到信宿的传输过程中，信息是安全的。

在介绍调制思想的时候，我们强调的一个重点是要在给定信道上确保消息信号的传输，另一个重点就是由于使用调制而可以采用复用。复用就是把几个消息信号组合在一起同时在一个信道上传输。下面是三种常用的复用方式：

- 频分复用(FDM)，通过分配给载波不同的载波频率，用连续波把各个消息信号移到信道通带内的一个特定频率间隔中。在接收端，用一组滤波器把不同的已调信号分离开，为各自分别解调做好准备。
- 时分复用(TDM)，利用脉冲调制把不同消息信号的抽样值放到互不重叠的时隙上。
- 码分复用(CDM)，这种方式中每一个消息信号通过不同的码来区分。

在频分复用中，从时域上看，消息信号在信道输入端是相互重叠的。因此，如果信道是非

线性的,系统就会发生串话(即消息信号间的相互干扰)。在时分复用中,消息信号在频率上重叠,时间上分离。在码分复用中,各个消息信号在信道的时间和频率上都是相互重叠的。

这里也要提到专门用于光纤的波分复用(WDM)。通过接入光纤的不同波长光谱(即不同的颜色)同时工作,使波长也成为一种新的自由度。不过,考虑到电磁波的波长和频率是成反比的,我们可以说波分复用是频分复用的一种形式。

模拟和数字通信

通常,在设计一个通信系统时,信源、信道和信宿都是指定的。主要的任务就是按如下原则设计发射机和接收机:

- 将信源产生的消息信号进行编码(或调制),发送到信道,然后在接收机输出端得出一个能满足终端用户的信号“估计值”。
- 所有功能的实现费用都应该在可接受的范围内。

我们可以选择采用数字或者模拟的通信系统。

首先考虑数字通信系统,方框图如图 9 所示,其基本原理是建立在信息论基础之上的。从信道远端开始,发射机和接收机的功能模块均是成对出现的:

- 信源编码器/译码器
- 信道编码器/译码器
- 调制器/解调器

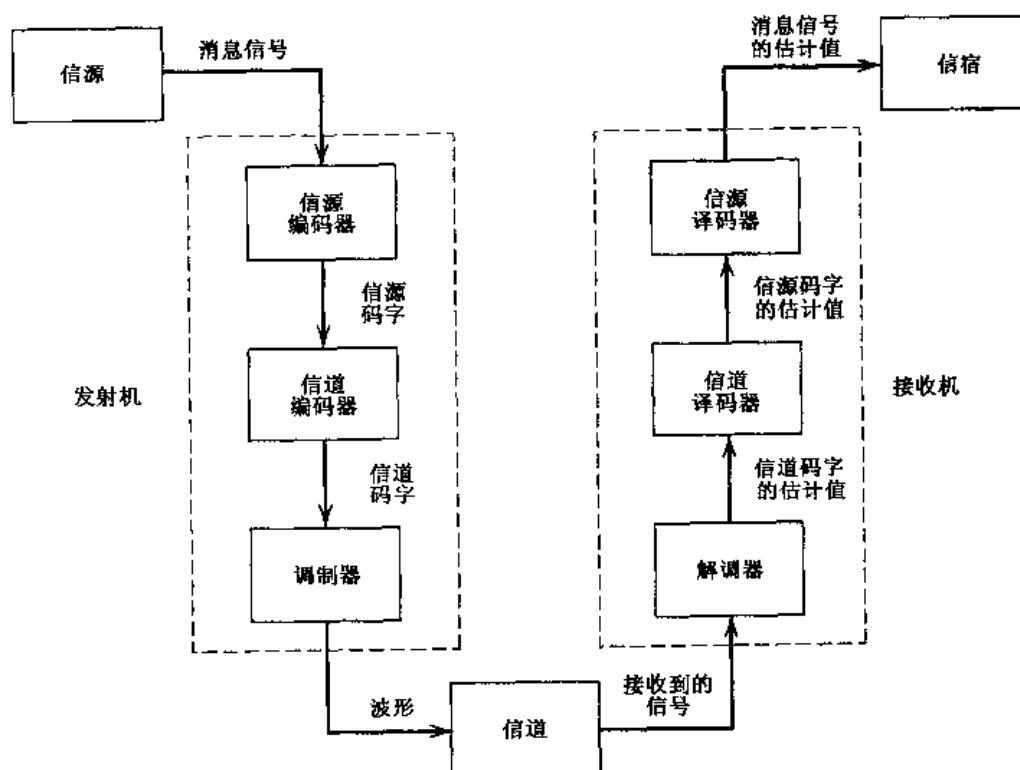


图 9 数字通信系统框图

信源编码去除消息信号中的冗余信息,以便有效利用信道。完成信源编码后的字符串称为信源码字。信源码字在信道编码器中进行处理。信道编码器生成一个新的称为信道码字的符号序列。由于加入了用于控制的冗余信息,信道码字比信源码字要长。最后,调制器把每一个信道码字符调制成相应的模拟符号。这个模拟符号是从有限个可能的模拟符号中适当地挑选出来的。调制器产生的适于在信道中传输的模拟符号串称为波形。对接收机中信道输出(即接收到的信号)的处理过程与发射机中相反,从而复原出一个可识别的原始消息信号,复原的消息信号在信宿最终传给用户。从上述描述可以看出,数字通信系统在概念上是相当复杂的,但是实际上建立起来并不难。此外,数字通信系统具有鲁棒性,与模拟系统相比,对自然界影响(如温度变化、老化、机械振动等)的承受能力更强。

相反,模拟通信系统的设计在概念上很简单,但建立起来就有些困难,因为它对线性度和系统调试的要求比较苛刻。比如,语音通信就要求非线性失真至少比所要消息信号低 40 dB。在信号处理方面,发射机由调制器构成,接收机由解调器构成,它们的具体情况主要由所采用的连续波调制类型决定。

模拟通信概念之所以简单,是由于模拟调制技术只对消息信号进行了相对简单的变换,例如已在广播和电视领域得到广泛应用的一些技术。更明确地讲,系统设计没有在调整发射信号波形以适应信道传输方面做更深层次的努力。此外,数字通信系统应尽力找出一组有限的波形以尽可能地匹配信道特性,从而增强系统承受信道损伤的能力,以便在信道上建立起可靠的通信。在选择适于噪声信道中数字通信的波形时,整个设计仅考虑了信道特性的影响。一旦选定了适于信道传输的波形,源信息就能被编码为信道波形,并且能够确保从信源到信宿的有效传输。总之,采用数字通信可以保证信息传输的有效性和可靠性。

从上述讨论可以看出,采用数字通信需要相当数量的电子线路。但随着制造在硅片上的超大规模集成电路(VLSI)技术的不断发展,电子元器件已经不再昂贵了,过去选择采用模拟通信或数字通信时所涉及的价格因素已经不是问题了。

虽然当前的趋势是采用数字通信,但基于以下两个重要原因,我们还需要学习模拟系统:

1. 既然我们通过广播和电视仍接触到模拟通信,就还有必要理解这些通信系统是如何工作的。此外,学习模拟调制可以激发我们学习数字调制技术的兴趣。
2. 模拟器件和电路的一个特性是能在很高的速率下运行,且与数字器件和电路相比功耗非常低。因此,以高速或者低功耗为原则的通信系统应当选用模拟方式。

香农信息容量定理

通信系统设计人员的目的就是配置一个系统,以便把消息信号从信源通过有噪声信道传递到信道另一端的用户,并兼顾以下原则:

消息信号能被有效和可靠地发送给用户,并能服从一些设计限制,如允许的发射功率、可用的信道带宽和可承受的系统造价。

在通信系统中,可靠性通常用接收机输出端测得的误比特率(BER)来衡量。更确切地说,误比特率越小,系统的可靠性就越好。我们就自然会考虑到这样一个问题:是否可以设计一个通信系统,在噪声信道下也能达到误比特率为零?在理想情况下,问题的答案显然是肯定的,

这个回答就包含于著名的香农定理^[10]之一的信息容量定理之中。

假设信道带宽为 B , 接收信噪比为 SNR。理论上信息容量定理可以用这两个参数来描述:

$$C = B \log_2(1 + \text{SNR}) \text{ bps} \quad (1)$$

这里 C 表示信道的信息容量。信息容量定义为通过信道无差错地传输信息的最大速率, 用每秒的比特数来衡量。对于给定的信道带宽 B 和接收信噪比, 通过信息容量定理可知, 即便是在信道有噪声的情况下, 消息信号也可以无差错地通过通信系统。假如实际信号速率为 R (bps), 则在有噪声信道中, R 要小于信息容量 C 。

遗憾的是, 香农信息容量定理没有告诉我们如何去实现这个系统。尽管如此, 从设计的角度来看, 这个理论还是很有价值的。这是因为:

1. 信息容量定理提供了一个在给定信道带宽 B 和接收信噪比的情况下, 理论上能达到的数据传输速率。基于这一点, 我们用比值

$$\eta = \frac{R}{C}$$

来衡量所研究的数字通信系统的有效性。 η 越接近 1, 系统的有效性越好。

2. 式(1)提供了一个在信道带宽 B 和接收信噪比之间进行权衡的基础。明确地讲, 对给定的信号速率 R , 我们可以通过增加信道带宽来降低对信噪比的要求, 这就促使我们采用宽带调制方式(如 PCM)来改善系统的抗噪声性能。
3. 式(1)提供了比较两种调制方式的抗噪声性能的理想化框架。

一个数字通信问题

我们说数字通信具有较低的误比特率是指接收机译码时, 在一长串二进制符号流中仅有小段发生了错误。接收机判定一个通过有噪声信道发送并译码的二进制符号是否正确, 是数字通信系统最重要的问题。接下来我们对这个问题进行简要的讨论, 以激发读者研究通信系统的兴趣。

假定有一随机信号 $m(t)$, 由符号 1 和 0 大致等概地组成, 符号 1 常用电平 +1 表示, 符号 0 则常用电平 -1 表示。每个符号均持续一段时间 T , 这样的信号就能够表示数字计算机的输出或者数字化的语音信号。为便于信号在信道中的传输, 采用了 PSK(相移键控)的简单调制方式。明确地说, 调制后的信号用承载有信息的信号 $m(t)$ 与正弦载波信号 $A_c \cos(2\pi f_c t)$ 的乘积表示。其中 A_c 表示载波幅度, f_c 表示载波频率, t 表示时间。图 10(a)是一个发射机的框图, 图中发射机的输出定义为

$$s(t) = \begin{cases} A_c \cos(2\pi f_c t) & \text{符号 1} \\ -A_c \cos(2\pi f_c t) & \text{符号 0} \end{cases} \quad (2)$$

其中 $0 \leq t \leq T$, 载波频率 $f_c = 1/T$ 。

如图 10(b)所示的信道, 假定除了噪声外没有其他失真。图中的接收信号 $x(t)$ 定义为

$$x(t) = s(t) + w(t) \quad (3)$$

式中 $w(t)$ 是加性信道噪声。

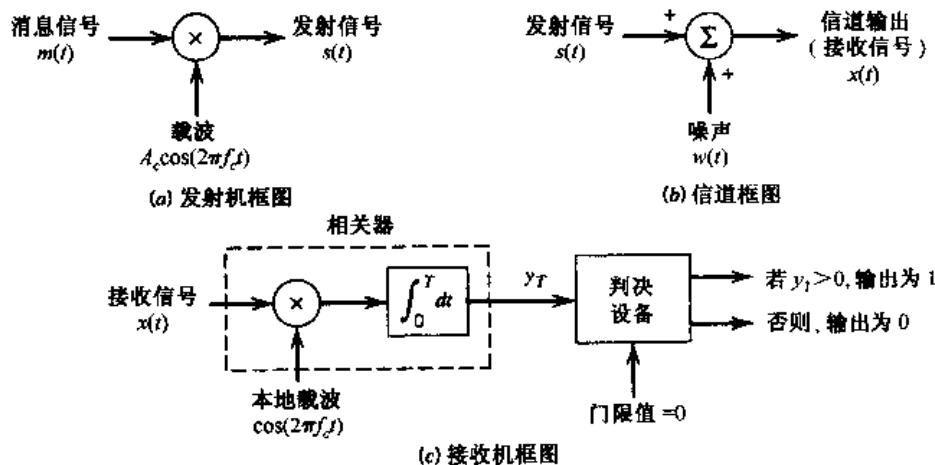


图 10 数字通信系统各要素

图 10(c)所示的接收机由一个相关器和连在其后的判决设备组成。相关器先把接收到的信号 $x(t)$ 与本地产生的载波 $\cos(2\pi f_c t)$ 相乘, 然后在 $0 \leq t \leq T$ 上对乘积求积分, 得到输出

$$y_T = \int_0^T x(t) \cos(2\pi f_c t) dt \quad (4)$$

将式(2), 式(3)代入式(4)中, 结合载波频率 $f_c = 1/T$, 经化简可得

$$y_T = \begin{cases} +\frac{A_c}{2} + w_T & \text{对于符号 1} \\ -\frac{A_c}{2} + w_T & \text{对于符号 0} \end{cases} \quad (5)$$

此处相关器输出中的 w_T 分量是信道噪声 $w(t)$ 造成的, 为复原原始二进制信号 $m(t)$, 相关器输出中 y_T 分量在判决设备中与 0 V 阈值进行比较。判决设备的判决规则如下:

如果相关器的输出 y_T 大于 0, 则接收机输出符号 1, 否则输出符号 0。

经过上述讨论, 可得出以下基本结论。通过傅里叶分析可以发现脉冲信号的时间和带宽的乘积是一个常数, 这就表示持续时间为 T 的矩形脉冲的带宽与 T 成反比关系。图 10(a)的发射信号就是这种矩形脉冲和正弦载波信号 $A_c \cos(2\pi f_c t)$ 的乘积。信号和正弦信号相乘的效果就是把信号的傅里叶变换向左右各平移 f_c , 但信号的幅度只有原来的 $1/2$ 。这就可以得出, 发射信号 $m(t)$ 的带宽也就是信道所需带宽, 与符号持续时间 T 的倒数成正比。在这个问题中, T 的倒数也就是系统的信号速率, 单位为 bps。

不过, 还有一些其他问题需要在理论上进行分析:

1. 采用如图 10(c)所示接收机结构的理由是什么?
2. 根据式(3)和式(4), 噪声分量 w_T 是一个随机变量 W 和某个真实信道噪声 $w(t)$ 在 $t = T$ 时刻的抽样值的乘积。那么, 如何把随机变量 W 的统计特性和信道噪声的统计特性联系起来?
3. 如图 10(c)所示, 由于相关器输出的随机性, 接收机会发生一些偶然性的错误, 如把实际发送

是 1 的符号判断成符号 0, 或者发生相反的情况。那么此类错误的发生概率又是多少?

此外, 还有一些重要的实际问题需要引起重视:

1. 信道带宽是极其珍贵的资源, 那么, 如何选择一种能有效利用带宽的调制方式?
2. 二进制信号 $m(t)$ 中还包含着一些冗余符号, 这些符号是为了防止信道噪声的干扰而在信道编码时加入的。那么, 该如何用一种物理可实现的方式设计发射机的编码器和接收机的译码器, 使得信道容量非常接近香农信道容量定理得出的理论值?
3. 图 10(c) 中, 接收机本地产生的载波和发射机用于调制的载波是分开的。而再生原始二进制信号时的判决是基于等式(4)的, 这就需要接收机和发射机的载波相位、符号和定时均同步。这些同步如何实现?

图 10 所示的简单数字通信系统所引出的理论和实际问题将在接下来的几章中进行讨论。

大事记^[11]

为完整起见, 下面对通信的历史进行简要介绍。这里列出一些与通信相关的重大历史事件, 希望能给读者以启发和鼓励。

1837 年, 画家莫尔斯(Morse)改进了电报机。1844 年, 他又用电报机在华盛顿特区和马里兰的巴尔的摩之间发送了内容为“What hath God wrought”的信息, 从而引发了实时远距离通信方式的彻底革命。由于莫尔斯码是采用点、线、短空格和长空格 4 种符号的可变长度码, 短串代表常用字母, 而长串代表非常用字母。电报可以说是数字通信的前身。这种信号发送适用于人工键盘输入。1875 年, Emile Baudot 开发出电报定长二位码。适用于电传机的 Baudot 电报码中, 每一个码字都是由 5 个等长的电码元素组成, 而每个电码元素有两种可能的状态: 一个传号或一个空号(也就是今天术语中的符号 1 和 0)。

1864 年, 麦克斯韦建立了光的电磁理论并预言了无线电波的存在。因此, 基于此理论的一组方程式以他的名字命名。1887 年赫兹用实验证明了无线电波的存在。Oliver Lodge 于 1894 年展示了相当短距离(150 yd, 1 yd = 0.9144 m)内的无线通信。1901 年 12 月 12 日, Guglielmo Marconi 在纽芬兰的 Signal 山接收到了来自 1700 mile(1 mile = 1609.3 m)以外大西洋彼岸的英国 Cornwall 发送的无线电信号, 极大地拓宽了通信的范围。1906 年, 自学成才的 Reginald Fessenden 历史性地第一次进行了无线电广播。

1875 年, 聋哑学校教师贝尔发明了电话。电话通过电编码和声音真实回传使语音实时传输成为现实。最初的电话简陋且信号微弱, 只能进行短距离通话。电话投入使用仅仅几年, 人们就对电话的自动化产生了兴趣。特别是在 1897 年, 来自密苏里州堪萨斯城的殡仪员 Strowger 发明了以自己名字命名的步进开关。在随后的许多年中, Strowger 开关在所有的电子机械开关中一直最受欢迎, 并被广泛使用。

1904 年, John Ambrose Fleming 发明了真空二极管, 为 1906 年 Lee de Forest 发明真空三极管开辟了道路。真空三极管的发明促进了 1913 年洲际电话的开发, 标志着无线有声通信的来临。的确, 在晶体管出现和完善之前, 真空三极管是设计电子扬声器的最佳器件。

1918 年, Edwin H. Armstrong 发明了超外差式无线电接收机。直至今日, 几乎所有无线电接收机都属于这一类型。Armstrong 又于 1933 年提出了其称为频率调制系统这一革命性的概念,

并于 1936 年发表了制作调频无线电的论文。

1928 年, Philo T. Farnsworth 首次展示了全电子电视系统。Vladimir K. Zworykin 于 1929 年再一次演示了该系统。到 1939 年, 英国广播公司(BBC)开始发送电视信号, 进行商业运作。

1928 年, 奈奎斯特发表了关于电报通信中信号传输的经典论文。奈奎斯特还特别制定了在无噪声条件下正确接收通过弥散信道传输的电报信号的标准。奈奎斯特很多早期的成果后来都应用于在弥散信道上进行的数字数据传输。

1937 年, Alec Reeves 发明了 PCM, 用于语音信号的数字编码。该技术于二次世界大战时得到改进, 使语音信号加密成为可能。实际上, 二战末期美军已经在战场上使用了一个完全的 24 信道系统。尽管如此, 直到晶体管的发明以及随后的大规模集成电路的开发之后, PCM 的商用价值才真正得以体现。

1943 年, D.O. North 设计出匹配滤波器, 用于在加性白噪声下对一个已知信号进行最佳检测。1946 年, J.H. Van Vleck 和 D. Middleton 也独立地得出了相似的结果。

1947 年, V.A. Kotelnikov 在呈给莫斯科 Molotov 能源研究所学术委员会的博士论文中提出了信号的几何描绘方法。随后, 在 John M. Wozencraft 和 Irwin M. Jacobs 于 1965 年出版的具有里程碑意义的教科书中, 这一方法臻于完美。

1948 年, 香农的论文“*A Mathematical Theory of Communication*”为数字通信奠定了理论基础。该论文一发表, 就迎来了一片赞扬声。也许正是这赞扬声, 促使香农在次年与 Warren Weaver 合作出书时, 将书中这篇论文更名为“*The Mathematical Theory of Communication*”。值得一提的是, 在 1948 年发表这篇经典论文之前, 人们还认为增加信道的信息传输速率必然会增加差错概率。但香农证明了并非如此, 理论上只要信息传输速率低于信道容量就可获得接近无差错的传输。这一结论让通信理论界大吃一惊。在该论文发表之后, 编码理论又取得了如下几项重大进步:

- M.J. E. Golay 于 1949 年, Richard W. Hamming 于 1950 年开发了第一个重要的纠错码。
- C. Berrou, A. Glavieux 和 P. Thitimajshima 于 1993 年发明了 Turbo 码。Turbo 码在香农意义上提供了近乎最佳的纠错编码和译码性能。

Walter H. Brattain, John Bardeen 和 William Shockley 于 1948 年在贝尔实验室发明了晶体管。Robert Noyce 于 1958 年制造了第一个硅芯片集成电路。这一由固态物质和集成电路构成的划时代的新发明也促进了超大规模集成电路(VLSI)和单芯片微处理器的诞生。与此同时, 信号处理和通信产业也发生了根本性的转变。

1948 年, 晶体管的出现激发了电路在交换与数字通信中的应用, 其目的是为了增强可靠性、增加容量和降低成本。1958 年 3 月, 在贝尔实验室, 人们第一次通过存储程序系统拨通了电话。1960 年, 伊利诺州 Morris 市首先开通了数字交换的商业电话服务业务。贝尔实验室于 1962 年安装了第一个 T1 载波传输系统。

1943 年至 1946 年间, 费城大学摩尔电气工程学院在 J. Presper Eckert, Jr. 和 John W. Mauchly 的技术指导下, 制造了第一台名为 ENIAC 的电子数字计算机。根据 1945 年起草的一份报告初稿所述, John von Neumann 为早期数字计算机理论、设计和应用做出了重大的贡献。50 年代早期, 计算机和终端开始实现了远距离通信。早期用于连接的是传输速度慢、以声音分级的电话信道(300 ~ 1200 bps)。很多因素带来了数据传输率的显著提高, 其中最重要的有 1965 年由 Robert Lucky 率先提

出的自适应均衡理论和 G.Ungerboeck 于 1982 年开始倡导的高效调制技术。另一在计算机通信中广为应用的理论是自动重复请求(ARQ)。ARQ 由 H.C.A.van Duuren 于二战期间发明并于 1946 年发表,该方法适用于改进远距离电报传输的无线电话技术。

1950 年至 1970 年间,人们对计算机网络进行了各种研究。就对计算机通信所产生的影响而言,其中最重要的一项成果是于 1971 年投入使用 ARPANET(Advanced Research Project Agency Network)。ARPANET 由美国国防部下属的高级研究计划局支持开发。早期的分组交换就是在 ARPANET 上实现的。1985 年 ARPANET 更名为因特网(Internet)。因特网演变的转折点是在 1990 年,Tim Berners-Lee 提出在因特网中加入超媒体软件接口,也就是他所说的万维网。^[12]万维网从诞生之日到受到全世界的欢迎,仅用了不到两年的时间,就于 1994 年达到了商业运营最大化。我们如何解释因特网爆炸性增长的现象呢?原因可能有以下几点:^[13]

- 在万维网爆炸性诞生之前,促使其产生的因素就已经出现。特别是超大规模集成电路的出现,使 PC 得到了普及,越来越多的 PC 装上了 Modem,用于与外部世界进行联系。
- 因特网已稳步发展了近 20 年(虽然仅限于有限用户群),用户数已达到基于电子邮件和文本传输的用户数临界值。
- 文本描述和传输的标准,即超文本链接标记语言(HTML)和超文本传输协议(HTTP)得以采用。

因此,创造万维网的各种条件准备就绪,除了两个关键因素:一个简单的用户接口和一个绝妙的服务理念。

1955 年,John R.Pierce 建议利用卫星进行通信。而在这个提议之前,早在 1945 年,Arthur C.Clark 就在其论文中提出了利用一颗绕地球运转的卫星作为两个地面站之间的通信转接点。1957 年,苏联发射了 Sputnik I 卫星。该卫星可传送遥测信号达 21 天之久。美国于 1958 年发射了 Explorer 1 卫星,该卫星可传送遥测信号近 5 个月。1962 年 7 月 10 日,Telstar I 卫星在 Canaveral 角发射升空,卫星通信技术迈出了实验性的重大一步。Telstar I 卫星由贝尔实验室承建,其间从 Pierce 的前期工作中获得了大量的经验。该卫星能够实现跨越大西洋的电视信号转发,用户使用参量接收机和大型天线即可接收。

运用光学原理传输信息的技术可追溯到史前时代(如用烟或火为信号)。但直到 1966 年,英国标准电话实验室的 K.C.Kao 和 G.A.Hockham 提出用涂层玻璃纤维作为一种绝缘体波导,光通信才有了突破性的发展。激光(Laser, light amplification by stimulated emission of radiation, 即受激辐射引起的光放大)于 1959 年问世,并在 1960 年得到改进。Kao 和 Hockham 指出:(1)光纤中的衰减是由于玻璃中成分不纯;(2)由瑞利散射导致的损耗很低。他们预言 20 dB/km 的损耗是可以实现的。在做出这一惊人的预言时,玻璃纤维中能量损耗大约有 1000 dB/km,这一预言后来得以证实。现在,传输中的损耗已经低至 0.1 dB/km。

我们目睹了微电路、数字计算机、光波系统领域的蓬勃发展,这些领域的技术今后仍将不断地推陈出新。这一切造就了电信环境的巨大变革。很多变革正在进行,更多的变革将随着时间的推移不断涌现。

注释与参考

[1] 关于描述早期通信和其他相关学科(如电子学、计算机、雷达、无线电天文学、卫星

等)的论文,见 Overhage(1962)。尤其是 L.V.Berkner 撰写的“通信”一章(第 35 至 50 页)。

- [2] JPEG 图像编码标准由 Wallace 于 1991 年进行了讨论。另见 Madisetti 和 Williams 于 1998 年编写的手册中 T.A.Ramstad 所写的文章。
- [3] 8×8 源像样值的 DCT 及其反变换的定义分别如下:

$$F(u, v) = \frac{1}{4} C(u)C(v) \left[\sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos\left(\frac{(2x+1)u\pi}{16}\right) \cos\left(\frac{(2y+1)v\pi}{16}\right) \right]$$

$$f(x, y) = \frac{1}{4} \left[\sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v) F(u, v) \cos\left(\frac{(2x+1)u\pi}{16}\right) \cos\left(\frac{(2y+1)v\pi}{16}\right) \right]$$

其中

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & u = 0, v = 0 \\ 1 & \text{其他} \end{cases}$$

要全面了解 DCT, 见 Rao 和 Yip(1990)。

- [4] MPEG-1 视频编码标准由 Gall 于 1991 年在论文中提出讨论。另见 Madisetti 和 Williams(1998)编写的手册中 A.M.Tekalp 的文章, 文中还讨论了 MPEG 视频编码标准的后续版本。
- [5] MPEG-1 音频编码标准由 Brandenburg 和 Stoll 于 1994 年、Pan 于 1993 年提出讨论, 另见 Madisetti 和 Williams(1998)编写的手册中 P.Noll 的文章, 文中还讨论了 MPEG 音频编码标准的后续版本。另外, 广泛使用的更为流行的现代标准 MPEG-3 音频标准一定程度上是为了解决侵犯版权的问题而产生的, 它可减少以往磁带被盗版的问题。
- [6] 更详细的关于网络的讨论见 Tanenbaum(1996)。
- [7] OSI 参考模型是由 ISO(国际标准化组织)的附属委员会于 1977 年开发的, 7 层模型每一层的原则和描述见 Tanenbaum(1996)。
- [8] SONET 最初由 Telcordia 技术公司(以及后来的 Bellcore)提出, 并由 ANSI(美国国家标准协会)进行标准化。后来 CCITT 采用了 SONET 标准并提出了一套推荐的并行标准, 称为同步数字序列(SDH)。SDH 和 SONET 的区别是很小的。
- [9] 要对光纤中光波的传播进行彻底精确的研究, 就需要把它当做波导电介质并使用 Maxwell 方程组进行分析, 这种分析实质上是一种高难度的数学工作。对于这种分析方法的描述见 Green,Jr.(1993)的第 3 章。
- [10] 读者可参阅 Lucky(1989)的“Silicon Dreams”一书, 书中以通俗的方式给出了信息论中香农定理的半技术性概要。
- [11] 对通信史的通俗描述见 Lebow(1995)。
- [12] 对因特网发展史的描述见 Leiner(1997)等人的文章。
- [13] 与新的电信业务以及社会对电信发展的反应相关的论文见 Lucky 于 1997 年撰写的论文。文中提出了 Metcalf 定律, 按照这一定律, 似乎任何新的电信业务都必须等待很长一段时间才能建立普遍的支持。Lucky 引用国际因特网作为 Metcalf 定律的反例, 并给出了这一现象的解释。

第1章 随机过程

本章以二阶统计量为重点,介绍了平稳随机过程。特别讨论了以下问题:

- 随机过程的概念。
- 一个随机过程是平稳随机过程必须满足的条件。
- 对随机过程的均值、相关函数和协方差函数的分别描述。
- 平稳随机过程具有遍历性必须满足的条件。这为采用集平均来代替时间平均提供了可能。
- 当一个平稳随机过程通过一个线性时不变滤波器时会发生变化。
- 采用功率谱密度对随机过程进行频域描述。
- 一种重要随机过程——高斯过程的特性。
- 噪声源及其窄带形式;
- 通信系统研究中产生的两类重要的概率分布:瑞利分布和莱斯分布。

1.1 简介

采用数学模型来描述一个物理现象的方法在物理学和工程学中得到了广泛的应用。在这里,要区分两类数学模型:确定的和随机的。如果模型的时间相关性在任一时刻都没有不确定性,那么就称这种模型为确定性模型。然而,因为物理现象的发生与太多不可知的因素有关,所以现实世界中的许多问题都不适宜采用确定性模型。尽管如此,仍然可以考虑一个用概率术语描述的模型,在这个模型中,每一个即将出现的可能值存在于两个特定的极限之间。这种模型称为随机性模型。对概率论的简要回顾见附录 A。

例如,考虑一个无线电通信系统。在这样一个系统里,接收到的信号通常包括承载信息的信号分量、随机的干扰分量和信道噪声。承载信息的信号分量代表的可能是一个声信号,这个信号往往包含持续随机时间的随机间隔的能量串。干扰分量代表的则可能是由无线电接收机附近的其他通信系统工作产生的寄生的电磁波。信道噪声的主要来源是热噪声,它是由接收设备前端导体和设备内的电子的随机运动引起的。因此,接收信号本质上是随机的,是不能事先预测其准确值的信号。但根据统计参数,例如平均功率和功率谱密度,仍可对其进行描述,这将在本章中予以讨论。

1.2 随机过程的数学定义

从上面的介绍可以看出,随机过程有两个属性:首先,它是时间的函数;其次,它在理论上是随机的,也就是说,在进行一次实验前,不可能确切定义未来将观察到的波形。

采用样本空间来描述随机实验是很方便的。每一次实验的结果对应于一个样本点。所有样本点的总和对应于实验所有可能的输出,称之为样本空间。样本空间中的每一个样本点都

是时间的函数。由时间函数构成的样本空间或者整体称之为随机过程。^[1]假设随机过程在定义在样本空间的集合上的概率分布存在，就能够对不同事件的概率进行讨论了。

接下来，讨论一个由样本空间 S 中的输出、 S 上定义的事件和这些事件的概率确定的一个随机实验。可以假定每个样本点 s 是符合下面式子的时间函数：

$$X(t, s), \quad -T \leq t \leq T \quad (1.1)$$

此处 $2T$ 为总的观察间隔。对于一个固定的样本点 s_j ，函数 $X(t, s_j)$ 对时间 t 的曲线称之为一次实现，或是随机过程的一个样本函数，为了简化符号，将样本函数表示为

$$x_j(t) = X(t, s_j) \quad (1.2)$$

图 1.1 给出了一组样本函数 $\{x_j(t) | j = 1, 2, 3, \dots, n\}$ ，从这张图中，注意到在观测间隔中的每一个固定的时刻 t_k ，这组数字

$$\{x_1(t_k), x_2(t_k), \dots, x_n(t_k)\} = \{X(t_k, s_1), X(t_k, s_2), \dots, X(t_k, s_n)\}$$

构成一个随机变量。从而，就找到了随机变量 $\{X(t, s)\}$ 所构成的一个整体，即一个随机过程。为了简化表示，在平时的实际应用中隐藏 s ，仅简单地用 $X(t)$ 来表示一个随机过程。现在，可以正式定义一个随机过程 $X(t)$ 为具有一定概率规则的时间函数的集合，这个概率规则就是为任何与随机过程的一个样本函数观察值有关的有意义事件都分配一定的概率。可以根据以下几点来区分随机变量和随机过程：

- 对于随机变量来说，随机实验的结果对应于一个数。
- 对于随机过程来说，随机实验的结果对应于时间函数的波形。

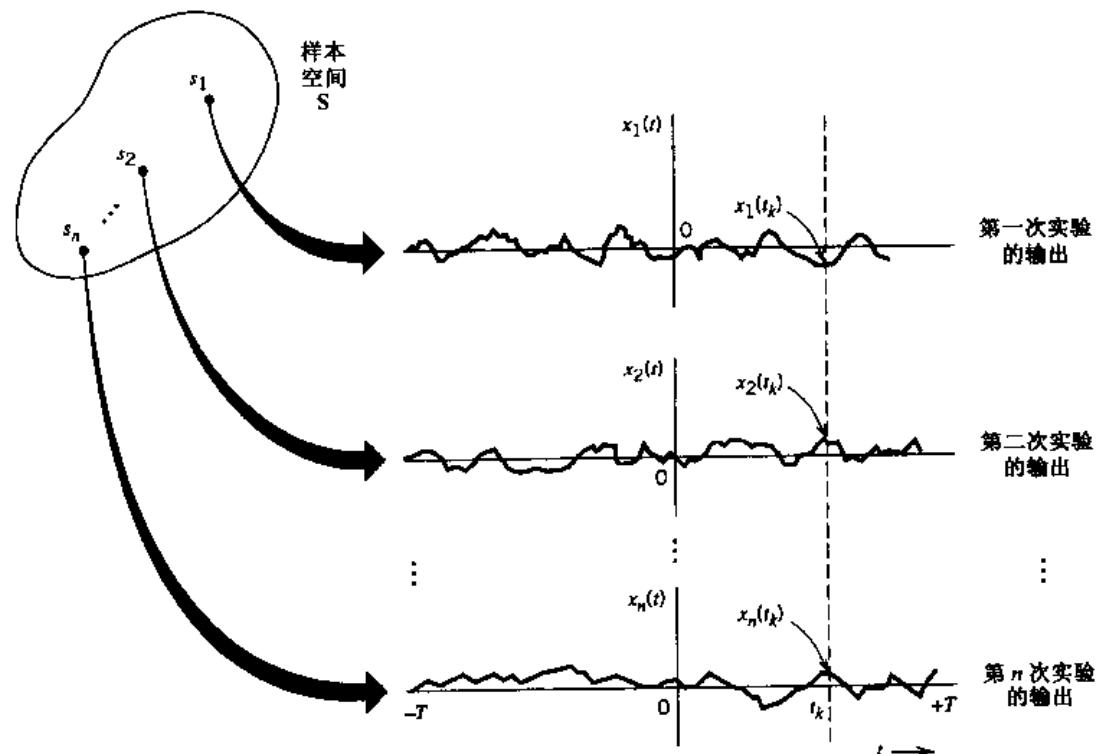


图 1.1 一组样本函数

1.3 平稳过程

在处理现实世界中所遇到的随机过程时,研究者往往发现有些过程的统计特性与观测开始的时间无关。并且,如果这个过程被分成许多时间段,不同的时间段都会显示出本质上相同的统计特性。这样的过程被称为是平稳的,反之则称为是不平稳的。一般说来,平稳随机过程源于稳定的物理现象,而非平稳随机过程则源于不稳定的物理现象。

可以更具体地考虑一个从 $t = -\infty$ 开始的随机过程 $X(t)$ 。 $X(t_1), X(t_2), \dots, X(t_k)$ 为 t_1, t_2, \dots, t_k 时刻观察随机过程所得的随机变量。这组随机变量对应的联合分布函数为 $F_{X(t_1), \dots, X(t_k)}(x_1, \dots, x_k)$ 。假定将观测时间都向前推移一个固定值 τ ,由此获得一组新的随机变量 $X(t_1 + \tau), X(t_2 + \tau), \dots, X(t_k + \tau)$ 。这组新的随机变量的联合分布函数为 $F_{X(t_1 + \tau), \dots, X(t_k + \tau)}(x_1, \dots, x_k)$ 。如果满足下面的条件,称随机过程 $X(t)$ 对所有的时间偏移量 τ 、所有的 k 和所有可能的观测时刻 t_1, \dots, t_k 是严格意义上平稳的或者说是严平稳的:

$$F_{X(t_1 + \tau), \dots, X(t_k + \tau)}(x_1, \dots, x_k) = F_{X(t_1), \dots, X(t_k)}(x_1, \dots, x_k) \quad (1.3)$$

上式对于所有的 τ, k 以及全部可能的观测时刻 t_1, \dots, t_k 都成立。

换言之,如果观测随机过程 $X(t)$ (从 $t = -\infty$ 开始)所得到的每一组随机变量的联合分布函数对于取定不同的时间原点 $t = 0$ 是时不变的,那么随机过程 $X(t)$ 是严平稳的。注意,式(1.3)中,有限维的分布只与随机变量间的相对时间差有关,而与它们的绝对时间无关。也就是说,随机过程在整个时间段上有相同的概率特性。

同样,如果两组随机变量 $X(t_1), \dots, X(t_k)$ 和 $Y(t'_1), \dots, Y(t'_j)$ 的联合有限维分布函数对所有的 k 和 j 以及观测时刻 t_1, \dots, t_k 和 t'_1, \dots, t'_j 而言,不随起始时刻 $t = 0$ 的变化而变化,则称这两个随机过程 $X(t)$ 和 $Y(t)$ 是联合严平稳的。

以下是式(1.3)的两种特殊情况:

1. 令 $k = 1$, 有

$$F_{X(t)}(x) = F_{X(t+\tau)}(x) = F_X(x) \quad \text{对于所有 } t \text{ 和 } \tau \quad (1.4)$$

也就是说,平稳过程的一维分布函数与时间无关。

2. 令 $k = 2$, 并令 $\tau = -t_1$, 有

$$F_{X(t_1), X(t_2)}(x_1, x_2) = F_{X(0), X(t_2 - t_1)}(x_1, x_2) \quad \text{对于所有 } t_1 \text{ 和 } t_2 \quad (1.5)$$

也就是说,平稳过程的二维分布函数只与观测的时间差有关,与观测的具体时间无关。

这两个特性对平稳过程的统计参数化具有非常深远的意义,这个问题将在 1.4 节中进行讨论。

例 1.1

假设图 1.2 在时刻 t_1, t_2, t_3 有三个空间窗。希望得到随机过程 $X(t)$ 的一个样本函数 $x(t)$ 穿过这组空间窗的概率,也就是联合事件概率

$$A = \{a_i < X(t_i) \leq b_i\}, \quad i = 1, 2, 3$$

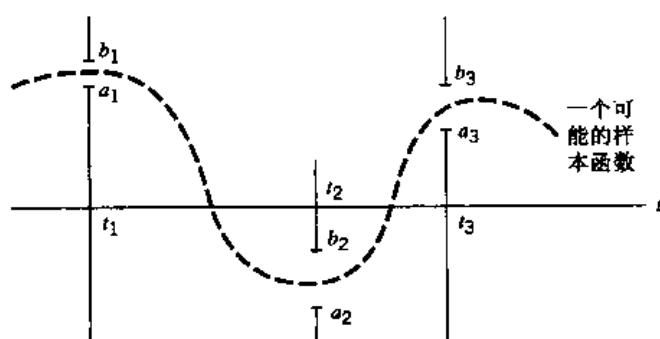
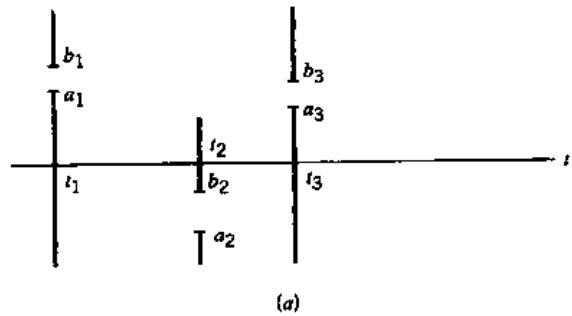


图 1.2 阐述一个联合事件的概率

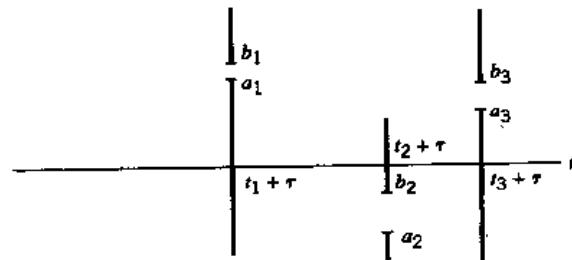
根据联合分布函数, 所求概率为

$$P(A) = F_{X(t_1), X(t_2), X(t_3)}(b_1, b_2, b_3) - F_{X(t_1), X(t_2), X(t_3)}(a_1, a_2, a_3)$$

现在假定随机过程 $X(t)$ 是严平稳的, 这就表示穿过图 1.3(a) 中时间窗的样本函数组的概率与穿过对应的图 1.3(b) 中平移后的时间窗的样本函数组的概率相等。不过应注意, 这并非说两组包含同样的样本函数。



(a)



(b)

图 1.3 阐述例 1.1 中平稳性的概念

1.4 均值、相关函数和协方差函数

设有一个严平稳的过程 $X(t)$, 定义 $X(t)$ 的均值为在某一时间 t 处观测所得到的随机变量的数学期望, 如下式:

$$\begin{aligned}\mu_X(t) &= E[X(t)] \\ &= \int_{-\infty}^{\infty} xf_{X(t)}(x) dx\end{aligned}\quad (1.6)$$

这里 $f_{X(t)}(x)$ 是过程的一维概率密度函数。由式(1.4)可推知,对于一个严平稳的随机过程, $f_{X(t)}(x)$ 是独立于时间 t 的。因此,严平稳随机过程的均值是一个常量,如下式:

$$\mu_X(t) = \mu_X \quad \text{对于所有 } t \quad (1.7)$$

定义 $X(t)$ 的自相关函数为在 t_1 和 t_2 时刻观察 $X(t)$ 所得的两个随机变量 $X(t_1)$ 和 $X(t_2)$ 乘积的数学期望,记为

$$\begin{aligned}R_X(t_1, t_2) &= E[X(t_1)X(t_2)] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{X(t_1), X(t_2)}(x_1, x_2) dx_1 dx_2\end{aligned}\quad (1.8)$$

在这里 $f_{X(t_1), X(t_2)}(x_1, x_2)$ 是过程的二维概率密度函数。由式(1.5)可推知,对于一个严平稳的随机过程, $f_{X(t_1), X(t_2)}(x_1, x_2)$ 只与观测时间 t_1 和 t_2 的差有关。由此可得出严平稳随机过程的自相关函数只与时间差 $t_2 - t_1$ 有关的结论,如下式:

$$R_X(t_1, t_2) = R_X(t_2 - t_1) \quad \text{对于所有 } t_1 \text{ 和 } t_2 \quad (1.9)$$

同样,严平稳随机过程 $X(t)$ 的自协方差函数可写为

$$\begin{aligned}C_X(t_1, t_2) &= E[(X(t_1) - \mu_X)(X(t_2) - \mu_X)] \\ &= R_X(t_2 - t_1) - \mu_X^2\end{aligned}\quad (1.10)$$

式(1.10)表明,与自相关函数相似,严平稳随机过程 $X(t)$ 的自协方差函数也只与时间差 $t_2 - t_1$ 有关。这个公式还表明,如果确定了随机过程的均值和自相关函数,就能惟一确定地得到它的自协方差函数。因此,用均值和自相关函数已经能够描述平稳随机过程的一阶和二阶矩。

不过,还有两个要点必须注意:

1. 均值和自相关函数只给出了随机过程 $X(t)$ 分布的部分描述。
2. 仅满足式(1.7)和式(1.9),即相应的均值和自相关函数的条件,并不能证明随机过程 $X(t)$ 是严平稳的。

尽管如此,实际应用中通常还是仅考虑由均值和自相关函数给出的随机过程的部分描述。满足式(1.7)和式(1.9)的随机过程有许多名称,如二维平稳过程、宽平稳过程或弱平稳过程。本书约定简单地用平稳过程来代替这些名称。^[2]

一个平稳过程不一定是严平稳的。这是因为式(1.7)和式(1.9)显然并没有说明对于所有的 k ,式(1.3)中的 k 维联合分布函数关于 τ 是时不变的。另一方面,一个严平稳随机过程也不一定需要满足式(1.7)和式(1.9),这是因为它的一阶和二阶矩可能并不存在。不过,很显然有限二阶矩的严平稳随机过程组成的集合是所有平稳过程组成的集合的子集。

1.4.1 自相关函数的特性

为了简化符号,我们重新定义一个平稳随机过程 $X(t)$ 的自相关函数为

$$R_X(\tau) = E[X(t + \tau)X(t)] \quad \text{对于所有 } t \quad (1.11)$$

这个自相关函数有以下一些重要性质：

1. 随机过程的均方根值可以通过在式(1.11)中令 $\tau = 0$ 由 $R_X(\tau)$ 简单地得到,如下式:

$$R_X(0) = E[X^2(t)] \quad (1.12)$$

2. 自相关函数 $R_X(\tau)$ 是 τ 的偶函数,也就是说

$$R_X(\tau) = R_X(-\tau) \quad (1.13)$$

这一性质是直接从式(1.11)中导出的。相应地,也可以定义自相关函数 $R_X(\tau)$ 为

$$R_X(\tau) = E[X(t)X(t - \tau)]$$

3. 自相关函数在 $\tau = 0$ 时有最大值,即

$$|R_X(\tau)| \leq R_X(0) \quad (1.14)$$

为了证明这一性质,考虑非负量

$$E[(X(t + \tau) \pm X(t))^2] \geq 0$$

展开上式并取各自的数学期望,得

$$E[X^2(t + \tau)] \pm 2E[X(t + \tau)X(t)] + E[X^2(t)] \geq 0$$

结合式(1.11)和式(1.12),可将上式化简为

$$2R_X(0) \pm 2R_X(\tau) \geq 0$$

或等价地写为

$$-R_X(0) \leq R_X(\tau) \leq R_X(0)$$

从而可以直接导出式(1.14)。

自相关函数 $R_X(\tau)$ 的重要物理意义在于,它给出了描述间隔 τ 秒观测随机过程 $X(t)$ 所得到的两个随机变量相互依赖程度的一种方法。如图 1.4 所示,显然随机过程 $X(t)$ 随时间变化越快,随着 τ 的增加,自相关函数 $R_X(\tau)$ 从最大值 $R_X(0)$ 减小得越快。这种减小可以用去相关时间 τ_0 ,来表示。当 $\tau > \tau_0$ 时,自相关函数 $R_X(\tau)$ 的值将低于一个指定的值。由此可以定义零均值平稳随机过程 $X(t)$ 的去相关时间 τ_0 为自相关函数在 τ_0 的值减小为其最大值 $R_X(0)$ 的 1% 时的时间值。

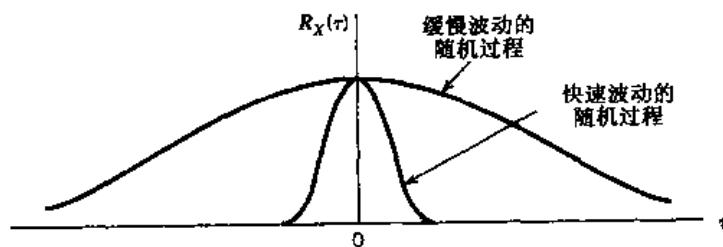


图 1.4 对快速和缓慢波动随机过程自相关函数的描述

例 1.2 随机相位正弦信号

设有一个随机相位正弦信号,由下式定义:

$$X(t) = A \cos(2\pi f_c t + \Theta) \quad (1.15)$$

这里 A 和 f_c 均为常数, Θ 为一个在 $[-\pi, \pi]$ 上均匀分布的随机变量, 即

$$f_\Theta(\theta) = \begin{cases} \frac{1}{2\pi}, & -\pi \leq \theta \leq \pi \\ 0, & \text{其他} \end{cases} \quad (1.16)$$

这意味着随机变量 Θ 等概地取区间 $[-\pi, \pi]$ 内的任意值, 每个 Θ 值对应于随机过程 $X(t)$ 的样本空间中的一个样本。

式(1.15)和式(1.16)定义的随机过程可以代表通信系统接收设备中随机生成的一个本地载波, 这个载波往往被用来对所接收信号进行解调。而随机变量 Θ 表示本地产生的正弦载波和发射端用于调制信号的正弦载波间的相位差。 $X(t)$ 的自相关函数为

$$\begin{aligned} R_X(\tau) &= E[X(t + \tau)X(t)] \\ &= E[A^2 \cos(2\pi f_c t + 2\pi f_c \tau + \Theta) \cos(2\pi f_c t + \Theta)] \\ &= \frac{A^2}{2} E[\cos(4\pi f_c t + 2\pi f_c \tau + 2\Theta)] + \frac{A^2}{2} E[\cos(2\pi f_c \tau)] \\ &= \frac{A^2}{2} \int_{-\pi}^{\pi} \frac{1}{2\pi} \cos(4\pi f_c t + 2\pi f_c \tau + 2\theta) d\theta + \frac{A^2}{2} \cos(2\pi f_c \tau) \end{aligned}$$

第一项积分为 0, 由此可得

$$R_X(\tau) = \frac{A^2}{2} \cos(2\pi f_c \tau) \quad (1.17)$$

对应的曲线如图 1.5 所示, 由此可见, 一个随机相位正弦信号是 τ 域而非原始时间域内的另一个同频的正弦信号。

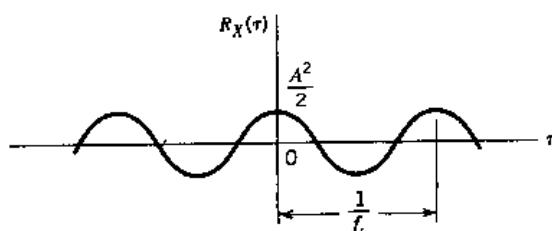


图 1.5 随机相位正弦信号的自相关函数

例 1.3 随机二进制信号

图 1.6 给出了包含一个随机二进制 0 和 1 序列的随机过程 $X(t)$ 的一个样本函数 $x(t)$, 并进行了如下假设:

1. 符号 1 和 0 分别用电压幅度为 $+A$ 和 $-A$ 的脉冲来表示, 分别持续 T 秒。
2. 脉冲不是同步的, 所以在正时间轴上, 第一个完整脉冲的开始时间 t_d 可以是 $0 \sim T$ 秒内的任一时刻。也就是说, t_d 是一个均匀分布的随机变量 T_d 的样本值, 其概率密度函数定义如下:

$$f_{T_d}(t_d) = \begin{cases} \frac{1}{T}, & 0 \leq t_d \leq T \\ 0, & \text{其他} \end{cases}$$

3. 在任一时间区间 $(n-1)T < t - t_d < nT$ (n 为一个整数) 内, 出现一个 1 还是一个 0 由投一枚均匀硬币决定, 如果结果是正面, 取 1; 如果是反面, 则取 0。因此这两个记号是等概出现的, 在任一时间区间内出现 1 或 0 与其他时间区间内的情况是相互独立的。

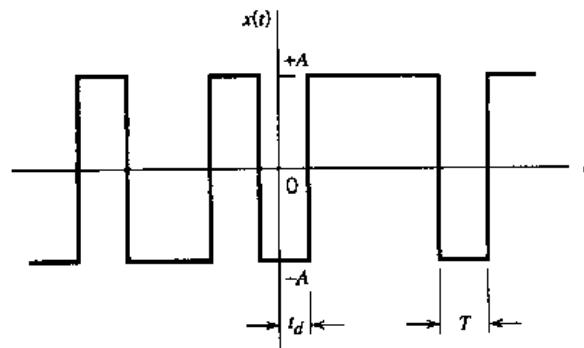


图 1.6 随机二进制波形的样本函数

由于幅度电平 $+A$ 和 $-A$ 等概出现, 可以得出, 对所有的 t 时刻有 $E[X(t)] = 0$, 即随机过程的均值为 0。

为了得出自相关函数 $R_X(t_k, t_i)$, 必须求出 $E[X(t_k)X(t_i)]$ 。此处 $X(t_k)$ 和 $X(t_i)$ 分别为在时刻 t_k 和 t_i 观测随机过程 $X(t)$ 所得的随机变量。

考虑第一种情况: $|t_k - t_i| > T$ 。在这一条件下, 随机变量 $X(t_k)$ 和 $X(t_i)$ 在不同的脉冲间隔内, 所以是相互独立的, 因此有

$$E[X(t_k)X(t_i)] = E[X(t_k)]E[X(t_i)] = 0, \quad |t_k - t_i| > T$$

考虑第二种情况: $|t_k - t_i| < T$, $t_k = 0$, $t_i < t_k$ 。在这一情况下, 由图 1.6 可知, 当且仅当 $t_d < T - |t_k - t_i|$ 时, 随机变量 $X(t_k)$ 和 $X(t_i)$ 在同一脉冲间隔内出现。由此可得条件期望

$$E[X(t_k)X(t_i)|t_d] = \begin{cases} A^2, & t_d < T - |t_k - t_i| \\ 0, & \text{其他} \end{cases}$$

在所有可能的 t_d 值上对这个结果取平均, 得

$$\begin{aligned} E[X(t_k)X(t_i)] &= \int_0^{T-|t_k-t_i|} A^2 f_{T_d}(t_d) dt_d \\ &= \int_0^{T-|t_k-t_i|} \frac{A^2}{T} dt_d \\ &= A^2 \left(1 - \frac{|t_k - t_i|}{T} \right), \quad |t_k - t_i| < T \end{aligned}$$

同理, 对于其他的 t_k 值, 也可以得出相同的结论: 具有图 1.6 所示样本函数的随机二进制波形的自相关函数仅仅是一个时间差 $\tau = t_k - t_i$ 的函数, 如下式所示:

$$R_X(\tau) = \begin{cases} A^2 \left(1 - \frac{|\tau|}{T} \right), & |\tau| < T \\ 0, & |\tau| \geq T \end{cases} \quad (1.18)$$

这一结果在图 1.7 中绘出。

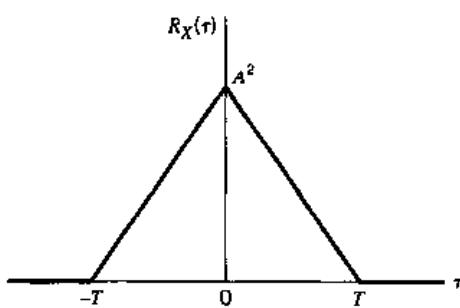


图 1.7 随机二进制波形的自相关函数

1.4.2 互相关函数

考虑一个更普遍的情况,两个随机过程 $X(t)$ 和 $Y(t)$, 自相关函数分别为 $R_X(t, u)$ 和 $R_Y(t, u)$, 则 $X(t)$ 和 $Y(t)$ 的两个互相关函数可以被定义为

$$R_{XY}(t, u) = E[X(t)Y(u)] \quad (1.19)$$

和

$$R_{YX}(t, u) = E[Y(t)X(u)] \quad (1.20)$$

这里 t 和 u 为随机过程的两个观测时间值。在这种情况下,两个随机过程 $X(t)$ 和 $Y(t)$ 的相关性可以更为方便地用矩阵形式表示如下:

$$\mathbf{R}(t, u) = \begin{bmatrix} R_X(t, u) & R_{XY}(t, u) \\ R_{YX}(t, u) & R_Y(t, u) \end{bmatrix}$$

这个矩阵称为随机过程 $X(t)$ 和 $Y(t)$ 的相关矩阵。如果 $X(t)$ 和 $Y(t)$ 都是平稳的,并且它们是联合平稳的,则自相关矩阵可以写为

$$\mathbf{R}(\tau) = \begin{bmatrix} R_X(\tau) & R_{XY}(\tau) \\ R_{YX}(\tau) & R_Y(\tau) \end{bmatrix} \quad (1.21)$$

其中 $\tau = t - u$ 。

与自相关函数不同,互相关函数往往不是 τ 的偶函数,初值不一定为最大值。但确实满足如下的对称关系(见例 1.9):

$$R_{XY}(\tau) = R_{YX}(-\tau) \quad (1.22)$$

例 1.4 正交调制过程

设有与平稳随机过程 $X(t)$ 存在如下关系的一对正交调制过程 $X_1(t)$ 和 $X_2(t)$:

$$\begin{aligned} X_1(t) &= X(t) \cos(2\pi f_c t + \Theta) \\ X_2(t) &= X(t) \sin(2\pi f_c t + \Theta) \end{aligned}$$

这里 f_c 是载波频率,随机变量 Θ 在区间 $[0, 2\pi]$ 上均匀分布。此外, Θ 是独立于 $X(t)$ 的。

$X_1(t)$ 和 $X_2(t)$ 的互相关函数由下式得到:

$$\begin{aligned}
 R_{12}(\tau) &= E[X_1(t)X_2(t - \tau)] \\
 &= E[X(t)X(t - \tau) \cos(2\pi f_c t + \Theta) \sin(2\pi f_c t - 2\pi f_c \tau + \Theta)] \\
 &= E[X(t)X(t - \tau)]E[\cos(2\pi f_c t + \Theta) \sin(2\pi f_c t - 2\pi f_c \tau + \Theta)] \\
 &= \frac{1}{2}R_X(\tau)[\sin(4\pi f_c t - 2\pi f_c \tau + 2\Theta) - \sin(2\pi f_c \tau)] \\
 &= -\frac{1}{2}R_X(\tau) \sin(2\pi f_c \tau)
 \end{aligned} \tag{1.23}$$

最后一行利用了表示相位的随机变量 Θ 的均匀分布。注意到 $\tau = 0$ 时, 因子 $\sin(2\pi f_c \tau)$ 为 0, 因此

$$\begin{aligned}
 R_{12}(0) &= E[X_1(t)X_2(t)] \\
 &= 0
 \end{aligned}$$

这表明, 在某一固定的时刻 t , 同时观测正交调制过程 $X_1(t)$ 和 $X_2(t)$, 所得的随机变量是相互正交的。

1.5 遍历过程

随机过程 $X(t)$ 的数学期望或集平均是“整体平均”。例如, 随机过程 $X(t)$ 在某一固定时刻 t_k 的均值, 也就是描述在时刻 $t = t_k$ 时过程所有可能值的随机变量 $X(t_k)$ 的数学期望。当然, 也可以定义长期样本平均或时间平均为整个过程上的平均。因此, 考虑用时间平均代替集平均。这意味着可能找到一种可行的方法来估计随机过程的集平均。当然, 重要的是, 什么时候可以用时间平均来代替集平均? 为了研究这个问题, 考虑一个观测区间定义在 $-T \leq t \leq T$ 内的平稳过程 $X(t)$ 的样本函数 $x(t)$ 。 $x(t)$ 的直流成分定义为时间平均:

$$\mu_x(T) = \frac{1}{2T} \int_{-T}^T x(t) dt \tag{1.24}$$

显然时间平均 $\mu_x(T)$ 是一个随机变量, 它的值由观测时间区间和在式(1.24)中采用的随机过程 $X(t)$ 所选定的样本函数决定。既然 $X(t)$ 被假定为平稳的, 那么时间平均 $\mu_x(T)$ 的均值由下式给出(已经交换了求积分和数学期望的顺序):

$$\begin{aligned}
 E[\mu_x(T)] &= \frac{1}{2T} \int_{-T}^T E[x(t)] dt \\
 &= \frac{1}{2T} \int_{-T}^T \mu_X dt \\
 &= \mu_X
 \end{aligned} \tag{1.25}$$

这里 μ_X 是整个过程 $X(t)$ 的均值。因此, 时间平均 $\mu_x(T)$ 代表集平均 μ_X 的一个无偏估计。如果满足下面两个条件, 则称 $X(t)$ 是均值遍历的:

- 时间平均 $\mu_x(T)$ 在观察间隔 T 趋于无穷时趋于集平均 μ_X , 即

$$\lim_{T \rightarrow \infty} \mu_x(T) = \mu_X$$

- 将 $\mu_x(T)$ 的方差作为随机变量来处理时, 当时间间隔 T 趋于无穷时其值趋于 0。也就是说

$$\lim_{T \rightarrow \infty} \text{var}[\mu_x(T)] = 0$$

此外,自相关函数 $R_x(\tau, T)$ 是在区间 $-T \leq t \leq T$ 内观测的样本函数 $x(t)$ 的时间平均。结合式(1.24),可以正式地定义样本函数 $x(t)$ 的时间平均自相关函数为

$$R_x(\tau, T) = \frac{1}{2T} \int_{-T}^T x(t + \tau)x(t) dt \quad (1.26)$$

这第二个时间平均也可以视为一个具有相同均值和方差的随机变量。与均值的遍历性同理,若满足下面两个极限条件,则称一个过程 $x(t)$ 是自相关函数遍历的:

$$\begin{aligned} \lim_{T \rightarrow \infty} R_x(\tau, T) &= R_x(\tau) \\ \lim_{T \rightarrow \infty} \text{var}[R_x(\tau, T)] &= 0 \end{aligned}$$

当然,也可以继续用相同的方法来定义过程 $X(t)$ 的其他高阶统计量的遍历性。但在通常情况下,这里所说的自相关函数和均值的遍历性已经足够了。同时需注意的是,在利用式(1.24)和式(1.26)计算时间平均 $\mu_x(T)$ 和 $R_x(\tau, T)$ 时,必须满足随机过程 $X(t)$ 是平稳的这一条件。总之,一个具有遍历性的随机过程一定是平稳的,但反之不一定成立。

1.6 随机过程通过一个线性时不变滤波器

假定如图 1.8 所示,随机过程 $X(t)$ 是一个冲激响应为 $b(t)$ 的线性时不变滤波器的输入,在滤波器的输出端产生一个新的随机过程 $Y(t)$ 。通常,即使输入随机过程在 $-\infty < t < \infty$ 上的概率分布函数已经完全给定,也很难对输出随机过程的概率分布函数进行描述。

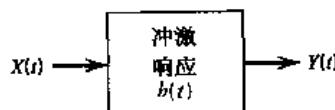


图 1.8 随机过程通过线性时不变滤波器的传输

本节假定 $X(t)$ 是一个平稳过程,通过滤波器的输入-输出关系的时域形式,可以用输入 $X(t)$ 均值和自相关函数来对输出 $Y(t)$ 的均值和自相关函数进行定义。

随机过程通过线性时不变滤波器后的传输是由卷积积分决定的,运算过程见附录 B。输出 $Y(t)$ 与输入 $X(t)$ 的关系为

$$Y(t) = \int_{-\infty}^{\infty} b(\tau_1)X(t - \tau_1) d\tau_1$$

其中 τ_1 为积分变量。因此 $Y(t)$ 的均值为

$$\begin{aligned} \mu_Y(t) &= E[Y(t)] \\ &= E\left[\int_{-\infty}^{\infty} b(\tau_1)X(t - \tau_1) d\tau_1\right] \end{aligned} \quad (1.27)$$

如果给定的数学期望 $E[X(t)]$ 对所有的 t 是有限的,并且系统是稳定的,则式(1.27)中积分和求期望的顺序可以交换,从而变为

$$\begin{aligned}\mu_Y(t) &= \int_{-\infty}^{\infty} b(\tau_1) E[X(t - \tau_1)] d\tau_1 \\ &= \int_{-\infty}^{\infty} b(\tau_1) \mu_X(t - \tau_1) d\tau_1\end{aligned}\quad (1.28)$$

当输入 $X(t)$ 是平稳过程, 均值 $\mu_X(t)$ 是常量 μ_X , 可将式(1.28)简化如下:

$$\begin{aligned}\mu_Y &= \mu_X \int_{-\infty}^{\infty} b(\tau_1) d\tau_1 \\ &= \mu_X H(0)\end{aligned}\quad (1.29)$$

这里 $H(0)$ 是系统的一个零频率(直流)响应。式(1.29)表明, 作为输入 $X(t)$ 的响应, 线性时不变滤波器的输出 $Y(t)$ 的均值等于 $X(t)$ 的均值乘以系统直流响应。

下面考虑输出 $Y(t)$ 的自相关函数。由定义, 有

$$R_Y(t, u) = E[Y(t)Y(u)]$$

这里 t 和 u 为观测输出过程的两个时刻, 因此可以利用卷积积分写为

$$R_Y(t, u) = E \left[\int_{-\infty}^{\infty} b(\tau_1) X(t - \tau_1) d\tau_1 \int_{-\infty}^{\infty} b(\tau_2) X(u - \tau_2) d\tau_2 \right] \quad (1.30)$$

这里仍假设均方值 $E[X^2(t)]$ 对所有的 t 是有限的, 并且系统是稳定的, 交换式(1.30)中 τ_1 与 τ_2 求积分和数学期望的顺序, 得到

$$\begin{aligned}R_Y(t, u) &= \int_{-\infty}^{\infty} d\tau_1 b(\tau_1) \int_{-\infty}^{\infty} d\tau_2 b(\tau_2) E[X(t - \tau_1)X(u - \tau_2)] \\ &= \int_{-\infty}^{\infty} d\tau_1 b(\tau_1) \int_{-\infty}^{\infty} d\tau_2 b(\tau_2) R_X(t - \tau_1, u - \tau_2)\end{aligned}\quad (1.31)$$

当输入 $X(t)$ 为平稳过程时, 自相关函数是观测时间差 $t - \tau_1$ 和 $u - \tau_2$ 的函数。在式(1.31)中令 $\tau = t - u$, 从而变为

$$R_Y(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\tau_1) b(\tau_2) R_X(\tau - \tau_1 + \tau_2) d\tau_1 d\tau_2 \quad (1.32)$$

将这个结论与前面所得到的关于均值 μ_Y 的结果相结合, 可以得出结论, 如果一个稳定的线性时不变滤波器的输入是一个平稳随机过程, 那么它的输出也是一个平稳随机过程。

在式(1.32)中令 $\tau = 0$, 可得到输出 $Y(t)$ 的均方值。结合 $R_Y(0) = E[Y^2(t)]$ 可得

$$E[Y^2(t)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b(\tau_1) b(\tau_2) R_X(\tau_2 - \tau_1) d\tau_1 d\tau_2 \quad (1.33)$$

结果为一个常数。

1.7 功率谱密度

到此为止, 我们已经讨论了平稳过程在线性系统中的时域特性。下面将利用频域的思想来讨论线性系统中的随机过程的特性。首先, 希望从式(1.33)的结果中获得频域的等价形式, 以便对滤波器输出的均方值进行定义。

由定义,一个线性时不变滤波器的冲激响应等于这个系统频率响应的傅里叶反变换,傅里叶反变换的内容见附录2。以 $H(f)$ 表示一个系统的频率响应,则冲激响应

$$b(\tau_1) = \int_{-\infty}^{\infty} H(f) \exp(j2\pi f \tau_1) df \quad (1.34)$$

将此式代入式(1.33),得

$$\begin{aligned} E[Y^2(t)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} H(f) \exp(j2\pi f \tau_1) df \right] b(\tau_2) R_X(\tau_2 - \tau_1) d\tau_1 d\tau_2 \\ &= \int_{-\infty}^{\infty} df H(f) \int_{-\infty}^{\infty} d\tau_2 b(\tau_2) \int_{-\infty}^{\infty} R_X(\tau_2 - \tau_1) \exp(j2\pi f \tau_1) d\tau_1 \end{aligned} \quad (1.35)$$

在式(1.35)右边最后一个积分项中,定义一个新的变量

$$\tau = \tau_2 - \tau_1$$

则式(1.35)可写为

$$E[Y^2(t)] = \int_{-\infty}^{\infty} df H(f) \int_{-\infty}^{\infty} d\tau_2 b(\tau_2) \exp(j2\pi f \tau_2) \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f \tau) d\tau \quad (1.36)$$

将式(1.36)中右边中间的积分简写为 $H^*(f)$,即滤波器频率响应的复共轭,于是得到

$$E[Y^2(t)] = \int_{-\infty}^{\infty} df |H(f)|^2 \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f \tau) d\tau \quad (1.37)$$

其中 $|H(f)|$ 是滤波器的幅度响应。最右边的积分是输入 $X(t)$ 的自相关函数 $R_X(\tau)$ 的傅里叶变换。为进一步简化上式,我们引入一个新的参数,其定义为

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f \tau) d\tau \quad (1.38)$$

函数 $S_X(f)$ 被称为平稳随机过程 $X(t)$ 的功率谱密度或功率谱。将式(1.38)代入式(1.37),得

$$E[Y^2(t)] = \int_{-\infty}^{\infty} |H(f)|^2 S_X(f) df \quad (1.39)$$

式(1.39)表明,一个稳定的线性时不变系统对一个平稳过程的响应输出 $Y(t)$ 的均方值,等于输入 $X(t)$ 的功率谱密度与滤波器幅度响应平方的乘积在整个频域上的积分。这就是所求的式(1.33)的时域关系的频域等价式。

为了研究功率谱密度的物理意义,假定随机过程 $X(t)$ 通过一个以 f_c 为中心的理想窄带滤波器,其幅度响应如图 1.9 所示,即

$$|H(f)| = \begin{cases} 1, & |f \pm f_c| < \frac{1}{2}\Delta f \\ 0, & |f \pm f_c| > \frac{1}{2}\Delta f \end{cases} \quad (1.40)$$

这里 Δf 是滤波器的带宽,由式(1.39)可见,如果滤波器的带宽 Δf 相对于中心频率 f_c 足够小,且 $S_X(f)$ 为一连续函数,则滤波器输出 $Y(t)$ 的均方值近似为

$$E[Y^2(t)] \approx (2\Delta f) S_X(f_c) \quad (1.41)$$

滤波器只允许输入 $X(t)$ 中那些在以 $\pm f_c$ 为中频且频宽为 Δf 的窄频带内的频率分量通过。于是,在 $f = f_c$ 时求得的 $S_X(f_c)$ 就代表 $X(t)$ 的平均功率密度。因此功率谱密度的单位是 W/Hz。

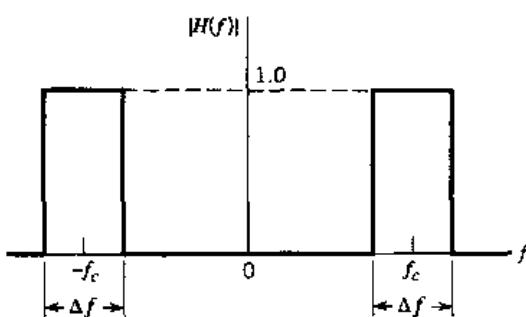


图 1.9 理想窄带滤波器的幅度响应

1.7.1 功率谱密度的特性

平稳过程 $X(t)$ 的功率谱密度 $S_X(f)$ 和自相关函数 $R_X(\tau)$ 组成一个以 τ 和 f 为自变量的傅里叶变换对, 即下面这对关系式:

$$S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f\tau) d\tau \quad (1.42)$$

$$R_X(\tau) = \int_{-\infty}^{\infty} S_X(f) \exp(j2\pi f\tau) df \quad (1.43)$$

式(1.42)和式(1.43)是随机过程谱分析理论的基本关系式, 它们一起构成的关系式通常称为 Einstein-Wiener-Khintchine 关系式。^[3]

Einstein-Wiener-Khintchine 关系式表明, 如果随机过程的自相关函数和功率谱密度中有一个已知, 那么另一个就能立即准确求出。这两个函数表示的是随机过程相关信息的不同方面。通常认为, 对于实际应用来说, 功率谱密度是更有用的一个参数。

下面利用这对关系式来得到平稳随机过程功率谱密度的几点基本性质。

性质 1

平稳过程功率谱密度的零频率值等于自相关函数曲线下的全部面积, 也就是

$$S_X(0) = \int_{-\infty}^{\infty} R_X(\tau) d\tau \quad (1.44)$$

在式(1.42)中令 $f = 0$ 即可得到这一性质。

性质 2

平稳过程的均方值等于功率谱密度曲线下的全部面积, 即

$$E[X^2(t)] = \int_{-\infty}^{\infty} S_X(f) df \quad (1.45)$$

在式(1.43)中令 $\tau = 0$, $R_X(0) = E[X^2(t)]$ 可直接导出这个性质。

性质 3

平稳过程的功率谱密度非负, 也就是说

$$S_X(f) \geq 0 \quad \text{对于所有 } f \quad (1.46)$$

这个性质是式(1.41)中, 均方值 $E[Y^2(t)]$ 必须非负的直接结论。

性质 4

一个实随机过程的功率谱密度是频率的偶函数。也就是说

$$S_X(-f) = S_X(f) \quad (1.47)$$

这个性质可以通过在式(1.42)中用 $-f$ 代替 f 得到

$$S_X(-f) = \int_{-\infty}^{\infty} R_X(\tau) \exp(j2\pi f\tau) d\tau$$

然后,用 $-\tau$ 代替 τ ,结合 $R_X(-\tau) = R_X(\tau)$,得到

$$S_X(-f) = \int_{-\infty}^{\infty} R_X(\tau) \exp(-j2\pi f\tau) d\tau = S_X(f)$$

这就是所求的结果。

性质 5

功率谱密度进行适当标准化后,往往具有概率密度函数的性质。

此处所用的标准化方法与功率谱密度曲线下的全部面积(也就是过程的均方值)有关。考虑函数

$$p_X(f) = \frac{S_X(f)}{\int_{-\infty}^{\infty} S_X(f) df} \quad (1.48)$$

根据性质 2 和 3,对所有 f 而言,有 $p_X(f) \geq 0$ 。此外,在 $p_X(f)$ 函数下全部面积为 1。因此式(1.48)所定义的标准化功率谱密度与概率密度函数有相似的性质。

例 1.5 随机相位正弦信号(续)

设随机过程 $X(t) = A \cos(2\pi f_c t + \Theta)$, 这里 Θ 是一个在区间 $[-\pi, \pi]$ 上均匀分布的随机变量, $X(t)$ 的自相关函数由式(1.17)给出,即

$$R_X(\tau) = \frac{A^2}{2} \cos(2\pi f_c \tau)$$

令 $\delta(f)$ 为 $f=0$ 时的 delta 函数,delta 函数的定义和性质见附录 2。对自相关函数 $R_X(\tau)$ 的关系式的两边同时进行傅里叶变换,得到正弦过程 $X(t)$ 的功率谱密度:

$$S_X(f) = \frac{A^2}{4} [\delta(f - f_c) + \delta(f + f_c)] \quad (1.49)$$

如图 1.10 所示,这个式子包含 $\pm f_c$ 处的加权值为 $A^2/4$ 的一对 delta 函数。定义 delta 函数下的全部面积为 1,因此 $S_X(f)$ 下所有面积为 $A^2/2$ 。

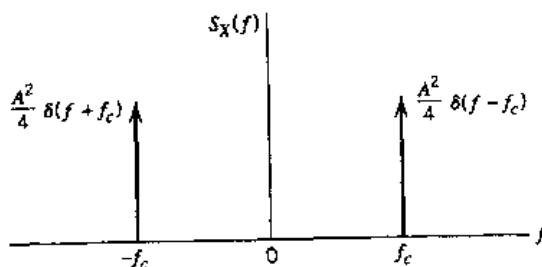


图 1.10 随机相位正弦信号的功率谱密度。 $\delta(f)$ 为 $f=0$ 时对应的 delta 函数

例 1.6 随机二进制信号(续)

再来看分别由 $\pm A$ 表示的 0 和 1 序列组成的随机二进制信号, 由例 1.3 可知, 这一随机过程的自相关函数具有三角形的波形, 可用下式表示:

$$R_X(\tau) = \begin{cases} A^2 \left(1 - \frac{|\tau|}{T}\right), & |\tau| < T \\ 0, & |\tau| \geq T \end{cases}$$

因此, 随机过程的功率谱密度为

$$S_X(f) = \int_{-T}^T A^2 \left(1 - \frac{|\tau|}{T}\right) \exp(-j2\pi f\tau) d\tau$$

利用三角形函数的傅里叶变换(见附录 6 表 A6.3)可得

$$S_X(f) = A^2 T \operatorname{sinc}^2(fT) \quad (1.50)$$

其图形在图 1.11 中给出。图中再次显示了功率谱密度对所有的 f 是非负的, 并且是 f 的偶函数。注意到 $R_X(0) = A^2$, 结合性质 2, 可知 $S_X(f)$ 曲线下的全部面积或这里所说的随机二进制信号的平均功率为 A^2 , 这与直观是相符的。

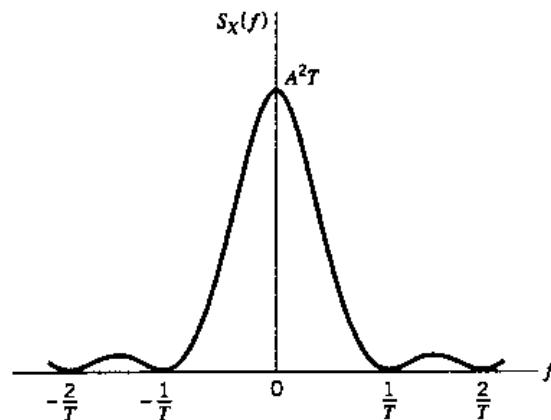


图 1.11 随机二进制信号的功率谱密度

式(1.50)的结果可以进行如下的推广。定义一个幅度为 A , 持续时间为 T 的矩形脉冲的能量谱密度(即傅里叶变换的幅度平方)为

$$\mathcal{E}_g(f) = A^2 T^2 \operatorname{sinc}^2(fT) \quad (1.51)$$

根据 $\mathcal{E}_g(f)$ 对式(1.50)进行简化得

$$S_X(f) = \frac{\mathcal{E}_g(f)}{T} \quad (1.52)$$

式(1.52)表明, 对于一个用 $g(t)$ 和 $-g(t)$ 分别代替前面所说的符号 1 和 0 的随机二进制信号, 其功率谱密度 $S_X(f)$ 等于波形为 $\mathcal{E}_g(f)$ 的符号的能量谱密度与其持续时间 T 的商。

例 1.7 随机过程与正弦过程的混合

实际情况中经常会遇到一个平稳过程 $X(t)$ 和一个正弦信号 $\cos(2\pi f_c t + \Theta)$ 混合(即相乘)的情况, 这里相位 Θ 是在区间 $[0, 2\pi]$ 上均匀分布的随机变量。在这个问题中, 随机相位

的引入仅仅是为了表明当 $X(t)$ 和 $\cos(2\pi f_c t + \Theta)$ 产生于物理上相互独立的来源时, 起始时间是可以任意选定的。事实上也往往如此。为了得到随机过程 $Y(t)$ 的功率谱密度, 定义 $Y(t)$ 如下:

$$Y(t) = X(t) \cos(2\pi f_c t + \Theta) \quad (1.53)$$

利用平稳过程自相关函数的定义, 并由随机变量 Θ 独立于 $X(t)$ 的性质, 得出 $Y(t)$ 的自相关函数为

$$\begin{aligned} R_Y(\tau) &= E[Y(t + \tau)Y(t)] \\ &= E[X(t + \tau) \cos(2\pi f_c t + 2\pi f_c \tau + \Theta)X(t) \cos(2\pi f_c t + \Theta)] \\ &= E[X(t + \tau)X(t)]E[\cos(2\pi f_c t + 2\pi f_c \tau + \Theta) \cos(2\pi f_c t + \Theta)] \\ &= \frac{1}{2}R_X(\tau)E[\cos(2\pi f_c \tau) + \cos(4\pi f_c t + 2\pi f_c \tau + 2\Theta)] \\ &= \frac{1}{2}R_X(\tau) \cos(2\pi f_c \tau) \end{aligned} \quad (1.54)$$

由于功率谱密度是自相关函数的傅里叶变换, 因此可得随机过程 $X(t)$ 与 $Y(t)$ 的功率谱密度的关系如下:

$$S_Y(f) = \frac{1}{4}[S_X(f - f_c) + S_X(f + f_c)] \quad (1.55)$$

根据式(1.55), 在式(1.53)中定义的随机过程 $Y(t)$ 的功率谱密度可以由以下方法得到: 将随机过程 $X(t)$ 的功率谱密度 $S_X(f)$ 分别向右和向左平移 f_c , 将两次所得结果相加, 再除以 4。

1.7.2 输入和输出随机过程功率谱密度的关系

令 $S_Y(f)$ 为随机过程 $X(t)$ 通过频率响应为 $H(f)$ 的线性滤波器后得到的输出 $Y(t)$ 的功率谱密度。那么由随机过程的功率谱密度等于其自相关函数的傅里叶变换的定义, 结合式(1.32)得到

$$\begin{aligned} S_Y(f) &= \int_{-\infty}^{\infty} R_Y(\tau) \exp(-j2\pi f\tau) d\tau \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\tau_1)h(\tau_2)R_X(\tau - \tau_1 + \tau_2) \exp(-j2\pi f\tau) d\tau_1 d\tau_2 d\tau \end{aligned} \quad (1.56)$$

令 $\tau - \tau_1 + \tau_2 = \tau_0$ 或等价地令 $\tau_0 + \tau_1 - \tau_2 = \tau$ 并代入式(1.56), 则 $S_Y(f)$ 可以写成三项的乘积: 滤波器的频率响应 $H(f)$, $H(f)$ 的复共轭, 输入随机过程 $X(t)$ 的功率谱密度 $S_X(f)$ 。由此可将式(1.56)简化为

$$S_Y(f) = H(f)H^*(f)S_X(f) \quad (1.57)$$

最后, 结合 $|H(f)|^2 = H(f)H^*(f)$, 可将频域内输入输出的功率谱密度关系写为

$$S_Y(f) = |H(f)|^2 S_X(f) \quad (1.58)$$

式(1.58)表明, 输出 $Y(t)$ 的功率谱密度等于输入 $X(t)$ 的功率密度与滤波器幅度响应平方的乘积。利用这个关系可得到随机过程通过一个稳定线性时不变滤波器的输出结果。从计算方面考虑, 式(1.58)往往比采用自相关函数的时域对应式(1.32)更易于处理。

1.7.3 样本函数的功率谱密度与幅度谱的关系

下面将找出平稳遍历的随机过程 $X(t)$ 的一个样本函数 $x(t)$ 的频谱特性与功率谱密度

$S_x(f)$ 的直接关系。一个可以进行傅里叶变换的样本函数 $x(t)$ 必须是完全可积的,也就是说

$$\int_{-\infty}^{\infty} |x(t)| dt < \infty \quad (1.59)$$

无限持续时间内的任何一个平稳样本函数 $x(t)$ 都不满足这一条件,为了采用傅里叶变换的分析方法,考虑定义在观测区间 $-T \leq t \leq T$ 上的 $x(t)$ 的一个截短部分。用 $X(f, T)$ 代表该截短样本函数的傅里叶变换,则

$$X(f, T) = \int_{-T}^{T} x(t) \exp(-j2\pi ft) dt \quad (1.60)$$

假定 $x(t)$ 也是遍历的,可以利用时间平均方程(见 1.5 节)求出 $X(f, T)$ 的自相关函数 $R_X(\tau)$:

$$R_X(\tau) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^{T} x(t + \tau) x(t) dt \quad (1.61)$$

习惯上认为样本函数 $x(t)$ 是一个功率信号(即功率有限的信号),因此,可以公式化如下的傅里叶变换对:

$$\frac{1}{2T} \int_{-T}^{T} x(t + \tau) x(t) dt \rightleftharpoons \frac{1}{2T} |X(f, T)|^2 \quad (1.62)$$

上式左边部分是自相关函数的时间平均;右边的部分称为周期图,它的量纲与功率谱相同。但这个术语用词并不确切,因为周期图是频率而不是周期的函数。尽管如此,它仍然得到了广泛的应用。这个量最初的应用是被统计学家用来确定周期性,例如确定数据周期性的趋势。

利用式(1.62)所示的傅里叶变换对中的反变换式,结合周期图将样本函数 $x(t)$ 时间平均的自相关函数表示如下:

$$\frac{1}{2T} \int_{-T}^{T} x(t + \tau) x(t) dt = \int_{-\infty}^{\infty} \frac{1}{2T} |X(f, T)|^2 \exp(j2\pi f T) df \quad (1.63)$$

由此,将式(1.63)代入式(1.61),得到

$$R_X(\tau) = \lim_{T \rightarrow \infty} \int_{-\infty}^{\infty} \frac{1}{2T} |X(f, T)|^2 \exp(j2\pi f \tau) df \quad (1.64)$$

对于一个固定的频率值,周期图是一个随机变量,因为它的值是随着随机过程不同样本函数的变化而随机变化的。因此,当 T 趋于无穷时,一个给定的样本函数 $x(t)$ 的周期图,在任何统计意义上都不会收敛于一个极限值。因此,不应当交换式(1.64)中积分和极限运算的顺序。不过,可以假定在式(1.64)两边取随机过程所有样本函数的总体平均,并且遍历性的自相关函数 $R_X(\tau)$ 在这一运算过程中保持不变。由于遍历过程的每个样本函数最终会经历其他样本函数的几乎所有状态,因此可以得到下式:

$$R_X(\tau) = \lim_{T \rightarrow \infty} \int_{-\infty}^{\infty} \frac{1}{2T} E[|X(f, T)|^2] \exp(j2\pi f \tau) df \quad (1.65)$$

现在,可以交换积分和极限运算的顺序并得到

$$R_X(\tau) = \int_{-\infty}^{\infty} \left\{ \lim_{T \rightarrow \infty} \frac{1}{2T} E[|X(f, T)|^2] \right\} \exp(j2\pi f \tau) df \quad (1.66)$$

比较式(1.66)和式(1.43)可得,所需的遍历过程的功率谱密度 $S_X(f)$ 与该过程截短样本函数幅度平方谱的关系为

$$\begin{aligned} S_X(f) &= \lim_{T \rightarrow \infty} \frac{1}{2T} E[|X(f, T)|^2] \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} E\left[\left|\int_{-T}^T x(t) \exp(-j2\pi f t) dt\right|^2\right] \end{aligned} \quad (1.67)$$

需要注意,在式(1.67)中不能先令 T 趋于无穷再取期望值,这一点十分重要。通过在 $[-T, T]$ 上观测随机过程的一个样本函数 $x(t)$,式(1.67)给出了估计^[4]遍历随机过程的功率谱密度的数学基础。

1.7.4 互功率谱密度

功率谱密度只给出了单个随机过程的频率分布的度量方法,而互功率谱密度则给出了两个随机过程在频域上相互关系的度量方法。令 $X(t)$ 和 $Y(t)$ 为两个联合平稳随机过程,它们的互相关函数为 $R_{XY}(\tau)$ 和 $R_{YX}(\tau)$ 。可将这对随机过程的互功率谱密度 $S_{XY}(f)$ 和 $S_{YX}(f)$ 定义为其相应的互相关函数的傅里叶变换,如下所示:

$$S_{XY}(f) = \int_{-\infty}^{\infty} R_{XY}(\tau) \exp(-j2\pi f \tau) d\tau \quad (1.68)$$

和

$$S_{YX}(f) = \int_{-\infty}^{\infty} R_{YX}(\tau) \exp(-j2\pi f \tau) d\tau \quad (1.69)$$

互相关函数和互功率谱密度构成一对傅里叶变换对。利用傅里叶反变换式,还可写为

$$R_{XY}(\tau) = \int_{-\infty}^{\infty} S_{XY}(f) \exp(j2\pi f \tau) df \quad (1.70)$$

和

$$R_{YX}(\tau) = \int_{-\infty}^{\infty} S_{YX}(f) \exp(j2\pi f \tau) df \quad (1.71)$$

互功率谱密度 $S_{XY}(f)$ 和 $S_{YX}(f)$ 并不需要是频率 f 的实函数。将关系式

$$R_{XY}(\tau) = R_{YX}(-\tau)$$

代入式(1.68)中,结合式(1.69),得到 $S_{XY}(f)$ 和 $S_{YX}(f)$ 的关系式为

$$S_{XY}(f) = S_{YX}(-f) = S_{YX}^*(f) \quad (1.72)$$

例 1.8

假定随机过程 $X(t)$ 和 $Y(t)$ 是零均值、独立、平稳的。假设和随机过程为

$$Z(t) = X(t) + Y(t)$$

试求 $Z(t)$ 的功率谱密度。

$Z(t)$ 的自相关函数为

$$\begin{aligned}
 R_Z(t, u) &= E[Z(t)Z(u)] \\
 &= E[(X(t) + Y(t))(X(u) + Y(u))] \\
 &= E[X(t)X(u)] + E[X(t)Y(u)] + E[Y(t)X(u)] + E[Y(t)Y(u)] \\
 &= R_X(t, u) + R_{XY}(t, u) + R_{YX}(t, u) + R_Y(t, u)
 \end{aligned}$$

定义 $\tau = t - u$, 则

$$R_Z(\tau) = R_X(\tau) + R_{XY}(\tau) + R_{YX}(\tau) + R_Y(\tau) \quad (1.73)$$

随机过程 $X(t)$ 和 $Y(t)$ 是联合平稳的, 在式(1.73)两边取傅里叶变换, 得

$$S_Z(f) = S_X(f) + S_{XY}(f) + S_{YX}(f) + S_Y(f) \quad (1.74)$$

由此看出, 随机过程之和的功率谱密度为互功率谱密度 $S_{XY}(f)$ 和 $S_{YX}(f)$ 与相关的随机过程各自功率谱密度之和。

当平稳过程 $X(t)$ 和 $Y(t)$ 不相关时, 互功率谱密度 $S_{XY}(f)$ 和 $S_{YX}(f)$ 等于 0, 所以式(1.74)简化为

$$S_Z(f) = S_X(f) + S_Y(f) \quad (1.75)$$

可以把上式的结果推广为, 当几个零均值平稳过程彼此不相关时, 它们和的功率谱密度等于各自功率谱密度之和。

例 1.9

两个联合平稳过程通过一对分开的稳定线性时不变滤波器, 如图 1.12 所示。假定随机过程 $X(t)$ 是冲激响应为 $b_1(t)$ 的滤波器的输入, 随机过程 $Y(t)$ 是冲激响应为 $b_2(t)$ 的滤波器的输入。令随机过程 $V(t)$ 和 $Z(t)$ 为相应滤波器的输出。则 $V(t)$ 和 $Z(t)$ 的互相关函数为

$$\begin{aligned}
 R_{VZ}(t, u) &= E[V(t)Z(u)] \\
 &= E\left[\int_{-\infty}^{\infty} b_1(\tau_1)X(t - \tau_1) d\tau_1 \int_{-\infty}^{\infty} b_2(\tau_2)Y(u - \tau_2) d\tau_2\right] \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b_1(\tau_1)b_2(\tau_2)E[X(t - \tau_1)Y(u - \tau_2)] d\tau_1 d\tau_2 \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b_1(\tau_1)b_2(\tau_2)R_{XY}(t - \tau_1, u - \tau_2) d\tau_1 d\tau_2
 \end{aligned} \quad (1.76)$$

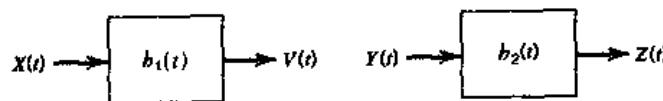


图 1.12 一对分开的线性时不变滤波器

这里 $R_{XY}(t, u)$ 为 $X(t)$ 和 $Y(t)$ 的互相关函数, 输入是联合平稳的(根据前面的假设), 令 $\tau = t - u$ 并将式(1.76)写为

$$R_{VZ}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b_1(\tau_1)b_2(\tau_2)R_{XY}(\tau - \tau_1 + \tau_2) d\tau_1 d\tau_2 \quad (1.77)$$

与式(1.39)同理, 在式(1.77)两边取傅里叶变换, 并进行推导, 最后得到

$$S_{VZ}(f) = H_1(f)H_2^*(f)S_{XY}(f) \quad (1.78)$$

这里 $H_1(f)$ 和 $H_2(f)$ 是图 1.12 中相应滤波器的频率响应, $H_2^*(f)$ 为 $H_2(f)$ 的复共轭。这就是所求的输出和输入过程功率谱密度的关系。

1.8 高斯过程

前面讨论的随机过程都是相当普遍的。在这一节,考虑随机过程家族中一类重要的过程:高斯过程。^[5]

假定观测一个 $t = 0$ 到 $t = T$ 区间内的随机过程 $X(t)$,用某一函数 $g(t)$ 来乘以随机过程 $X(t)$,并在这一观测区间上对乘积 $g(t)X(t)$ 进行积分,由此得到一个随机变量 Y :

$$Y = \int_0^T g(t)X(t) dt \quad (1.79)$$

Y 称为 $X(t)$ 的一个线性泛函。函数和泛函的区别应当引起注意。例如和式 $Y = \sum_{i=1}^N a_i X_i$ (a_i 为常数, X_i 为随机变量) 中 Y 就是 X_i 的线性函数。这是因为,对于随机变量 X_i 的每组观测值,都有一个随机变量 Y 与之对应。另一方面,在式(1.79)中,随机变量 Y 的值与从 0 到 T 的整个观测区间内的参数函数 $g(t)X(t)$ 有关。因此,一个泛函是与一个或几个函数的整个过程有关的量,而不仅仅是与一些离散变量有关。换言之,泛函域是一组满足要求的函数而不是一个对等空间的区域。

在式(1.79)中,如果加权函数 $g(t)$ 使得随机变量 Y 的均方值为有限,并且对于每一个 $g(t)$,随机变量 Y 都是高斯分布的,那么称 $X(t)$ 为高斯过程。换言之,如果 $X(t)$ 的所有的线性泛函都是高斯随机变量,则 $X(t)$ 为高斯过程。

如果概率密度函数有如下形式,称随机变量 Y 是高斯分布的:

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left[-\frac{(y - \mu_Y)^2}{2\sigma_Y^2}\right] \quad (1.80)$$

其中 μ_Y 是均值,而 σ_Y^2 是随机变量 Y 的方差。图 1.13 给出了特殊情况下,即高斯随机变量 Y 的均值 μ_Y 为 0、方差 σ_Y^2 为 1 时的概率密度曲线。这时

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right)$$

这一正态高斯分布通常写为 $N(0, 1)$ 。

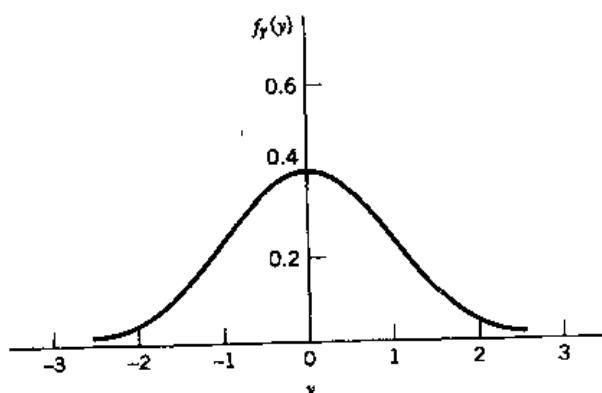


图 1.13 正态高斯分布

高斯过程有两个主要优点。第一，高斯过程有许多易于分析的性质，本节将会对这些性质进行讨论。第二，物理现象所产生的随机过程往往适合采用高斯模型。进一步说，用于描述物理现象的高斯模型通常都能被实验所证实。因此，很多物理现象都可以采用高斯模型，同时高斯过程非常便于进行数学处理，这使得高斯过程在通信系统学习中显得十分重要。

1.8.1 中心极限定理

在特定时刻观测到的随机变量实际上是大量相互独立的随机事件的总和，可以选择高斯过程作为这些互不相同的自然现象的模型，中心极限定理为此提供了数学依据。为了将这一重要定理公式化，令 $X_i (i = 1, 2, \dots, N)$ 为一组满足以下要求的随机变量：

1. X_i 是统计独立的。
2. X_i 具有相同的概率分布，均值为 μ_X ，方差为 σ_X^2 。

这些 X_i 可以组成一组独立同分布的随机变量。对这些随机变量进行如下归一化：

$$Y_i = \frac{1}{\sigma_X} (X_i - \mu_X), \quad i = 1, 2, \dots, N$$

于是有

$$E[Y_i] = 0$$

且

$$\text{var}[Y_i] = 1$$

定义随机变量

$$V_N = \frac{1}{\sqrt{N}} \sum_{i=1}^N Y_i$$

中心极限定理表明， V_N 的概率分布在随机变量 N 趋于无穷时的极限逼近于标准高斯分布 $\mathcal{N}(0, 1)$ 。

不过，需要注意的是，中心极限定理只给出了当 N 趋于无穷时，标准化随机变量 V_N 的概率分布的极限形式。当 N 为有限值时，即使 N 相当大，高斯极限有时也只能提供 V_N 的一个很差的相对近似。

1.8.2 高斯过程的性质

下面给出高斯过程的一些有用的性质。

性质 1

如果高斯过程 $X(t)$ 通过一个稳定线性滤波器，那么，滤波器的输出 $Y(t)$ 也是高斯型的。

这个性质由基于式(1.79)的高斯过程的定义很容易得到。如图 1.8 所示，有一个冲激响应为 $b(t)$ 的线性时不变滤波器，随机过程 $X(t)$ 为输入，随机过程 $Y(t)$ 为输出。假定 $X(t)$ 是高斯过程，随机过程 $X(t)$ 和 $Y(t)$ 的关系由下面的卷积积分给出：

$$Y(t) = \int_0^T b(t - \tau) X(\tau) d\tau, \quad 0 \leq t < \infty \quad (1.81)$$

假定冲激响应 $b(t)$ 使得对于 $0 \leq t < \infty$ 内的所有 t , 输出 $Y(t)$ 的均方值有限, 则在此范围内 $Y(t)$ 是有定义的。要证明输出 $Y(t)$ 也是高斯型的, 必须证明它的任一泛函为高斯型随机变量。也就是说, 如果定义随机变量

$$Z = \int_0^\infty g_Y(t) \int_0^T b(t-\tau) X(\tau) d\tau dt \quad (1.82)$$

那么, 对每一个函数 $g_Y(t)$, Z 必须是高斯随机变量, 同时它的均方值有限。改变式(1.82)中的积分顺序, 得到

$$Z = \int_0^T g(\tau) X(\tau) d\tau \quad (1.83)$$

其中

$$g(\tau) = \int_0^\infty g_Y(t) b(t-\tau) d\tau \quad (1.84)$$

既然 $X(t)$ 是一个高斯过程, 由式(1.83)得到的 Z 就一定是一个高斯随机变量。因此, 如果线性滤波器的输入 $X(t)$ 是高斯过程, 则输出 $Y(t)$ 也是高斯过程。值得注意的是, 虽然证明是在线性时不变滤波器的前提下得到的, 但这个性质对任意稳定的线性滤波器都成立。

性质 2

设在 t_1, t_2, \dots, t_n 时刻观测随机过程 $X(t)$ 所得的样本或随机变量组为 $X(t_1), X(t_2), \dots, X(t_n)$ 。如果 $X(t)$ 是高斯型的, 那么对任意的 n 来说, 这组随机变量是联合高斯型的, 它们的 n 维联合概率密度函数可以用它们的均值组

$$\mu_{X(t_i)} = E[X(t_i)], \quad i = 1, 2, \dots, n$$

和协方差组

$$C_X(t_k, t_i) = E[(X(t_k) - \mu_{X(t_k)})(X(t_i) - \mu_{X(t_i)})], \quad k, i = 1, 2, \dots, n$$

完全确定。

用 $n \times 1$ 的矢量 \mathbf{X} 表示高斯过程 $X(t)$ 在 t_1, \dots, t_n 时刻抽样所获得的随机变量组 $X(t_1), \dots, X(t_n)$ 。由 \mathbf{x} 表示 \mathbf{X} 的一个值, 由性质 2 可知, 随机矢量 \mathbf{X} 是多维高斯分布的, 这个分布由下面的矩阵式定义:

$$f_{X(t_1), \dots, X(t_n)}(\mathbf{x}_1, \dots, \mathbf{x}_n) = \frac{1}{(2\pi)^{n/2} \Delta^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\right) \quad (1.85)$$

这里, 上标 T 为转置符号, 并且

$\boldsymbol{\mu}$ = 均值矢量

$$= [\mu_1, \mu_2, \dots, \mu_n]^T$$

$\boldsymbol{\Sigma}$ = 协方差矩阵

$$= \{C_X(t_k, t_j)\}_{k,j=1}^n$$

$\boldsymbol{\Sigma}^{-1}$ = 协方差矩阵的逆阵

Δ = 协方差矩阵 $\boldsymbol{\Sigma}$ 的行列式

性质 2 常被作为高斯过程的定义。但这一定义在用来估计高斯过程通过滤波器的结果

时,要比用式(1.79)的定义来进行估计更为困难。

可以将性质2扩展到两个(或更多)随机过程的情况。设在 $\{t_i, i = 1, 2, \dots, n\}$ 时刻观测第一个随机过程 $X(t)$ 和在 $\{\mu_k, k = 1, 2, 3, \dots, m\}$ 时刻观测第二个随机过程 $Y(t)$ 所得到的随机变量组分别为 $X(t_1), X(t_2), \dots, X(t_n)$ 和 $Y(u_1), Y(u_2), \dots, Y(u_m)$,如果对任意 n 和 m ,这两组随机变量是联合高斯分布的,则称过程 $X(t)$ 和 $Y(t)$ 是联合高斯分布的。除了要知道 $X(t)$ 和 $Y(t)$ 的方差和自相关函数外,还要知道任意观测时刻 (t_i, u_k) 的互协方差函数

$$E[(X(t_i) - \mu_{X(t_i)})(Y(u_k) - \mu_{Y(u_k)})] = R_{XY}(t_i, u_k) - \mu_{X(t_i)}\mu_{Y(u_k)}$$

这一额外信息可以由过程 $X(t)$ 和 $Y(t)$ 的互相关函数 $R_{XY}(t_i, u_k)$ 来得到。

性质3

如果高斯过程是平稳的,那么,一定也是严平稳的。

这可以由性质2直接导出。

性质4

如果在时刻 t_1, t_2, \dots, t_n 对高斯过程 $X(t)$ 抽样所得的随机变量 $X(t_1), X(t_2), \dots, X(t_n)$ 是不相关的,也就是

$$E[(X(t_k) - \mu_{X(t_k)})(X(t_i) - \mu_{X(t_i)})] = 0, \quad i \neq k$$

则称这些随机变量是统计独立的。

$X(t_1), \dots, X(t_n)$ 不相关意味着协方差矩阵 Σ 是一个对角阵,即

$$\Sigma = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix}$$

其中

$$\sigma_i^2 = E[(X(t_i) - E[X(t_i)])^2], \quad i = 1, 2, \dots, n$$

在这一条件下,式(1.85)的多维高斯分布简化为

$$f_X(\mathbf{x}) = \prod_{i=1}^n f_{X_i}(x_i)$$

其中 $X_i = X(t_i)$,并且

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x_i - \mu_{X_i})^2}{2\sigma_i^2}\right)$$

总之,如果高斯随机变量 $X(t_1), \dots, X(t_n)$ 是不相关的,那么它们是统计独立的,这同时也意味着这组随机变量的联合概率密度分布可以表示为该组各个随机变量概率密度函数的乘积。

1.9 噪声

噪声一词通常用来指在通信系统中的干扰信号和无用信号。对于噪声,往往不能做到完

全控制。通信系统中有许多潜在的噪声源。噪声源可以是系统外部的(如空气中的电波噪声、宇宙噪声、人为噪声等),也可以是系统内部的。系统内部的噪声包括由电源的电压或电流自发波动引起的一类重要噪声。⁶这一类噪声代表了在使用电子设备的通信系统中传输和检测信号的基本极限。两个最常见的由电路自发波动引起的噪声是散弹噪声和热噪声,下面将进行描述。

1.9.1 散弹噪声

散弹噪声是由二极管和晶体管这类电子元件中的电流流动的离散特性引起的。例如,在一个光电探测电路中,由于受到恒定强度光源的照射,电子由阴极发射,从而产生电流脉冲。电子的发射时刻用 τ_k 表示,其中 $-\infty < k < \infty$ 。假定电子的随机发射持续时间很长,因此,流经光电探测器的总电流可以认为是无限电流脉冲的和,即

$$X(t) = \sum_{k=-\infty}^{\infty} b(t - \tau_k) \quad (1.86)$$

其中 $b(t - \tau_k)$ 为 τ_k 时刻产生的电流脉冲。 $X(t)$ 为一个平稳随机过程,称为散弹噪声。

时间区间 $[0, t]$ 内发射的电子数量 $N(t)$ 为离散随机过程,其值随着电子的发射而增大。图 1.14 给出了这一过程的一个样本函数。在时间区间 $[t, t + t_0]$ 内发射的电子数 ν 的均值的定义为

$$E[\nu] = \lambda t_0 \quad (1.87)$$

参数 λ 是一个称为过程率的常量,则在区间 $[t, t + t_0]$ 内发射的所有电子数为

$$\nu = N(t + t_0) - N(t)$$

符合均值为 λt_0 的泊松分布。区间 $[t, t + t_0]$ 内发射 k 个电子的概率定义为

$$P(\nu = k) = \frac{(\lambda t_0)^k}{k!} e^{-\lambda t_0} \quad k = 0, 1, \dots \quad (1.88)$$

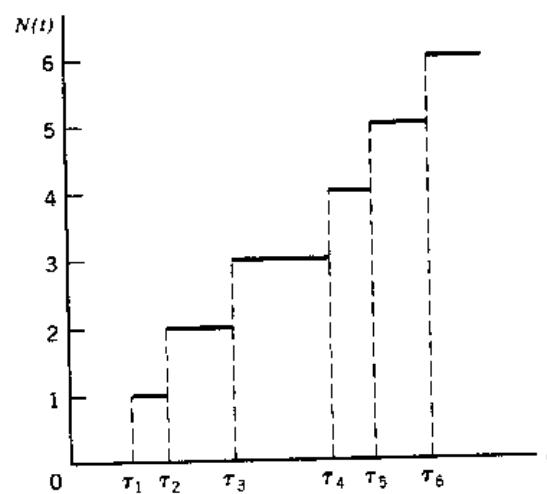


图 1.14 泊松过程的样本函数

遗憾的是,要对式(1.86)给出的散弹噪声过程的特性进行详细描述,在数学上是一件非常

困难的事情。这里只能简单地得到散弹噪声过程的前二阶矩：

- $X(t)$ 的均值为

$$\mu_X = \lambda \int_{-\infty}^{\infty} b(t) dt \quad (1.89)$$

其中 λ 为过程率, $b(t)$ 为电流脉冲的波形。

- $X(t)$ 的自协方差函数为

$$C_X(\tau) = \lambda \int_{-\infty}^{\infty} b(t)b(t + \tau) dt \quad (1.90)$$

这个结论就是坎贝儿定理。

作为特例, 波形 $b(t)$ 由幅度为 A 、持续时间为 T 的矩形脉冲构成, 散弹噪声过程 $X(t)$ 的均值为 λAT , 自协方差函数为

$$C_X(\tau) = \begin{cases} \lambda A^2(T - |\tau|), & |\tau| < T \\ 0, & |\tau| \geq T \end{cases}$$

它有与图 1.7 相似的三角形图像。

1.9.2 热噪声

热噪声是由于导体中电子的随机运动而产生的电噪声。出于实用的考虑, 将 Δf Hz 带宽上的电阻两端热噪声电压 V_{TN} 的均方值表示为

$$E[V_{TN}^2] = 4kTR \Delta f \text{ V}^2 \quad (1.91)$$

其中 k 为玻尔兹曼常数, 等于 1.38×10^{-23} J/K, T 为用 K(开尔文) 表示的绝对温度, R 为用 Ω 表示的电阻值。可以用戴维南等效电路来对噪声电阻进行建模, 即在无噪声电阻上串联一个均方值为 $E[V_{TN}^2]$ 的噪声电压源, 如图 1.15(a) 所示。也可用诺顿等效电路来表示, 即在一个无噪声电导上并联一个噪声电流源, 如图 1.15(b) 所示。噪声电流源的均方值为

$$\begin{aligned} E[I_{TN}^2] &= \frac{1}{R^2} E[V_{TN}^2] \\ &= 4kTG \Delta f \text{ A}^2 \end{aligned} \quad (1.92)$$

其中电导 $G = 1/R$ 。由于电阻中的电子数目巨大, 其随机运动又是相互统计独立的, 因此由中心极限定理可知, 热噪声是零均值的高斯分布。

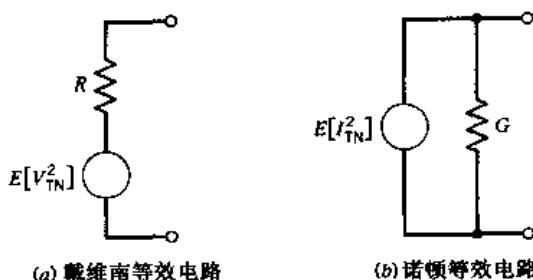


图 1.15 噪声电阻模型

噪声计算涉及到功率输出,而最大功率输出定理适用于噪声计算。最大功率输出定理表明,当负载电阻 R_L 等于电源内阻 R 时,将得到最大可能的功率输出。在这种匹配状态下,电源产生的功率将被平均地分配到电源内阻和负载阻抗上,这时输出到负载上的功率被称为有效功率。对图 1.15(a)所示的戴维南等效电路和 1.15(b)所示的诺顿等效电路运用最大功率输出定理,结果表明噪声电阻产生了一个值为 $kT\Delta f$ 瓦的有效噪声功率。

1.9.3 白噪声

通信系统的噪声分析通常是基于一种称为白噪声的理想噪声形式,白噪声的噪声功率谱密度与工作频率无关。白光在可见电磁辐射的频率上是等量分布的,形容词“白”就是借用这个概念。样本函数为 $w(t)$ 的白噪声的功率谱密度为

$$S_W(f) = \frac{N_0}{2} \quad (1.93)$$

如图 1.16(a)所示。 N_0 的单位是 W/Hz(瓦每赫)。参数 N_0 通常称为通信系统接收机的输入量级。其表达式为

$$N_0 = kT_e \quad (1.94)$$

其中 k 为玻尔兹曼常数, T_e 为接收机的等效噪声温度。^[1] 等效噪声温度是维持噪声电阻所需的温度。噪声电阻被连接到无噪声系统的输入端,它在系统的输出端产生与实际系统所有噪声源等效的噪声。等效噪声温度的重要特性在于它只取决于系统的参数。

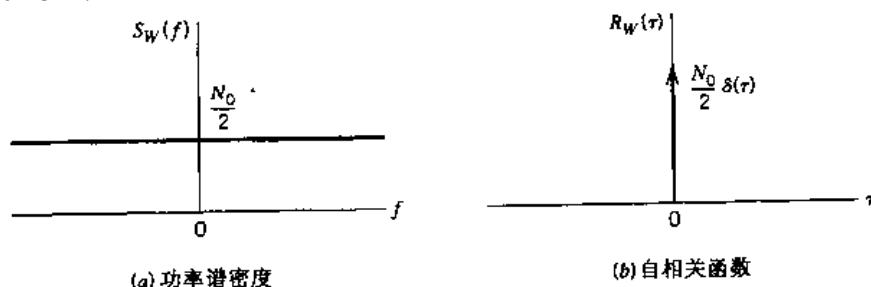


图 1.16 白噪声特性

由于自相关函数是功率谱密度的傅里叶反变换,于是得到白噪声的自相关函数

$$R_W(\tau) = \frac{N_0}{2} \delta(\tau) \quad (1.95)$$

也就是说,白噪声的自相关函数为 $\tau = 0$ 时刻加权因子为 $N_0/2$ 的冲激函数,如图 1.16(b)所示。注意到在 $\tau \neq 0$ 时 $R_W(\tau) = 0$ 。因此,两个不同的白噪声样本,不管它们在时间上靠得多么近,都是不相关的。如果白噪声 $w(t)$ 是高斯型的,那么它们也是统计独立的。在这个意义上,高斯白噪声代表“随机性”的极限。

严格地讲,白噪声有无限的平均功率,所以在物理上是不可实现的。但白噪声可用式(1.93)和式(1.95)进行简单的数学表达,这使得白噪声在统计系统分析中很有用。

在线性系统分析中,白噪声的作用与冲激函数相同。与冲激脉冲一样,必须让白噪声通过与带限系统类似的系统,才能观察到它的影响。因此,只要系统输入端的噪声带宽比系统带宽大得多,就可以认为系统是白噪声的。

例 1.10 白噪声通过理想低通滤波器

假设一个零均值、功率谱密度为 $N_0/2$ 的白噪声通过一个带宽为 B 、幅度为 1 的低通滤波器。滤波器输出端噪声 $n(t)$ 的功率谱密度为

$$S_N(f) = \begin{cases} \frac{N_0}{2}, & -B < f < B \\ 0, & |f| > B \end{cases} \quad (1.96)$$

$n(t)$ 的自相关函数是如图 1.17(a) 所示的功率谱密度的傅里叶反变换：

$$\begin{aligned} R_N(\tau) &= \int_{-B}^B \frac{N_0}{2} \exp(j2\pi f\tau) df \\ &= N_0 B \operatorname{sinc}(2B\tau) \end{aligned} \quad (1.97)$$

自相关函数如图 1.17(b) 所示, $R_N(\tau)$ 在原点具有最大值 $N_0 B$, 在 $\tau = \pm k/2B$, $k = 1, 2, 3, \dots$ 处, 其值为 0。

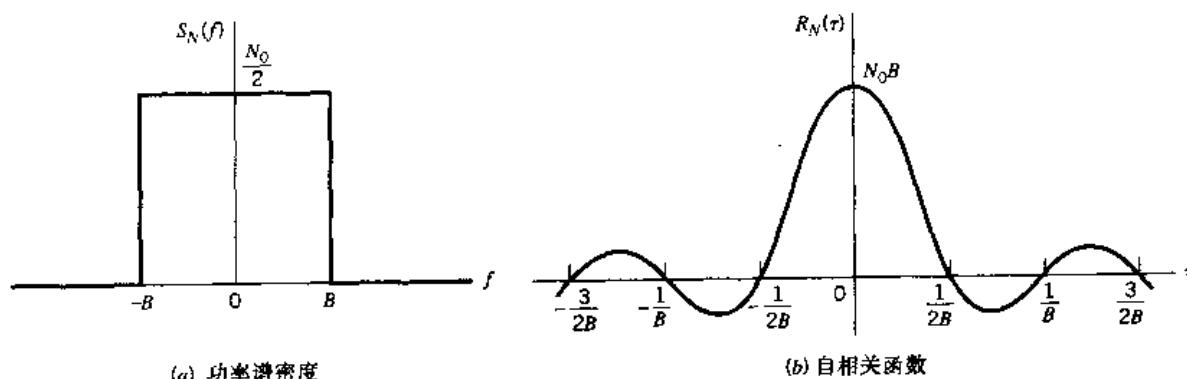


图 1.17 白噪声通过低通滤波器的特性

由于输入噪声 $w(t)$ 是高斯型的, 则滤波器输出端的带限噪声也是高斯型的。假设对 $n(t)$ 以 $2B$ 次/秒的速率进行抽样, 由图 1.17(b) 可知, 输出端的噪声抽样是不相关的, 因而也是统计独立的。相应地, 它们的联合概率密度就是各个概率密度的乘积。注意, 各个抽样的均值都为零, 方差都为 $N_0 B$ 。

例 1.11 白噪声和正弦信号的相关

考虑样本函数

$$w'(t) = \sqrt{\frac{2}{T}} \int_0^T w(t) \cos(2\pi f_c t) dt \quad (1.98)$$

它是具有高斯白噪声 $w(t)$ 的以正弦信号 $\sqrt{2/T} \cos(2\pi f_c t)$ 作为输入的相关器的输出; 系数因子 $\sqrt{2/T}$ 的作用是使输入的正弦信号在区间 $0 \leq t \leq T$ 上具有单位能量(这个问题在“背景与预览”中曾经遇到, 但没有进行详细阐述)。 $w(t)$ 是零均值的, 因此相关输出 $w'(t)$ 也是零均值的。相关器输出的方差为

$$\begin{aligned}\sigma^2 &= E \left[\frac{2}{T} \int_0^T \int_0^T w(t_1) \cos(2\pi f_c t_1) w(t_2) \cos(2\pi f_c t_2) dt_1 dt_2 \right] \\ &= \frac{2}{T} \int_0^T \int_0^T E[w(t_1)w(t_2)] \cos(2\pi f_c t_1) \cos(2\pi f_c t_2) dt_1 dt_2 \\ &= \frac{2}{T} \int_0^T \int_0^T R_w(t_1, t_2) \cos(2\pi f_c t_1) \cos(2\pi f_c t_2) dt_1 dt_2\end{aligned}$$

其中 $R_w(t_1, t_2)$ 为白噪声 $w(t)$ 的相关函数。由公式(1.95)可得

$$R_w(t_1, t_2) = \frac{N_0}{2} \delta(t_1 - t_2)$$

其中 $N_0/2$ 为白噪声 $w(t)$ 的功率谱密度。相应地, 方差 σ^2 的表达式可以简化为

$$\sigma^2 = \frac{N_0}{2} \times \frac{2}{T} \int_0^T \int_0^T \delta(t_1 - t_2) \cos(2\pi f_c t_1) \cos(2\pi f_c t_2) dt_1 dt_2$$

引用冲激函数的移位特性, 即

$$\int_{-\infty}^{\infty} g(t) \delta(t) dt = g(0)$$

其中 $g(t)$ 是时间连续的函数, 假设 $g(0)$ 为其在 $t=0$ 时的值。于是, 方差 σ^2 可以进一步简化为

$$\sigma^2 = \frac{N_0}{2} \times \frac{2}{T} \int_0^T \cos^2(2\pi f_c t) dt = \frac{N_0}{2} \quad (1.99)$$

这里, 假设正弦信号输入的频率 f_c 为 T 的倒数的整数倍。

1.10 窄带噪声

通信系统接收机通常包括用于预处理接收信号的一些装置。预处理装置采用窄带滤波器, 其带宽正好足够使接收信号的调制成分无失真地通过, 同时又能限制多余的噪声通过。滤波器输出端的噪声称为窄带噪声。如图 1.18(a)所示, 窄带噪声的频谱成分集中在中心频率 $\pm f_c$ 的附近, 其样本函数 $n(t)$ 与频率为 f_c 的正弦信号类似, 如图 1.18(b)所示, 不同的是 $n(t)$ 在幅度和相位上都有缓慢的起伏。

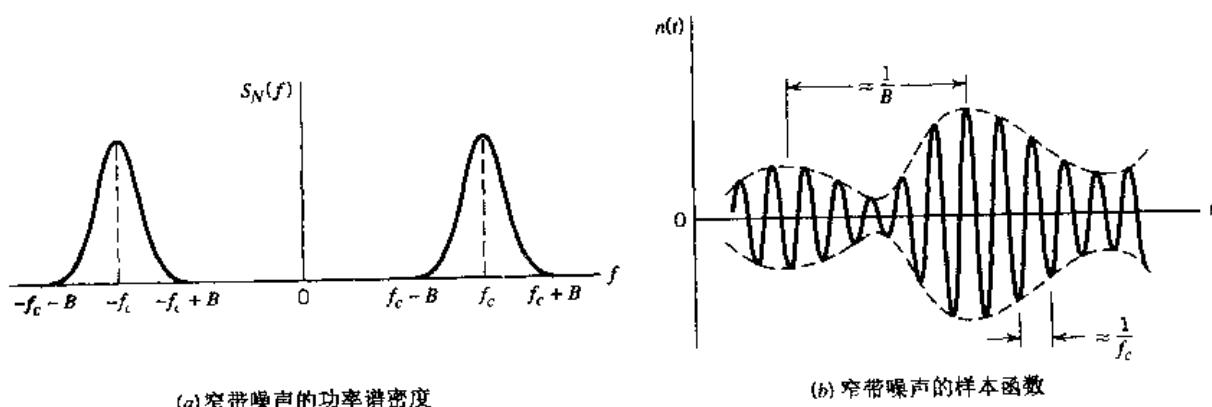


图 1.18 窄带噪声的功率谱密度和样本函数

要分析窄带噪声对通信系统性能的影响,首先要得到它的数学表达式。在应用上,窄带噪声有两种具体的表达方式:

1. 根据一对被称为同相和正交的分量来定义窄带噪声。
2. 根据两个被称为包络和相位的分量来定义窄带噪声。

给定同相和正交分量就能够得到幅度和相位分量,反之亦然。这两种独立的表示不仅是进行通信系统噪声分析的基础,也是窄带噪声本身的特性。下面分别对这两种表示法进行详细阐述。

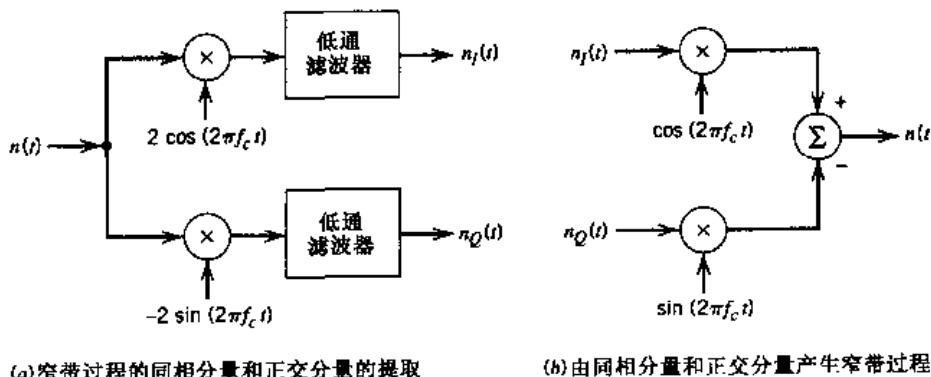
1.11 基于同相和正交分量的窄带噪声表示法

考虑带宽为 $2B$ 、中心频率为 f_c 的窄带噪声 $n(t)$,如图 1.18 所示,根据附录 2 中带通信号与系统的相关理论,可将 $n(t)$ 表达成规范(标准)的形式:

$$n(t) = n_I(t) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t) \quad (1.100)$$

其中 $n_I(t)$ 为 $n(t)$ 的同相分量,而 $n_Q(t)$ 则是 $n(t)$ 的正交分量。 $n_I(t)$ 和 $n_Q(t)$ 都是低通信号。它们和中心频率 f_c 一起,完整地表达了窄带噪声 $n(t)$ 。

对于给定的窄带噪声 $n(t)$,可以采用图 1.19 所示的方法提取同相和正交分量。假定方案中两个低通滤波器都是理想的,并具有带宽 B ,也就是窄带噪声 $n(t)$ 带宽的一半。图 1.19(a)的方案是从等式(1.100)得来的。如图 1.19(b)所示,给定同相和正交分量,即可利用式(1.100)得到窄带噪声 $n(t)$ 。图 1.19(a)和图 1.19(b)所示的方案可以分别视为窄带噪声分析器和合成器。



(a)窄带过程的同相分量和正交分量的提取

(b)由同相分量和正交分量产生窄带过程

图 1.19 窄带过程的同相分量和正交分量的提取与合成

窄带噪声的同相和正交分量具有一些重要特性,总结如下:

1. 窄带噪声 $n(t)$ 的同相分量 $n_I(t)$ 和正交分量 $n_Q(t)$ 具有零均值。
2. 如果窄带噪声 $n(t)$ 是高斯型的,则同相分量 $n_I(t)$ 和正交分量 $n_Q(t)$ 也都是高斯型的。
3. 如果窄带噪声 $n(t)$ 是平稳的,则同相分量 $n_I(t)$ 和正交分量 $n_Q(t)$ 也都是平稳的。
4. 同相分量 $n_I(t)$ 和正交分量 $n_Q(t)$ 具有相同的功率谱密度。这个功率谱密度与窄带噪声 $n(t)$ 的功率谱密度 $S_N(f)$ 关系为

$$S_{N_I}(f) = S_{N_Q}(f) = \begin{cases} S_N(f - f_c) + S_N(f + f_c), & -B \leq f \leq B \\ 0, & \text{其他} \end{cases} \quad (1.101)$$

其中,假设 $S_N(f)$ 占用了频率间隔 $f_c - B \leq |f| \leq f_c + B$ 且 $f_c > B$ 。

5. 同相分量 $n_I(t)$ 和正交分量 $n_Q(t)$ 具有相同的方差,其值等于 $n(t)$ 的方差。

6. 窄带噪声 $n(t)$ 的同相分量和正交分量的互功率谱密度为纯虚数,如下式:

$$\begin{aligned} S_{N_I N_Q}(f) &= -S_{N_Q N_I}(f) \\ &= \begin{cases} j[S_N(f + f_c) - S_N(f - f_c)], & -B \leq f \leq B \\ 0, & \text{其他} \end{cases} \end{aligned} \quad (1.102)$$

7. 假定窄带噪声 $n(t)$ 是高斯型的,并且其功率谱密度 $S_N(f)$ 关于中心频率 f_c 对称,那么同相分量 $n_I(t)$ 和正交分量 $n_Q(t)$ 就是统计独立的。

关于这些特性的进一步讨论,读者可以参考习题 1.28 和习题 1.29。

例 1.12 白噪声通过理想带通滤波器

零均值且功率谱密度为 $N_0/2$ 的高斯白噪声通过理想带通滤波器,其幅度响应为 1,中心频率为 f_c ,带宽为 $2B$ 。经过滤波器的噪声 $n(t)$ 的功率谱密度特性如图 1.20(a)所示,试确定 $n(t)$ 的自相关函数、同相分量和正交分量。

$n(t)$ 的自相关函数是如图 1.20(a)所示的功率谱密度特性的傅里叶反变换:

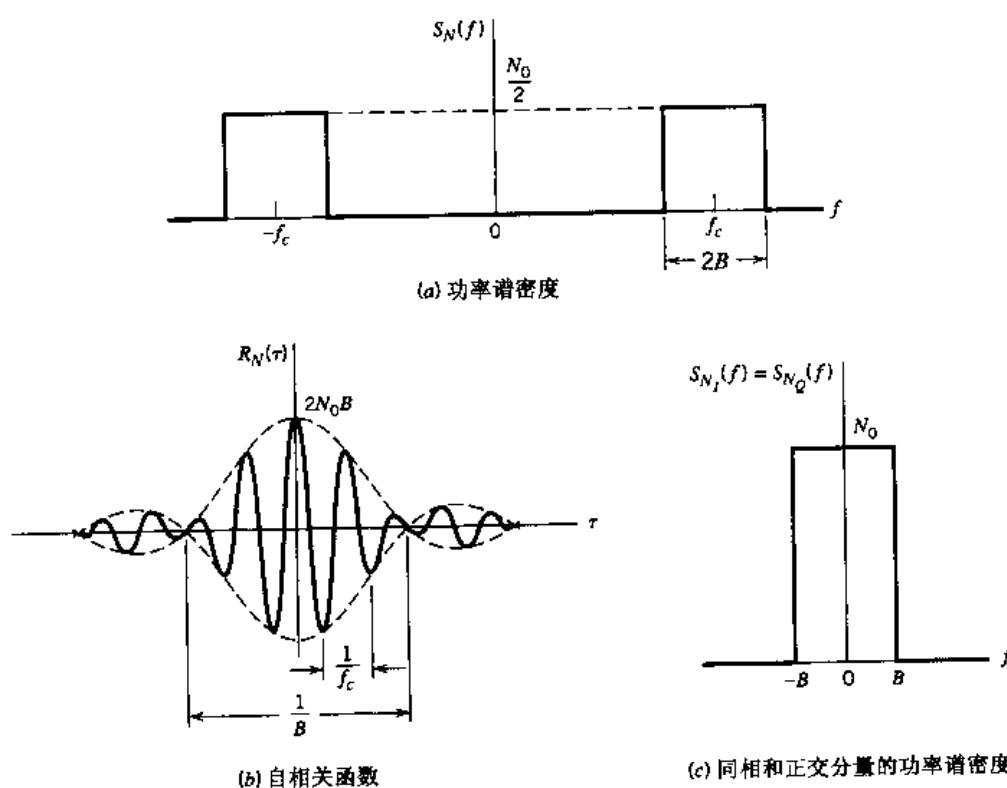


图 1.20 白噪声通过理想带通滤波器的特性

$$\begin{aligned}
 R_N(\tau) &= \int_{-f_c-B}^{-f_c+B} \frac{N_0}{2} \exp(j2\pi f\tau) df + \int_{f_c-B}^{f_c+B} \frac{N_0}{2} \exp(j2\pi f\tau) df \\
 &= N_0 B \operatorname{sinc}(2B\tau) [\exp(-j2\pi f_c\tau) + \exp(j2\pi f_c\tau)] \\
 &= 2N_0 B \operatorname{sinc}(2B\tau) \cos(2\pi f_c\tau)
 \end{aligned} \tag{1.103}$$

自相关函数如图 1.20(b) 所示。

图 1.20(a) 所示的功率谱密度关于 $\pm f_c$ 对称。同相分量 $n_I(t)$ 和正交分量 $n_Q(t)$ 的相应功率谱密度特性如图 1.21(c) 所示。 $n_I(t)$ 和 $n_Q(t)$ 的自相关函数为(参考例 1.10)：

$$R_{N_I}(\tau) = R_{N_Q}(\tau) = 2N_0 B \operatorname{sinc}(2B\tau) \tag{1.104}$$

1.12 基于包络和相位分量的窄带噪声表示法

1.11 节研究了基于同相和正交分量的窄带噪声 $n(t)$ 表示法。我们也可以用包络和相位分量来表示窄带噪声 $n(t)$ ：

$$n(t) = r(t) \cos[2\pi f_c t + \psi(t)] \tag{1.105}$$

其中

$$r(t) = [n_I^2(t) + n_Q^2(t)]^{1/2} \tag{1.106}$$

且

$$\psi(t) = \tan^{-1} \left[\frac{n_Q(t)}{n_I(t)} \right] \tag{1.107}$$

函数 $r(t)$ 称为 $n(t)$ 的幅度, $\psi(t)$ 称为 $n(t)$ 的相位。

幅度 $r(t)$ 和相位 $\psi(t)$ 都是低通随机过程的样本函数。如图 1.18(b) 所示, $r(t)$ 包络的两个相邻峰值之间的时间间隔约为 $1/B$, 其中 $2B$ 为窄带噪声 $n(t)$ 的带宽。

幅度 $r(t)$ 和相位 $\psi(t)$ 的概率分布可以从 $n_I(t)$ 和 $n_Q(t)$ 的概率分布获得。令 N_I 和 N_Q 分别表示通过观测(在某些固定时刻)由样本函数 $n_I(t)$ 和 $n_Q(t)$ 表示的随机过程所得到的随机变量。 N_I 和 N_Q 为独立的高斯随机变量, 其均值为零, 方差为 σ^2 , 于是联合概率密度函数为

$$f_{N_I, N_Q}(n_I, n_Q) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{n_I^2 + n_Q^2}{2\sigma^2}\right) \tag{1.108}$$

相应地, N_I 落在 n_I 和 $n_I + dn_I$ 之间, N_Q 落在 n_Q 和 $n_Q + dn_Q$ 之间。(即随机变量 N_I 和 N_Q 位于图 1.21(a) 所示的阴影区域内。)二者的联合事件概率由下式给出

$$f_{N_I, N_Q}(n_I, n_Q) dn_I dn_Q = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{n_I^2 + n_Q^2}{2\sigma^2}\right) dn_I dn_Q \tag{1.109}$$

如图 1.21(a), 定义变换:

$$n_I = r \cos \psi \tag{1.110}$$

$$n_Q = r \sin \psi \tag{1.111}$$

从极限意义上说, 图 1.21(a) 和图 1.21(b) 所示阴影中的两个增量区域是等同的, 即

$$dn_I dn_Q = r dr d\psi \quad (1.112)$$

令 R 和 Ψ 分别为通过在时刻 t 观察由包络 $r(t)$ 和相位 $\psi(t)$ 表示的随机过程所得到的随机变量。将式(1.110)~(1.112)代入式(1.109), 则随机变量 R 和 Ψ 落在图1.21(b) 所示阴影部分的概率为

$$\frac{r}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr d\psi$$

也就是说 R 和 Ψ 的联合概率密度函数为

$$f_{R,\Psi}(r, \psi) = \frac{r}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right) \quad (1.113)$$

其中, 概率密度函数与角度 ψ 无关, 这说明随机变量 R 和 Ψ 是统计独立的。可将 $f_{R,\Psi}(r, \psi)$ 表示为 $f_R(r)$ 和 $f_\Psi(\psi)$ 的乘积。特别地, 代表相位的随机变量 Ψ 在 $0 \sim 2\pi$ 之间为均匀分布, 即

$$f_\Psi(\psi) = \begin{cases} \frac{1}{2\pi}, & 0 \leq \psi \leq 2\pi \\ 0, & \text{其他} \end{cases} \quad (1.114)$$

从而, 随机变量 R 的概率密度函数为

$$f_R(r) = \begin{cases} \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), & r \geq 0 \\ 0, & \text{其他} \end{cases} \quad (1.115)$$

其中, σ^2 为原始窄带噪声 $n(t)$ 的方差。具有式(1.115)所示概率密度的随机变量被称为是瑞利分布的。^[8]

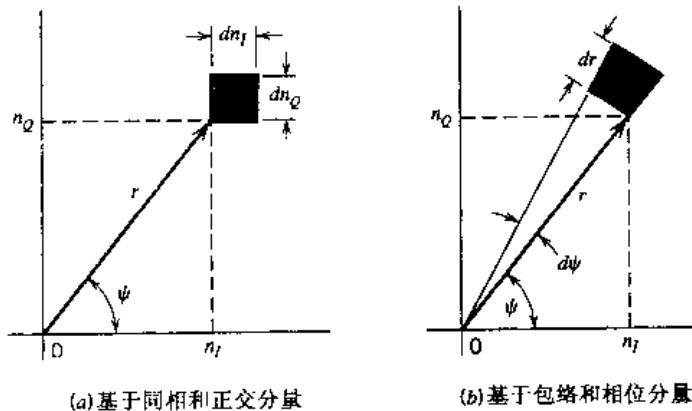


图 1.21 窄带噪声两种表示法的坐标图解

为了图表表示的方便, 令

$$v = \frac{r}{\sigma} \quad (1.116)$$

$$f_V(v) = \sigma f_R(r) \quad (1.117)$$

由此可得出式(1.115)的瑞利分布的标准化形式:

$$f_V(v) = \begin{cases} v \exp\left(-\frac{v^2}{2}\right), & v \geq 0 \\ 0, & \text{其他} \end{cases} \quad (1.118)$$

如图 1.22 所示。密度函数 $f_V(v)$ 的峰值出现在 $v = 1$ 处, 其值为 0.607。同样需要注意的是, 与高斯分布不同, 瑞利分布在 $v < 0$ 时等于零, 这是因为幅度 $r(t)$ 只能为非负值。

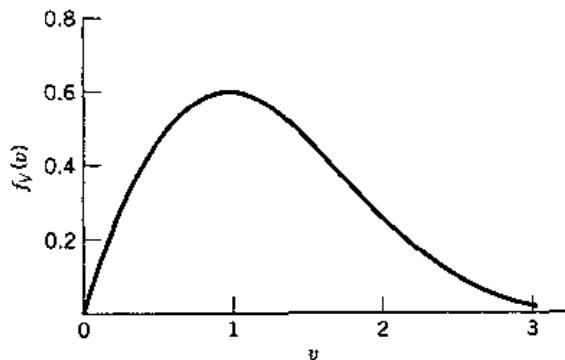


图 1.22 标准瑞利分布

1.13 正弦信号加窄带噪声

假设在正弦信号 $A \cos(2\pi f_c t)$ 上叠加窄带噪声 $n(t)$, 其中 A 和 f_c 都是常量。假设正弦信号的频率与噪声的标称载波频率相等。则正弦信号上叠加噪声的一个样本函数可表示为

$$x(t) = A \cos(2\pi f_c t) + n(t) \quad (1.119)$$

由窄带噪声 $n(t)$ 的同相和正交分量表示法, 可得

$$x(t) = n_I(t) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t) \quad (1.120)$$

其中

$$n_I(t) = A + n_I(t) \quad (1.121)$$

假设 $n(t)$ 是均值为零、方差为 σ^2 的高斯过程。则

1. $n'_I(t)$ 和 $n_Q(t)$ 为高斯型, 并且是统计独立的。
2. $n'_I(t)$ 的均值为 A , $n_Q(t)$ 的均值为零。
3. $n'_I(t)$ 和 $n_Q(t)$ 的方差都是 σ^2 。

分别与 $n'_I(t)$ 和 $n_Q(t)$ 对应的随机变量 N'_I 和 N_Q 的联合密度函数为

$$f_{N'_I, N_Q}(n'_I, n_Q) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{(n'_I - A)^2 + n_Q^2}{2\sigma^2}\right] \quad (1.122)$$

令 $r(t)$ 表示 $x(t)$ 的包络, $\psi(t)$ 表示其相位。由等式(1.120)可知

$$r(t) = \{[n'_I(t)]^2 + n_Q^2(t)\}^{1/2} \quad (1.123)$$

和

$$\psi(t) = \tan^{-1} \left[\frac{n_Q(t)}{n_I(t)} \right] \quad (1.124)$$

与1.12节瑞利分布的推导过程同理,对于固定时间 t ,对应于 $r(t)$ 和 $\psi(t)$ 的随机变量 R 和 Ψ 的联合概率密度函数为

$$f_{R,\Psi}(r, \psi) = \frac{r}{2\pi\sigma^2} \exp \left(-\frac{r^2 + A^2 - 2Ar \cos \psi}{2\sigma^2} \right) \quad (1.125)$$

但这里不能将联合概率密度函数 $f_{R,\Psi}(r, \psi)$ 表示成 $f_R(r)$ 和 $f_\Psi(\psi)$ 的乘积。这是因为上式中有一项包含了两个随机变量的乘积 $r \cos \psi$ 。因此,在正弦信号分量的幅度 A 为非零值的情况下,随机变量 R 和 Ψ 是相关的。

我们最感兴趣的是 R 的概率密度函数。为确定这个概率密度函数,将式(1.125)对 ψ 进行积分,从而获得边缘概率密度

$$\begin{aligned} f_R(r) &= \int_0^{2\pi} f_{R,\Psi}(r, \psi) d\psi \\ &= \frac{r}{2\pi\sigma^2} \exp \left(-\frac{r^2 + A^2}{2\sigma^2} \right) \int_0^{2\pi} \exp \left(\frac{Ar}{\sigma^2} \cos \psi \right) d\psi \end{aligned} \quad (1.126)$$

可以根据定义第一类零阶修正贝塞尔函数(见附录3)的积分式来确定式(1.126)右侧的积分值,即

$$I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp(x \cos \psi) d\psi \quad (1.127)$$

进一步地,令 $x = Ar/\sigma^2$,则等式(1.126)可简化为

$$f_R(r) = \frac{r}{\sigma^2} \exp \left(-\frac{r^2 + A^2}{2\sigma^2} \right) I_0 \left(\frac{Ar}{\sigma^2} \right) \quad (1.128)$$

这个关系式被称为莱斯分布。^[9]

与瑞利分布类似,莱斯分布的表示可进行如下简化:

$$v = \frac{r}{\sigma} \quad (1.129)$$

$$a = \frac{A}{\sigma} \quad (1.130)$$

$$f_V(v) = \sigma f_R(r) \quad (1.131)$$

由此可得式(1.128)中莱斯分布的标准化形式:

$$f_V(v) = v \exp \left(-\frac{v^2 + a^2}{2} \right) I_0(av) \quad (1.132)$$

图1.23显示了参数 a 分别为0,1,2,3,5时的情况。由这些曲线,可以得出以下结论:

1. 当 $a = 0$ 时,莱斯分布退化为瑞利分布。
2. 在 a 足够大也就是正弦信号幅度 A 大于 σ 到足够的程度时,包络分布近似于围绕 $v = a$ 的高斯分布, σ 为噪声 $n(t)$ 平均功率的平方根。

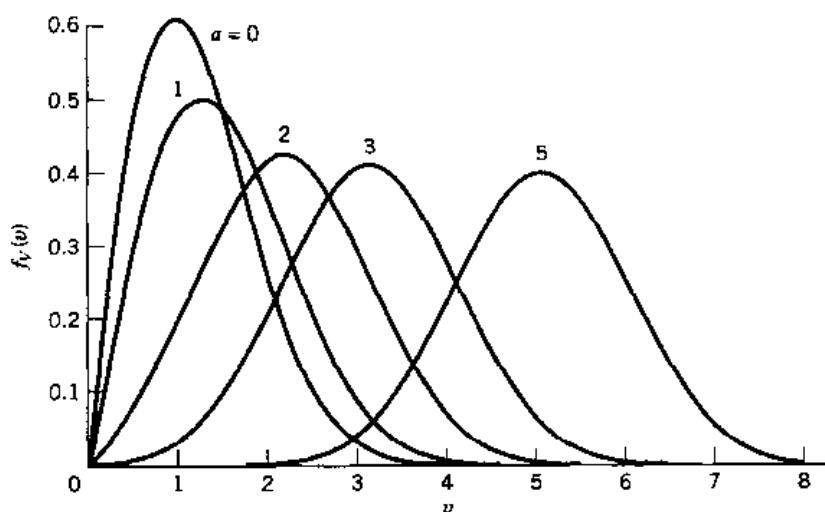


图 1.23 标准莱斯分布

1.14 计算机实验：平衰落信道

这一节将利用计算机来仿真以瑞利衰落为特征的多径信道，其模型源于无线通信以及电离层的长距离无线电传输。沿不同路径传输的信号相互干扰导致了衰落的产生，最后这些信号在相对不同的时刻到达接收机。最终的结果就是信号的幅度和相位都发生了较大的变化。在特定条件下，接收信号包络的统计时变特性可以近似地用瑞利分布来进行描述。

图 1.24 给出了一个多径模型。它包括在空间随机分布的大量散射物质，因此，单入射波束在接收天线处已经变为相当数量的散射波束。令被传输的信号等于 $A \cos(2\pi f_c t)$ ，同时假设所有的散射波束都以相同的速率传输。由于传输损耗和时延的不同，这些波束在幅度和相位上都有很大不同。第 k 个散射波束可写为 $A_k \cos(2\pi f_c t + \Theta_k)$ ，其中 A_k 和 Θ_k 为随时间缓慢变化的随机变量。 Θ_k 是相互独立的，并在 $[0, 2\pi]$ 内均匀分布。这里所描述的由多径引起的衰落被称为“平衰落”，因为传输信号的频谱特性在信道输出端被完整地保留了。但是，由于多径现象引起的信道增益的随机波动，输出信号的强度是随时间变化的。

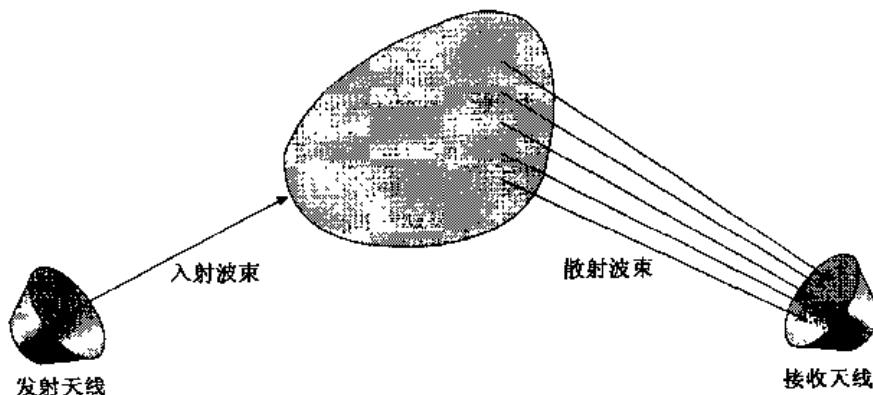


图 1.24 多径信道模型

假设有 N 条散射波束，则接收信号对应的随机过程为

$$X(t) = \sum_{k=1}^N A_k \cos(2\pi f_c t + \Theta_k) \quad (1.133)$$

其等价形式为

$$X(t) = X_I \cos(2\pi f_c t) - X_Q \sin(2\pi f_c t) \quad (1.134)$$

X_I 和 X_Q 的定义分别为

$$X_I = \sum_{k=1}^N A_k \cos \Theta_k \quad (1.135)$$

和

$$X_Q = \sum_{k=1}^N A_k \sin \Theta_k \quad (1.136)$$

为了表述方便且不失一般性，对于所有 k ，都假定 A_k 位于闭区间 $[-1, 1]$ 内。

1.14.1 实验 1：高斯分布

由中心极限定理，当散射波束 N 趋于无限时， X_I 和 X_Q 都将逼近高斯随机变量。为验证其有效性，取 $N = 10, 100, 1000$ 和 10000 来计算同相分量 X_I 和正交分量 X_Q 的概率分布。首先需要进行拟合性测量，以证明不同的 N 对应的抽样数据的概率分布测量值与理论上的高斯分布具有等效性。一种验证方法是利用概率分布的中心矩（最高可达四阶）来定义以下两个参量

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \quad (1.137)$$

和

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \quad (1.138)$$

其中， μ_2, μ_3, μ_4 分别为二、三、四阶中心矩。参量 β_1 和 β_2 的测量值用于检验概率分布的不对称性。测量值越接近理论分布所对应的 β_1 和 β_2 的数值，拟合性越好。对均值为 μ_X 、方差为 σ_X^2 的随机高斯变量 X ，有

$$\begin{aligned} \mu_2 &= \sigma_X^2 \\ \mu_3 &= 0 \\ \mu_4 &= 3\sigma_X^4 \end{aligned}$$

从而得到

$$\beta_1 = 0$$

和

$$\beta_2 = 3$$

表 1.1 提供了根据不同 N 值对应的同相分量 X_I 和正交分量 X_Q 计算所得的 β_1 和 β_2 的数值。将表中数据与高斯分布的相应数值进行比较, 可清楚看到随着散射波束数目 N 的增加, X_I 和 X_Q 都逼近零均值高斯分布, 这与中心极限定理的结论是一致的。

表 1.1 同相分量和正交分量的 β 值

(a) 实测量分布

		散 射 数			
		10	100	1000	10 000
同相分量 X_I	β_1	0.2443	0.0255	0.0065	0.0003
	β_2	2.1567	2.8759	2.8587	3.0075
正交分量 X_Q	β_1	0.0874	0.0017	0.0004	0.0000
	β_2	1.9621	2.7109	3.1663	3.0135

(b) 理论高斯分布

$$\beta_1 = 0$$

$$\beta_2 = 3$$

1.14.2 实验 2: 瑞利分布

在式(1.134)中, 随机过程 $X(t)$ 是用同相和正交分量来表示的。等效地, 可以用包络和相位来表示 $X(t)$:

$$X(t) = R \cos(2\pi f_c t + \Psi) \quad (1.139)$$

其中

$$R = \sqrt{X_I^2 + X_Q^2} \quad (1.140)$$

且

$$\Psi = \tan^{-1}\left(\frac{X_Q}{X_I}\right) \quad (1.141)$$

需注意实验中同相分量 X_I 、正交分量 X_Q 、包络分量 R 和相位分量 Ψ 都与时间无关。

由 1.12 节所述的理论可知, 如果 X_I 和 X_Q 随着 N 的增大而逼近高斯随机变量, 则 R 将逼近瑞利分布, Ψ 将逼近均匀分布。图 1.25 为 $N = 10000$ 时的实际概率密度函数, 图中计算了 100 个直方图和 100 个总体均值。图中还同时包含了理论曲线。从图中可以看出两条曲线的一致性很好, 这就证实了接收信号的包络 R 逼近瑞利分布的说法。

图 1.26 说明了瑞利衰落对接收信号 $x(t)$ 波形的影响, 其中 $x(t)$ 为 $X(t)$ 的样本函数, 发射信号 $X(t)$ 为单位幅度(即 $A = 1$)的 $f_c = 1$ MHz 的正弦信号。发射信号和相应的接收信号分别如图 1.26 的(a)和(b)部分所示。比较两个波形, 可以看到, 经过图 1.24 所示的多径信道的传输, 将得到幅度和相位随时间随机变化的接收信号。

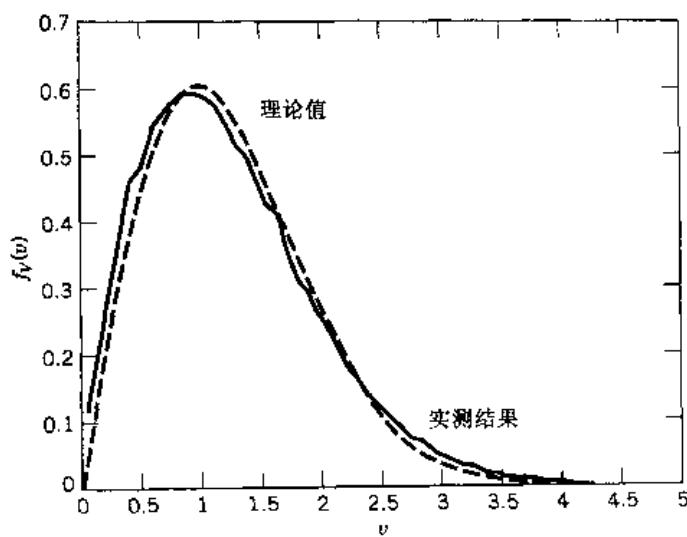
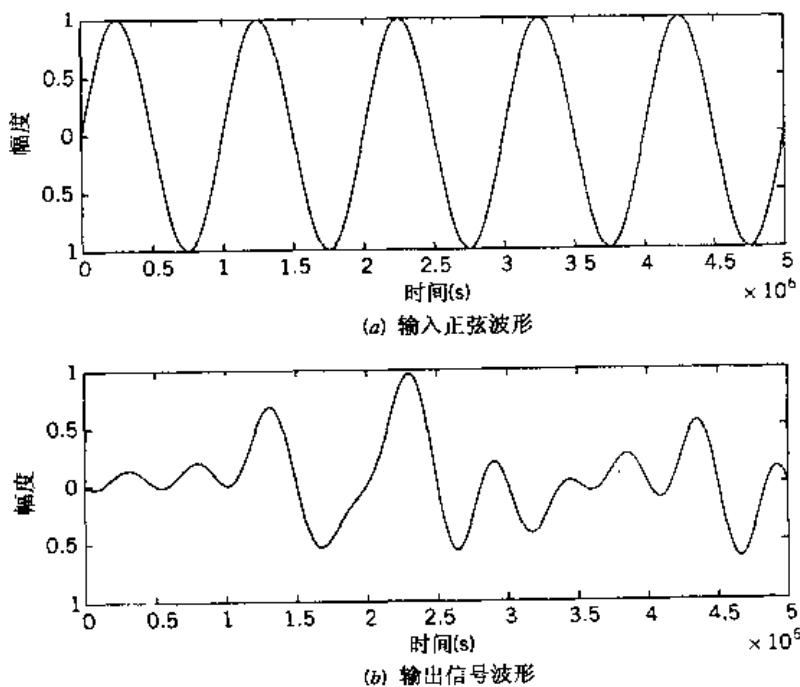
图 1.25 随机过程 $X(t)$ 包络的概率密度函数: 理论与实验的比较

图 1.26 瑞利衰落对正弦波形的影响

1.15 总结与讨论

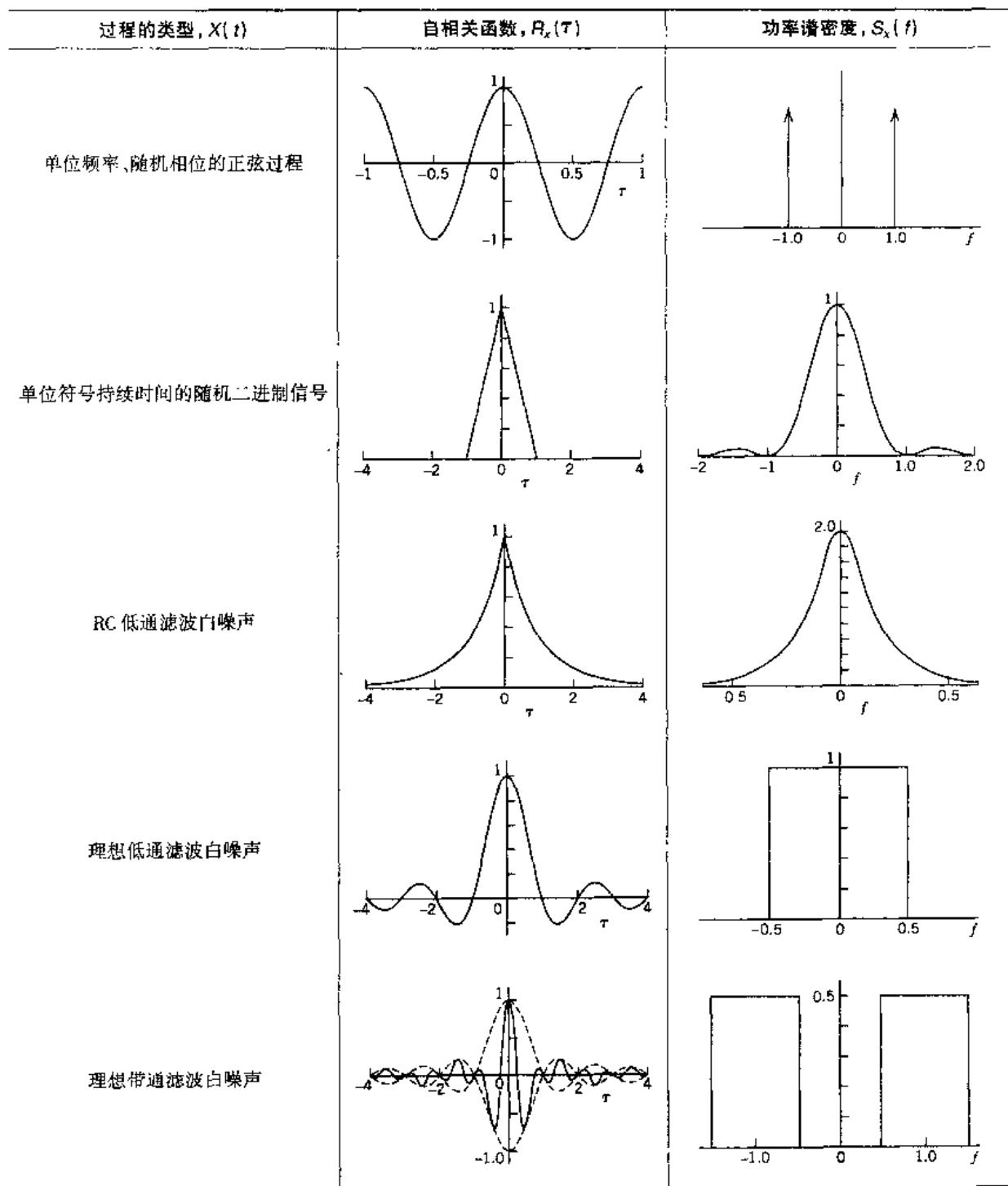
本章中许多内容都涉及到了特定类型的随机过程所具有的特性: 平稳性和各态历经性。广义地讲, 平稳的含义就是可以用两种集平均参数来对随机过程做出部分描述: (1) 均值与时间无关; (2) 自相关函数仅与观测的两个时间点的差值有关。^[10] 各态历经性使得我们可以用时间平均来代替参数的估计。时间平均可以用随机过程的样本函数(即单个实现)来计算。

另一个重要参数是功率谱密度。自相关函数和功率谱密度构成一对傅里叶变换对。可以

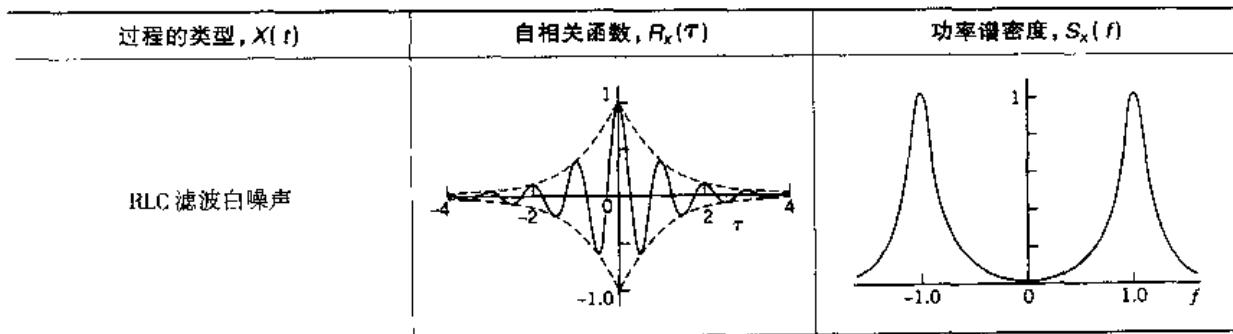
根据自相关函数来定义功率谱密度, 反之亦可。这种关系就被称为 Einstein-Wiener-Khintchine 关系。

表 1.2 给出了重要随机过程的自相关函数和功率谱密度的图形总结。假定表中所有过程都具有零均值和单位方差。读者从表中可以获得这样的感性认识:(1)随机过程的自相关函数和功率谱密度的相互作用;(2)线性滤波在形成自相关函数或白噪声功率谱密度中的作用。

表 1.2 零均值和单位方差随机过程的自相关函数和功率谱密度的图形总结



(续表)



本章后面部分讨论了窄带高斯噪声过程,这是理想通信接收机前端的一类滤波噪声。所谓高斯特性的含义是,在一些固定时刻观察滤波器输出所获得的随机变量具有高斯分布。噪声的窄带特性含义是它可以用同相和正交分量来表示。这两个分量都是低通的高斯过程,都具有零均值,并且方差与原窄带噪声相同。另外,一个窄带高斯噪声也可以用瑞利分布的包络和均匀分布的相位来表示。这两种表示法都有其特定的应用领域,本书后面的章节中会述及。

注释与参考

- [1] 对随机过程的详细论述可参看 Doob(1953), Loève(1963) 以及 Cramér 与 Leadbetter (1967) 的经典书籍。
- [2] 实践中往往会遇到另一种重要的随机过程,它的均值和自相关函数呈现出周期性:

$$\begin{aligned}\mu_X(t_1 + T) &= \mu_X(t_1) \\ R_X(t_1 + T, t_2 + T) &= R_X(t_1, t_2)\end{aligned}$$

上式对所有 t_1 和 t_2 都成立。满足这两个条件的随机过程 $X(t)$ 被称为循环平稳的(广义)。要将随机过程 $X(t)$ 构造成循环平稳的,就需要在随机过程的部分描述中加上一个新的自变量,即周期 T 。循环平稳过程的例子包括通过光栅扫描随机视频区域得到的电视信号,以及改变载波的幅度、相位或频率所得到的调制后的随机过程。对循环平稳过程的进一步讨论可参看 Franks(1969), pp. 204-214 以及 Gardner 与 Franks(1975)。

- [3] 通常,出于对 Norbert Wiener 和 A. I. Khintchine 的创始工作的认可,式(1.42) 和式(1.43) 在文献中常被称为 Wiener-Khintchine 关系;他们早期的论文参见 Wiener (1930) 和 Khintchine(1934)。爱因斯坦的一篇曾经被遗忘的关于时间序列分析的文章最终发表于 Basel 成立的瑞典物理协会 1914 年 2 月的会议集,这表明爱因斯坦先于 Wiener 和 Khintchine 很多年就已经讨论了时间序列自相关函数及其与频谱的关系。爱因斯坦论文的一篇英文翻译版在 1987 年 10 月被重新发表在 IEEE ASSP Magazine, vol. 4。这期特刊还包含 W. A. Gardner 和 A. M. Yaglom 的文章,文章对爱因斯坦的创始工作进行了阐述。
- [4] 如需进一步了解功率谱估计的情况,请分别参看 Blackman 和 Tukey(1958), Box 和 Jenkins(1976), Marple(1987), 以及 Kay(1988)。
- [5] 高斯分布和相关的高斯过程是因伟大的数学家高斯而得名。高斯在 18 岁时,就发

明了最小均方法来找出一系列测量值的最优值。后来，高斯将最小均方法用于行星数据的测量，其过程于 1809 年发表在他的“Theory of Motion of the Heavenly Bodies”一书中。由于观测中存在的误差，高斯提出了高斯分布。这个分布即是众所周知的正态分布。由于某些历史原因，数学家通常使用正态这个词，而工程师和物理学家则通常使用高斯这个词。

- [6] 对电噪声的详细论述可参看 Van der Ziel(1970) 和 Gupta(1977) 编辑的论文集。在 Helstrom(1990) 中有对散弹噪声的介绍性论述。更详细的论述可参看 Yue, Luganani 和 Rice(1978)。
热噪声最初是由 J. B. Johnson 在 1928 年用实验方法进行研究的，因此它有时也被称为 Johnson 噪声。Johnson 的实验结果被奈奎斯特用理论进行了证明。
- [7] 接收机的噪声性能也可以用所谓噪声系数来衡量。噪声系数和噪声温度的关系将在第 8 章具体说明。
- [8] 瑞利分布是根据英国物理学家 J. W. Strutt, Lord Rayleigh 的名字来命名的。
- [9] 莱斯分布是为了纪念 Stephen O. Rice，他在 1944 年和 1945 年分别发表的两篇文章中均有原创性的贡献，这两篇文章曾被 Wax 再版(1954)。
- [10] 本书所提到的通信系统的统计特性仅限于有关随机过程的前二阶矩、均值和自相关函数(或自协方差函数)。当一个随机过程通过非线性系统时，那些有价值的信息存在于输出的随机过程的高阶矩中。时域中用来表示高阶矩特性的参数被称为累积量，它们的多维傅里叶变换称为多谱。对高阶累积量和多谱及其估计的讨论可参看 Nikias 和 Raghubeer(1987)。

习题

平稳性和各态历经性

1.1 假设随机过程 $X(t)$ 定义如下：

$$X(t) = \sin(2\pi f_c t)$$

其中 f_c 为 $[0, W]$ 上均匀分布的随机变量。证明 $X(t)$ 是非平稳的。提示：检查 $f = W/4, W/2, W$ 的随机过程 $X(t)$ 的特定样本函数。

1.2 假设正弦过程

$$X(t) = A \cos(2\pi f_c t)$$

其中 f_c 为常量，幅度 A 是均匀分布的：

$$f_A(a) = \begin{cases} 1, & 0 \leq a \leq 1 \\ 0, & \text{其他} \end{cases}$$

试确定这个过程是否严平稳。

1.3 随机过程 $X(t)$ ：

$$X(t) = A \cos(2\pi f_c t)$$

其中, A 为均值为零、方差为 σ_A^2 的高斯分布随机变量。这个随机过程通过理想积分器, 产生输出

$$Y(t) = \int_0^t X(\tau) d\tau$$

- (a) 确定输出 $Y(t)$ 在特定时刻 t_k 的概率密度函数。
 - (b) 确定 $Y(t)$ 是否是平稳的。
 - (c) 确定 $Y(t)$ 是否是各态历经的。
- 1.4 令 X 和 Y 为统计独立的高斯分布随机变量, 都具有零均值和单位方差。定义高斯过程

$$Z(t) = X \cos(2\pi t) + Y \sin(2\pi t)$$

- (a) 确定分别在时刻 t_1 和 t_2 对 $Z(t)$ 观察得到的随机变量 $Z(t_1)$ 和 $Z(t_2)$ 的联合概率密度函数。
- (b) 随机过程 $Z(t)$ 是平稳的吗? 为什么?

相关函数和谱密度函数

- 1.5 证明随机过程 $X(t)$ 的自相关函数 $R_X(\tau)$ 具有下面两个特性:
- (a) 如果 $X(t)$ 包含直流分量, $R_X(\tau)$ 将包含一个等于 A^2 的常量。
 - (b) 如果 $X(t)$ 包含正弦分量, $R_X(\tau)$ 将包含同样频率的正弦分量。
- 1.6 图 P1.6 中幅度常数为 A 、周期为 T_0 、时延为 t_d 的方波 $x(t)$ 为随机过程 $X(t)$ 的样本函数, 时延是随机的, 其概率密度函数为

$$f_{T_d}(t_d) = \begin{cases} \frac{1}{T_0}, & -\frac{1}{2} T_0 \leq t_d \leq \frac{1}{2} T_0 \\ 0, & \text{其他} \end{cases}$$

- (a) 确定在 t_k 时刻观察随机过程 $X(t)$ 得到的随机变量 $X(t_k)$ 的概率密度函数。
- (b) 采用集平均来确定 $X(t)$ 的均值和自相关函数。
- (c) 采用时间平均来确定 $X(t)$ 的均值和自相关函数。
- (d) 确定 $X(t)$ 的平稳性。在什么意义上, 它是各态历经的?

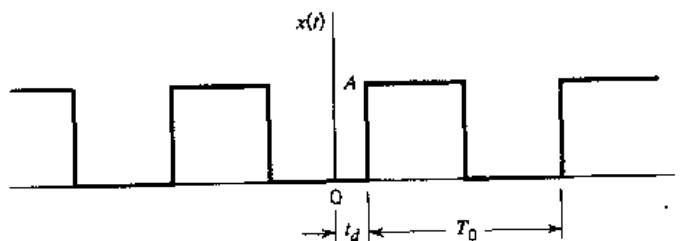


图 P1.6

- 1.7 二进制信号由符号 1 和 0 组成的随机序列构成。符号 1 由一幅度为 A 伏的脉冲表示, 而符号 0 由零电平表示。其他所有参数与例 1.3 中相同。试证明该新的随机二进制波形 $X(t)$:

(a) 自相关函数为

$$R_X(\tau) = \begin{cases} \frac{A^2}{4} + \frac{A^2}{4} \left(1 - \frac{|\tau|}{T}\right), & |\tau| < T \\ \frac{A^2}{4}, & |\tau| \geq T \end{cases}$$

(b) 功率谱密度为

$$S_X(f) = \frac{A^2}{4} \delta(f) + \frac{A^2 T}{4} \operatorname{sinc}^2(fT)$$

在该二进制波形中, 直流分量功率的比例是多少?

- 1.8 随机过程 $Y(t)$ 包含 $\sqrt{3}/2$ V 的直流分量, 周期性分量 $g(t)$ 和随机分量 $X(t)$ 。 $Y(t)$ 的自相关函数见图 P1.8。试确定
 (a) 周期性分量 $g(t)$ 的平均功率是多少?
 (b) 随机分量 $X(t)$ 的平均功率是多少?

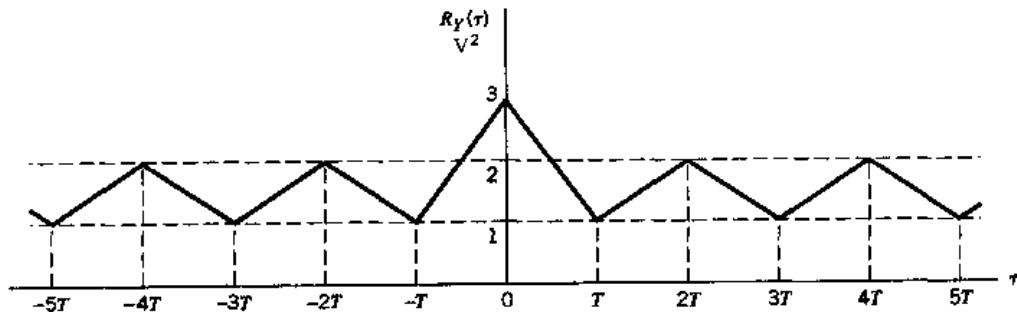


图 P1.8

- 1.9 假设一对平稳随机过程 $X(t)$ 和 $Y(t)$ 。证明其互相关函数 $R_{XY}(\tau)$ 和 $R_{YX}(\tau)$ 具有如下特性:

$$(a) R_{XY}(\tau) = R_{YX}(-\tau)$$

$$(b) |R_{XY}(\tau)| \leq \frac{1}{2} [R_X(0) + R_Y(0)]$$

其中, $R_X(\tau)$ 和 $R_Y(\tau)$ 分别为 $X(t)$ 和 $Y(t)$ 的自相关函数。

- 1.10 如图 P1.10 所示的两个串联线性滤波器, $X(t)$ 是自相关函数为 $R_X(\tau)$ 的平稳过程。第一个滤波器的输出为 $V(t)$, 第二个滤波器的输出为 $Y(t)$ 。

(a) 求 $Y(t)$ 的自相关函数。

(b) 求 $V(t)$ 和 $Y(t)$ 的互相关函数 $R_{VY}(\tau)$ 。

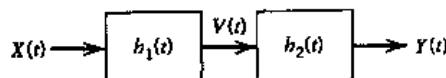


图 P1.10

- 1.11 平稳过程 $X(t)$ 通过一个冲激响应为 $b(t)$ 的线性时不变滤波器, 产生输出 $Y(t)$ 。

(a) 证明输出 $Y(t)$ 和输入 $X(t)$ 的互相关函数 $R_{YX}(\tau)$ 等于冲激响应 $b(\tau)$ 与输入信号的自相关函数 $R_X(\tau)$ 的卷积, 即

$$R_{YX}(\tau) = \int_{-\infty}^{\infty} h(u) R_X(\tau - u) du$$

证明另一个互相关函数为

$$R_{XY}(\tau) = \int_{-\infty}^{\infty} h(-u) R_X(\tau - u) du$$

(b) 求互功率谱密度 $S_{YX}(f)$ 和 $S_{XY}(f)$ 。

(c) 假定 $X(t)$ 是均值为零, 功率谱密度为 $N_0/2$ 的白噪声过程, 证明

$$R_{YX}(\tau) = \frac{N_0}{2} h(\tau)$$

并说明这个结果的实际意义。

1.12 一个随机过程 $X(t)$ 的功率谱密度如图 P1.12 所示, 它包括 $f=0$ 处的 delta 函数和一个三角函数分量。

(a) 求 $X(t)$ 的自相关函数 $R_X(\tau)$ 并画图。

(b) $X(t)$ 中的直流功率是多少?

(c) $X(t)$ 中的交流功率是多少?

(d) 用什么样的抽样速率才能得到 $X(t)$ 的不相关抽样? 这些抽样点统计独立吗?

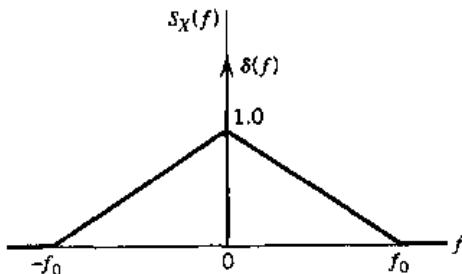


图 P1.12

1.13 - 对噪声过程 $n_1(t)$ 和 $n_2(t)$ 的关系为

$$n_2(t) = n_1(t) \cos(2\pi f_c t + \theta) - n_1(t) \sin(2\pi f_c t + \theta)$$

其中 f_c 为常量, θ 为随机变量 Θ 的值, 其概率密度函数为

$$f_\Theta(\theta) = \begin{cases} \frac{1}{2\pi}, & 0 \leq \theta \leq 2\pi \\ 0, & \text{其他} \end{cases}$$

假设噪声过程 $n_1(t)$ 是平稳的, 它的功率谱密度如图 P1.13 所示。求 $n_2(t)$ 的功率谱密度并画图。

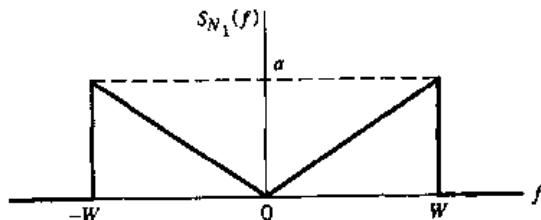


图 P1.13

1.14 一个随机电报信号 $X(t)$, 其自相关函数为

$$R_X(\tau) = \exp(-2\nu|\tau|)$$

其中 ν 是常量。 $X(t)$ 通过图 P1.14 所示的 RC 低通滤波器, 试求滤波器输出的功率谱密度和自相关函数。

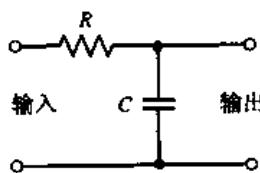


图 P1.14

1.15 一个连续积分器

$$y(t) = \int_{t-T}^t x(\tau) d\tau$$

其中, $x(t)$ 为输入, $y(t)$ 为输出, T 为积分周期。 $x(t)$ 和 $y(t)$ 分别为平稳过程 $X(t)$ 和 $Y(t)$ 的样本函数。证明积分器输出的功率谱密度与输入的功率谱密度关系为

$$S_Y(f) = T^2 \operatorname{sinc}^2(fT) S_X(f)$$

1.16 一个零均值平稳过程通过一个线性滤波器, 其冲激响应由截短指数函数定义:

$$h(t) = \begin{cases} ae^{-at}, & 0 \leq t \leq T \\ 0, & \text{其他} \end{cases}$$

证明滤波器输出过程 $Y(t)$ 的功率谱密度为

$$S_Y(f) = \frac{a^2}{a^2 + 4\pi^2 f^2} (1 - 2 \exp(-aT) \cos(2\pi fT) + \exp(-2aT)) S_X(f)$$

其中 $S_X(f)$ 是滤波器输入的功率谱密度。

1.17 振荡器输出过程描述为

$$X(t) = A \cos(2\pi ft - \Theta)$$

其中 A 为常量, f 和 Θ 为相互独立的随机变量。 Θ 的概率密度函数为

$$f_\Theta(\theta) = \begin{cases} \frac{1}{2\pi}, & 0 \leq \theta \leq 2\pi \\ 0, & \text{其他} \end{cases}$$

根据频率 f 的概率密度函数, 求出 $X(t)$ 的功率谱密度。如果 f 是常量, 功率谱密度将发生什么变化?

高斯过程

1.18 一个均值为零的平稳高斯过程 $X(t)$ 的功率谱密度为 $S_X(f)$, 求在时刻 t_k 观察 $X(t)$ 所得的随机变量的概率密度函数。

1.19 零均值且方差为 σ_x^2 的高斯过程 $X(t)$ 通过具有图 P1.19 中所示输入-输出关系的全

波整流器。证明在时刻 t_k 观察整流器的输出 $Y(t)$ 所得到的随机变量 $Y(t_k)$ 的概率密度函数为

$$f_{Y(t_k)}(y) = \begin{cases} \sqrt{\frac{2}{\pi}} \frac{1}{\sigma_X} \exp\left(-\frac{y^2}{2\sigma_X^2}\right), & y \geq 0 \\ 0, & y < 0 \end{cases}$$

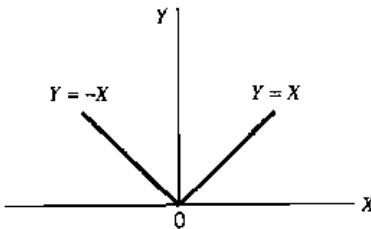


图 P1.19

- 1.20 令 $X(t)$ 为零均值平稳高斯过程, 其自相关函数为 $R_X(\tau)$ 。这个随机过程通过一个平方律器件, 对应输入-输出关系为

$$Y(t) = X^2(t)$$

其中 $Y(t)$ 为输出。

- (a) 证明 $Y(t)$ 的均值为 $R_X(0)$ 。
- (b) 证明 $Y(t)$ 的自协方函数为 $2R_X^2(\tau)$ 。

- 1.21 一个均值为 μ_X 、方差为 σ_X^2 的平稳高斯过程 $X(t)$ 通过冲激响应分别为 $b_1(t)$ 和 $b_2(t)$ 的两个线性滤波器, 产生两个输出 $Y(t)$ 和 $Z(t)$, 如图 P1.21 所示。

- (a) 求随机变量 $Y(t_1)$ 和 $Z(t_2)$ 的联合概率密度函数。
- (b) 为保证 $Y(t_1)$ 和 $Z(t_2)$ 统计独立, 什么条件是充分必要的?

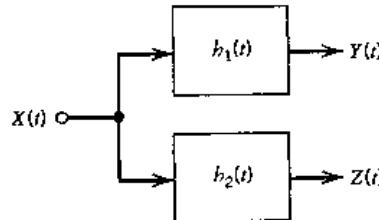


图 P1.21

- 1.22 一个零均值且功率谱密度为 $S_X(f)$ 的平稳高斯过程 $X(t)$ 通过如图 P1.22 所示的冲激响应为 $b(t)$ 的线性滤波器。在时刻 T 抽取滤波器输出一个样本 Y 。

- (a) 求 Y 的方差和均值。
- (b) Y 的概率密度函数。

噪声

- 1.23 假设一个零均值且功率谱密度为 $N_0/2$ 的高斯白噪声过程, 使其通过如图 P1.23 所示的高通 RL 滤波器。

(a) 求滤波器输出的自相关函数和功率谱密度。

(b) 输出的均值和方差是多少?

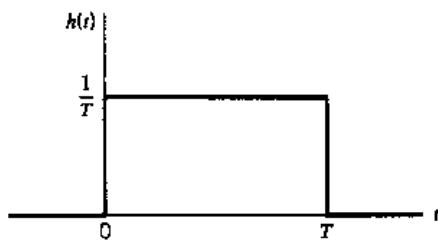


图 P1.22

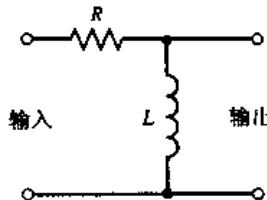


图 P1.23

1.24 一个功率谱密度为 $N_0/2$ 的白噪声过程 $w(t)$, 通过具有如下幅度响应的 n 阶巴特沃兹低通滤波器

$$|H(f)| = \frac{1}{[1 + (f/f_0)^{2n}]^{1/2}}$$

(a) 求低通滤波器的噪声等效带宽(参看附录 2 中的等效噪声带宽定义)。

(b) n 趋向于无穷时, 噪声等效带宽的极限值是多少?

1.25 由式(1.86)定义的散弹噪声 $X(t)$ 是平稳的。为什么?

1.26 一个零均值且功率谱密度为 $N_0/2$ 的高斯白噪声通过如图 P1.26(a)所示的滤波器组, 这两个滤波器的频率响应如图 P1.26(b)所示。低通滤波器输出端噪声由 $n(t)$ 表示。

(a) 求 $n(t)$ 的功率谱密度和自相关函数。

(b) 求 $n(t)$ 的均值和方差。

(c) 以什么样的速率对 $n(t)$ 进行抽样, 才能使得抽样结果是不相关的?

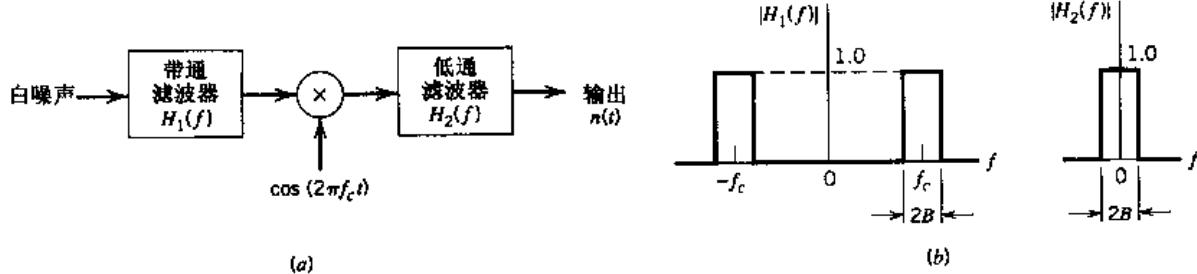


图 P1.26

1.27 $X(t)$ 是均值为零、自相关函数为 $R_X(\tau)$ 、功率谱密度为 $S_X(f)$ 的平稳过程。试找出一个冲激响应为 $b(t)$ 的线性滤波器, 使得当输入为功率谱密度是 $N_0/2$ 的白噪声时, 滤波器输出具有与 $X(t)$ 相同的统计特性。

(a) 指出为了达到此目的, 冲激响应 $b(t)$ 必须满足的条件。

(b) 其频率响应 $H(f)$ 的相应条件是什么?

窄带噪声

1.28 在如图 P1.19(a)所示的噪声分析仪中,低通滤波器是理想的,其带宽等于输入窄带噪声 $n(t)$ 带宽的一半。采用这种方案,推导以下结果:

- 等式(1.101),由 $n(t)$ 的功率谱密度,推出同相噪声分量 $n_I(t)$ 和正交噪声分量 $n_Q(t)$ 的功率谱密度。
- 等式(1.102), $n_I(t)$ 和 $n_Q(t)$ 的互功率谱密度的定义。

1.29 假定窄带噪声 $n(t)$ 是高斯型的,它的功率谱密度 $S_N(f)$ 是关于中心频率 f_c 对称的。

证明 $n(t)$ 的同相和正交分量是统计独立的。

1.30 窄带噪声 $n(t)$ 的功率谱密度如图 P1.30 所示,载频为 5 Hz。

- 求出 $n(t)$ 的同相和正交分量的功率谱密度。
- 求出它们的互功率谱密度。

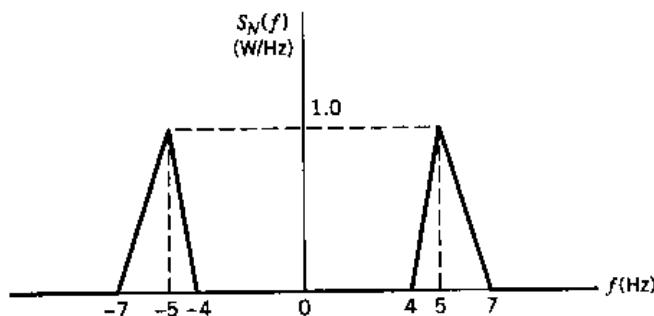


图 P1.30

1.31 假设有均值为零、功率谱密度为 $S_N(f)$ 的如图 P1.31 所示的高斯白噪声 $n(t)$ 。试求

- $n(t)$ 包络的概率密度函数。
- 包络的均值和方差。

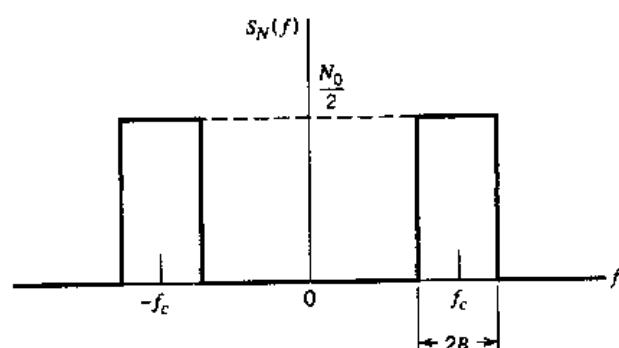


图 P1.31

计算机实验

1.32 实验中将研究随机过程 $X(t)$ 的统计特性。 $X(t)$ 的表达式为

$$X(t) = A \cos(2\pi f_c t + \Theta) + W(t)$$

其中,正弦分量的相位 Θ 是均匀分布在 $[-\pi, +\pi]$ 上的随机变量。 $W(t)$ 是零均值且功率谱密度为 $N_0/2$ 的高斯白噪声分量。 $X(t)$ 的两个分量是统计独立的,因此, $X(t)$ 的自相关函数为

$$R_X(\tau) = \frac{A^2}{2} \cos(2\pi f_c \tau) + \frac{N_0}{2} \delta(\tau)$$

这个等式表明当 $|\tau| > 0$ 时,自相关函数 $R_X(\tau)$ 有与 $X(t)$ 信号相同的正弦波形。

实验的内容是采用以下两种方法计算 $R_X(\tau)$:

- (a) 集平均。从 $X(t)$ 中随机选取 $M = 50$ 个实现。然后,对某个特定时间 t 计算乘积 $x(t + \tau)x(t)$, $x(t)$ 为 $X(t)$ 的一个实现。对 M 次实现重复计算 $x(t + \tau)x(t)$, 然后计算这 M 个结果的均值。对不同 τ , 重复这一系列计算。
- (b) 时间平均。计算时间平均自相关函数

$$R_x(\tau, T) = \frac{1}{2T} \int_{-T}^{T} x(t + \tau)x(t) dt$$

其中, $x(t)$ 为 $X(t)$ 的一个实现, $2T$ 为总观察间隔。使用傅里叶变换对

$$R_x(\tau, T) \rightleftharpoons \frac{1}{2T} |X_T(f)|^2$$

其中, $|X_T(f)|^2/2T$ 是过程 $X(t)$ 的周期图。特别地,计算窗口函数为

$$x_T(t) = \begin{cases} x(t), & -T \leq t \leq T \\ 0, & \text{其他} \end{cases}$$

的傅里叶变换 $X_T(f)$ 。最后,计算 $|X_T(f)|^2/2T$ 的傅里叶反变换。

最后,比较 $R_X(\tau)$ 的两种计算结果。

1.33 继续 1.14 节描述的多径信道的研究。特别地,考虑接收信号包含视距分量的情况,即

$$X(t) = \sum_{k=1}^N A_k \cos(2\pi f_c t + \Theta_k) + a \cos(2\pi f_c t)$$

其中, $a \cos(2\pi f_c t)$ 是直接接收分量。按 1.14 节所提供的资料,分别计算 $N = 10000$ 和 $a = 0, 1, 2, 3, 5$ 时 $X(t)$ 的包络。将计算结果与 1.13 节所研究的莱斯分布进行比较。

第2章 连续波调制

本章研究连续波调制,它是模拟通信系统的基本组成部分。本章由相关的两部分组成,第一部分别在时域和频域下讨论以下两种基本的连续波调制:

- 幅度调制,在幅度调制中,正弦载波的幅度随输入信号的变化而变化。
- 角度调制,在角度调制中,正弦载波的瞬时角频率或相位随输入信号的变化而变化。

第二部分讨论采用这些调制方法时,信道噪声对接收端造成的影响。

本章重点阐述了各种连续波调制方法的优缺点。

2.1 引言

通信系统的作用是通过发端与收端之间的通信信道传递承载信息的信号。承载信息的信号也称为基带信号。基带是指信源发出的原始信号的频率范围。通信信道将基带频率范围内的信号搬移到适合传送的其他频率上,并在到达接收端后对信号进行相应的变换以恢复出原始频率。例如,无线系统的工作频率为大于等于 30 kHz,而基带信号通常包含音频范围内的频率,因此必须进行某种频带变换以保证系统正常工作。信号的频谱搬移通过调制完成,调制是一个载波的某些特征量随着调制信号(信号)的变化而变化的过程。一种普遍采用的载波形式是正弦信号,我们将讨论采用正弦信号的连续波调制^[1]过程。基带信号称为调制信号,而调制过程的结果则称为已调信号,调制在通信系统的发射端进行。通常在系统的接收端恢复出原始的基带信号,这个过程可通过解调完成,调制与解调正好是互逆的两个过程。

在基本的信号处理过程中,模拟通信系统的发射机由调制器组成,接收机由解调器组成,如图 2.1 所示。

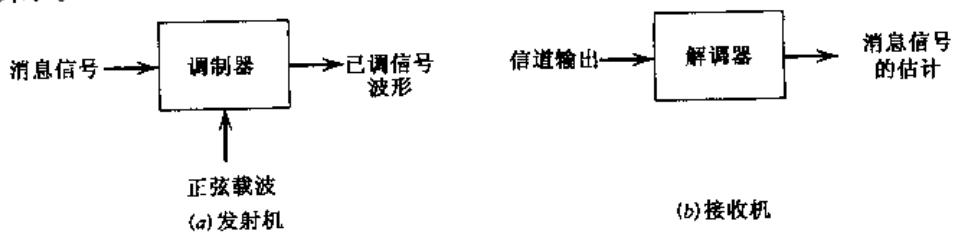


图 2.1 连续波调制系统的组成部分

接收机的输入除了来自发射机的信号外,还包含了信道噪声。信道噪声给接收机带来的性能损失大小由采用的调制类型决定。

本章研究的两种连续波(CW)调制系统分别称为幅度调制(AM)和角度调制。在幅度调制中,正弦载波的幅度随基带信号的变化而变化。在角度调制中,正弦载波的角度随基带信号变化而变化。图 2.2 给出了当调制信号为正弦信号时,幅度调制信号的波形和角度调制信号的波形。图中(a)和(b)分别为正弦载波和调制信号。(c)和(d)分别为相应的 AM 和 FM 信号波形。FM 即频率调制,是角度调制的一种形式。由该图可清楚地看出幅度调制和角度调制之间的基本差异,这些差异将在后面讨论。

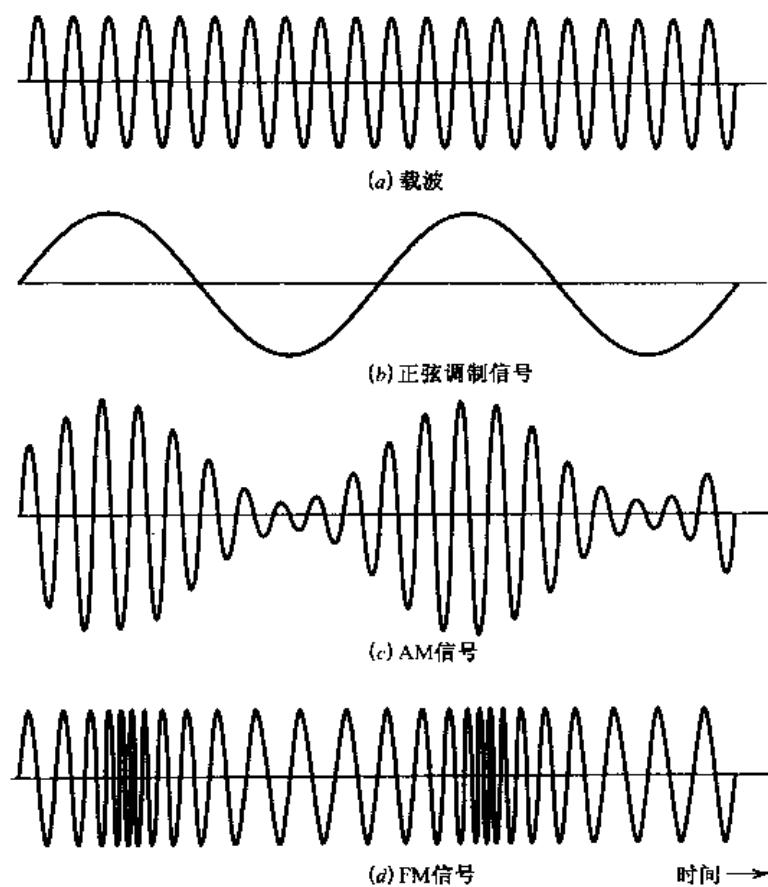


图 2.2 由单音频产生的 AM 和 FM 信号

2.2 幅度调制

考虑一个定义如下的正弦载波 $c(t)$:

$$c(t) = A_c \cos(2\pi f_c t) \quad (2.1)$$

式中, A_c 为载波幅度, f_c 为载波频率。为了使问题简化, 假定式(2.1)中载波的初始相位为 0, 这不会影响计算结果和最终的结论。令 $m(t)$ 表示携带消息的基带信号。产生载波 $c(t)$ 和 $m(t)$ 的信源是相互独立的。幅度调制(AM)是这样一个过程: 载波 $c(t)$ 的幅度围绕某个均值随基带信号 $m(t)$ 呈线性变化。因此 AM 信号最普遍的时域表达式为

$$s(t) = A_c [1 + k_a m(t)] \cos(2\pi f_c t) \quad (2.2)$$

其中, k_a 为常量, 称为产生已调信号 $s(t)$ 的调制器的调幅灵敏度。若载波幅度 A_c 和消息信号 $m(t)$ 的单位为 V(伏), 那么 k_a 的单位为 1/V。

图 2.3(a)为基带信号 $m(t)$ 的波形, 图 2.3(b)和 2.3(c)为调幅灵敏度 k_a 取两个不同值时对应的 AM 信号波形 $s(t)$, 其中载波幅度 $A_c = 1$ V。当满足以下两个条件时, $s(t)$ 的包络与基带信号 $m(t)$ 的形状相同:

1. $k_a m(t)$ 的幅值总小于 1, 即

$$|k_a m(t)| < 1 \quad \text{对于所有 } t \quad (2.3)$$

图 2.3(b) 满足此条件。这个条件保证了函数 $1 + k_a m(t)$ 总是取正值, 由于包络必须为正值函数, 因此可将式(2.2)中 AM 信号 $s(t)$ 的包络值用 $A_c[1 + k_a m(t)]$ 表示。当调制器的调幅灵敏度 k_a 足够大, 使得对所有的 t 值, 信号 $|k_a m(t)| > 1$, 此时载波变成了过调幅, 导致当因子 $1 + k_a m(t)$ 过零点时, 载波相位发生反转, 如图 2.3 所示, 此时已调信号出现了包络失真。为了避免过调幅, 对所有的时刻点, AM 信号的包络和调制信号之间都必须保持一个一一对应的关系——这个特点的作用将在后面得到体现。 $k_a m(t)$ 取绝对值后, 将其最大值乘以 100, 乘积称为调制百分比。

2. 载波频率 f_c 远远高于消息信号 $m(t)$ 的最高频率分量 W 。即

$$f_c \gg W \quad (2.4)$$

此处的 W 称为消息带宽。如果不满足式(2.4), 则无法看到(因而也无法检测到)正确的包络。

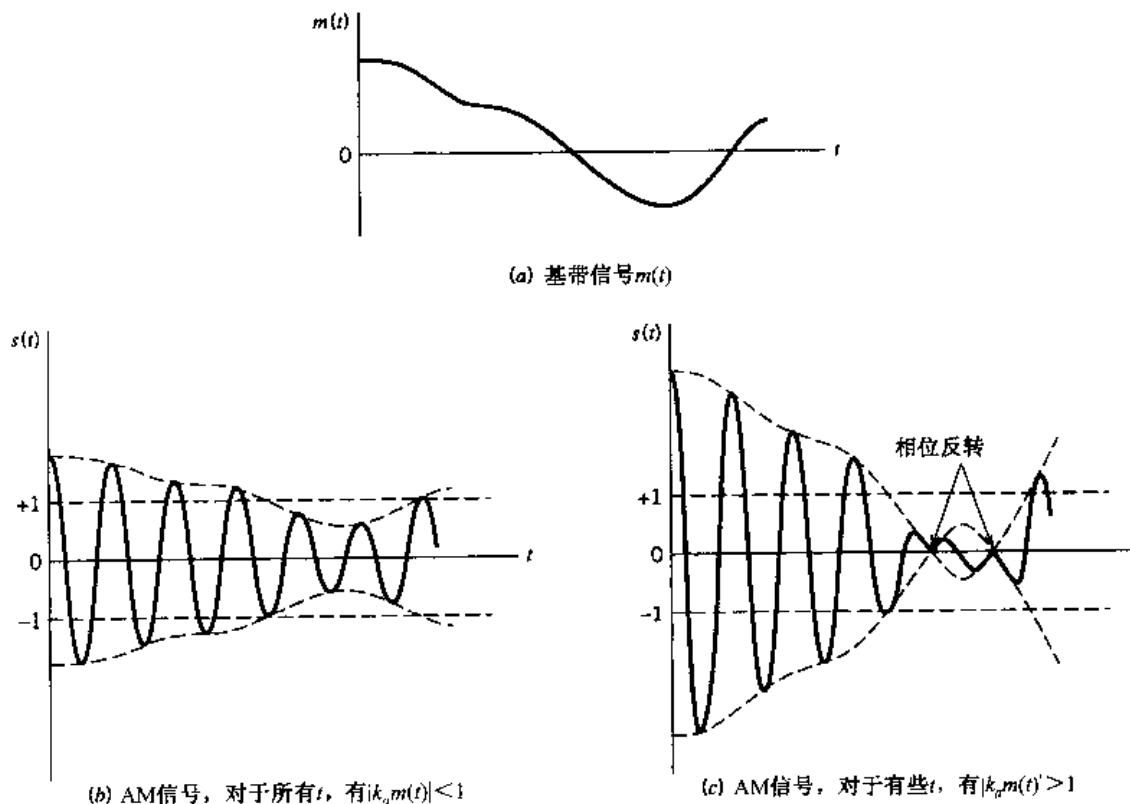


图 2.3 幅度调制过程

由式(2.2)得出 AM 信号 $s(t)$ 的傅里叶变换为

$$S(f) = \frac{A_c}{2} [\delta(f - f_c) + \delta(f + f_c)] + \frac{k_a A_c}{2} [M(f - f_c) + M(f + f_c)] \quad (2.5)$$

假定基带信号 $m(t)$ 带限于 $-W \leq f \leq W$, 如图 2.4(a), 该图只是示意图。由式(2.5), AM 信号

的频谱 $S(f)$ 在 $f_c > W$ 时可由图 2.4(b) 表示。频谱包括两个冲激函数, 冲激分别出现在 $\pm f_c$ 处, 大小均为 $A_c/2$ 。基带频谱在频率轴上分别平移了 $\pm f_c$, 幅值为 $k_a A_c/2$, 在图 2.4(b) 的频谱图中, 应当注意以下几点:

1. 当满足条件 $f_c > W$ 时, 由于调制过程的原因, 消息信号 $m(t)$ 在 $-W$ 到 0 之间的负频率频谱在正频率部分完全可见(可测), 其中“负”频率的概念很重要。
2. 对于正频率, AM 信号频谱中高于载频 f_c 的频率部分称为上边带, 而低于 f_c 的对称的频率部分称为下边带。对于负频率, 频谱中低于 $-f_c$ 的频率为上边带, 高于 $-f_c$ 的频率部分为下边带。约束条件 $f_c > W$ 确保了边带不会发生重叠。
3. 对于正频率, AM 信号的最高频率为 $f_c + W$, 最低频率为 $f_c - W$ 。两个频率的差值为 AM 信号的传输带宽 B_T , B_T 正好是信号带宽 W 的两倍, 即

$$B_T = 2W \quad (2.6)$$

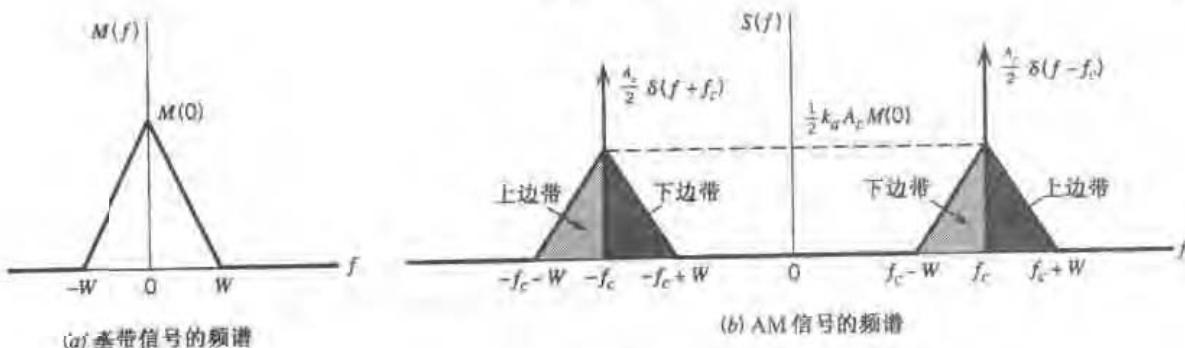


图 2.4 基带信号和 AM 信号的频谱

2.2.1 幅度调制的优缺点

幅度调制是最原始的一种调制方法。它最大的优点是实现方式简单。

- 在发射机中, 通过非线性装置就可完成幅度调制。例如, 习题 2.3 中的开关调制器, 将消息信号和载波之和送至二极管, 只要载波幅度足够大, 就可使二极管导通。对电阻两端的电压的傅里叶分析揭示了 AM 分量的产生机理, 通过带通滤波器可得到 AM 分量。
- 在接收机中, 采用非线性装置可完成解调。例如采用简单且高效的电路, 如习题 2.5 中讨论的包络检波器完成解调。该电路由一个二极管、一个电容和一个电阻的并联体串联而成。在许多商用的 AM 无线接收机中都存在这种电路形式。假定载频足够高且调制百分比低于 100%, 那么如果将电阻两侧作为解调器的输出端, 得到的波形与输入 AM 信号的包络就几乎是相同的, 因此称为“包络检波”。

发射功率和信道带宽是两个最应该有效利用的主要通信资源, 因此幅度调制的标准公式(2.2)有两个很大的缺陷:

1. 幅度调制浪费功率。载波 $c(t)$ 完全独立于携带信息的信号 $m(t)$, 因此载波的发送会造成功率的浪费, 这说明在幅度调制中整体的发射功率中只有一部分是由 $m(t)$ 决定的。
2. 幅度调制浪费带宽。由于 AM 信号的上边带和下边带关于载频对称, 因此上下边带之

间具有独特的关联性,只要给出了任一边带的幅度谱和相位谱,就可以得到另一个边带的相应参数。从信息传送的角度看,只传送一个边带就足够了,因此信道提供的带宽与基带信号的带宽相同即可。而幅度调制需要的传送带宽等于信号宽度的两倍,这显然是浪费带宽的。

为了克服这些缺陷,必须对幅度调制进行改进,如抑制载波并且改变AM信号的边带。这些改进势必会增加系统的复杂度。事实上,这样做是以增大系统的复杂度为代价提高了通信资源的利用率。这种方法的原理是线性调制(将在下一节进行讨论)。从严格的意义上说,标准幅度调制并不属于线性调制,因为标准幅度调制中出现了载波。

2.3 线性调制方案

线性调制最普遍的形式为

$$s(t) = s_I(t) \cos(2\pi f_c t) + s_Q(t) \sin(2\pi f_c t) \quad (2.7)$$

其中, $s_I(t)$ 为 $s(t)$ (已调信号)的同相分量, $s_Q(t)$ 为 $s(t)$ 的正交分量,式(2.7)是典型的窄带信号表达式,附录2中详细讨论了窄带信号。在线性调制中, $s_I(t)$ 和 $s_Q(t)$ 均为低通信号且与消息信号 $m(t)$ 线性相关。

根据 $s(t)$ 的两个分量的定义方法,可将含有一个消息信号的线性调制分为三种类型:

1. 抑制载波的双边带(DSB-SC)调制,只传送上边带和下边带。
2. 单边带(SSB)调制,只传送一个边带(上边带或下边带)。
3. 残留边带(VSB)调制,只传送某一边带的残余和另一边带相应修改的部分。

表2.1是对这三种线性调制类型的总结,需要注意以下两点:

1. 同相分量 $s_I(t)$ 只依赖于消息信号 $m(t)$ 。
2. 正交分量 $s_Q(t)$ 由 $m(t)$ 滤波得到。已调信号 $s(t)$ 频谱的改变仅仅是因为 $s_Q(t)$ 的存在。

更明确地说,正交分量(如果存在)的作用只是为了干预同相分量,目的是减小或排除已调信号 $s(t)$ 某一个边带的功率。正交分量的具体形式取决于正交分量的定义。

表2.1 线性调制的各种形式

调制类型	同相分量 $s_I(t)$	正交分量 $s_Q(t)$	说明
DSB-SC	$m(t)$	0	$m(t)$ =消息信号
SSB: [*]			
(a) 传送上边带	$\frac{1}{2}m(t)$	$\frac{1}{2}\dot{m}(t)$	$\dot{m}(t)=m(t)$ 的希尔伯特变换
(b) 传送下边带	$\frac{1}{2}m(t)$	$-\frac{1}{2}\dot{m}(t)$	
VSB:			
(a) 传送下边带的残留部分	$\frac{1}{2}m(t)$	$\frac{1}{2}m'(t)$	输入为 $m(t)$ 、频率响应为 $H_Q(f)$ 的 滤波器输出信号 $m'(t)$ 。 $H_Q(f)$ 的定义见式(2.16)。
(b) 传送上边带的残留部分	$\frac{1}{2}m(t)$	$-\frac{1}{2}m'(t)$	

* 对于SSB调制的数学描述见习题2.16。

2.3.1 DSB-SC 调制

这种线性调制通过使用乘积调制器, 将消息信号 $m(t)$ 与载波 $A_c \cos(2\pi f_c t)$ 相乘, 如图 2.5(a) 所示。可表示为

$$s(t) = A_c m(t) \cos(2\pi f_c t) \quad (2.8)$$

图 2.5(b) 的任意消息波形对应的已调信号波形 $s(t)$ 如图 2.5(c) 所示。当消息信号 $m(t)$ 过零点时, 已调信号 $s(t)$ 发生相位反转。因此, DSB-SC 已调信号的包络与消息信号不完全相同, 这一点与调制百分比小于 100% 的 AM 信号不同。

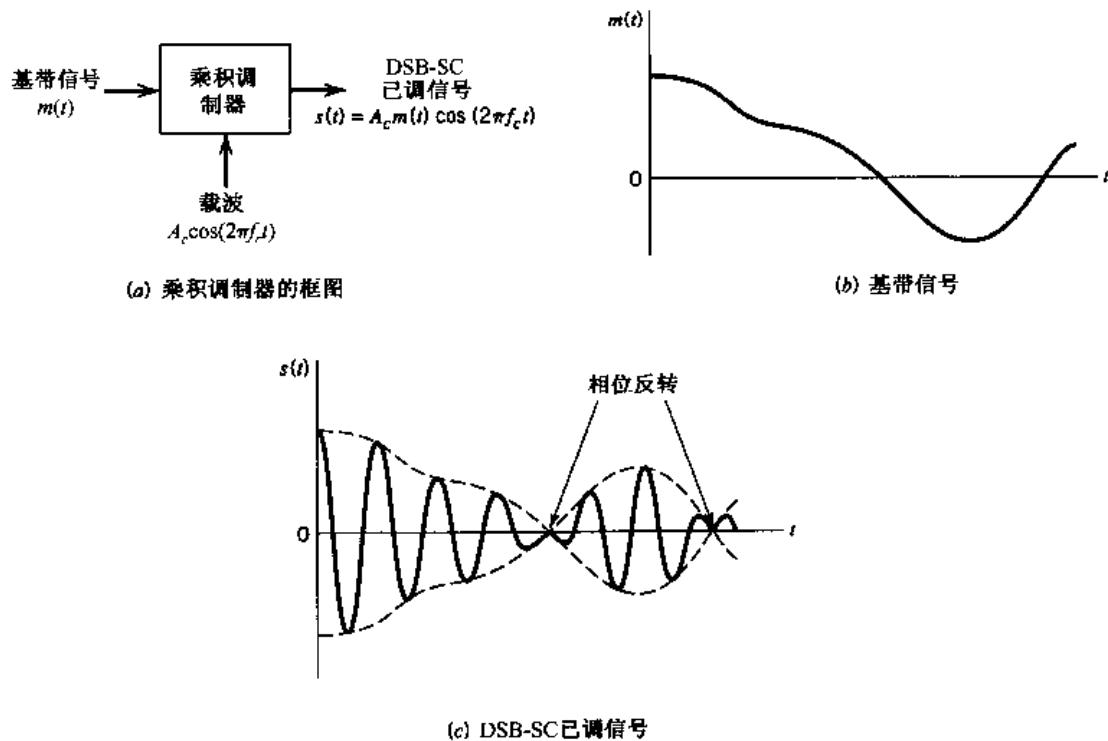


图 2.5 抑制载波的双边带(DSB-SC)调制

由式(2.8), 得到 $s(t)$ 的傅里叶变换为

$$S(f) = \frac{1}{2} A_c [M(f - f_c) + M(f + f_c)] \quad (2.9)$$

当基带信号如图 2.6(a) 那样被限制在间隔 $-W \leq f \leq W$ 时, DSB-SC 已调信号 $s(t)$ 的频谱 $S(f)$ 如图 2.6(b) 所示。除了比例因子的改变外, 调制过程其实就是简单地将基带信号的频谱搬到 $\pm f_c$ 处。DSB-SC 调制所需带宽与幅度调制一样, 均为 $2W$ 。

2.3.2 相干检测

从 DSB-SC 已调信号 $s(t)$ 中恢复基带信号 $m(t)$ 的步骤如下: 首先, 将 $s(t)$ 乘以一个本地正弦信号, 然后将乘积通过低通滤波器, 如图 2.7 所示。这里假定本地振荡器信号和乘积调制器中产生的 $s(t)$ 所用的载波 $c(t)$ 在频率和相位上是完全相干或同步的。这种解调方法称为相干检测或同步解调。

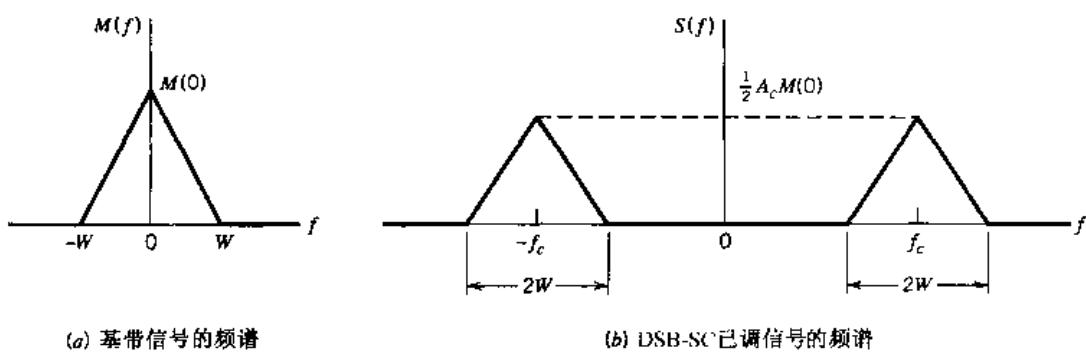


图 2.6 基带信号和 DSB-SC 已调信号的频谱

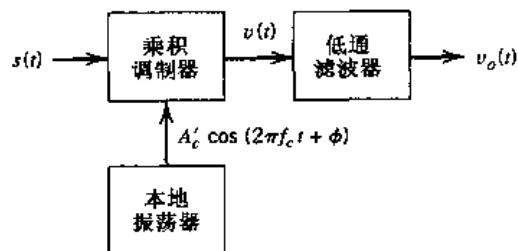


图 2.7 解调 DSB-SC 已调信号的相干检测器

可将相干检测扩展为更普遍使用的一种解调过程, 即采用与载波 $c(t)$ 同频但相位差为 ϕ 的本地振荡器信号进行解调。本地振荡器信号可表示为 $A'_c \cos(2\pi f_c t + \phi)$, 由式(2.8), 得到图 2.7 所示的乘积调制器输出结果为

$$\begin{aligned} v(t) &= A'_c \cos(2\pi f_c t + \phi) s(t) \\ &= A_c A'_c \cos(2\pi f_c t) \cos(2\pi f_c t + \phi) m(t) \\ &= \frac{1}{2} A_c A'_c \cos(4\pi f_c t + \phi) m(t) + \frac{1}{2} A_c A'_c \cos \phi m(t) \end{aligned} \quad (2.10)$$

式(2.10)的第一项表示载频为 $2f_c$ 的 DSB-SC 已调信号, 第二项与基带信号 $m(t)$ 成正比。图 2.8 中的频谱 $V(f)$ 进一步解释了式(2.10)。图 2.8 中, 假定基带信号 $m(t)$ 被限制在间隔 $-W \leq f \leq W$ 内。因此, 很明显公式(2.10)中的第一项通过图 2.7 中所示的低通滤波器后被过滤掉了。假定该低通滤波器的截止频率大于 W 但小于 $2f_c - W$, 为了满足这个条件应选择 $f_c > W$ 。则滤波器的输出端得到的信号表达式为

$$v_o(t) = \frac{1}{2} A_c A'_c \cos \phi m(t) \quad (2.11)$$

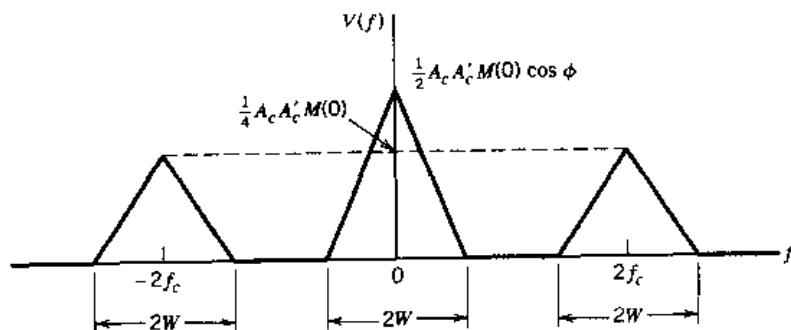


图 2.8 输入为 DSB-SC 已调信号时, 乘积调制器的输出信号频谱

因此,当相位误差 ϕ 为常量时,解调信号 $v_o(t)$ 与 $m(t)$ 成正比。当 $\phi = 0$ 时,解调信号的幅度达到最大值。当 $\phi = \pm \pi/2$ 时,解调信号的幅度达到最小值 0。解调信号为 0 表示相干检波的正交零化效应。因此,由本地振荡器引起的相位误差 ϕ 导致了检测器的输出信号以因子 $\cos\phi$ 减小。只要相位误差 ϕ 为常量,检测器就能够无失真地恢复出原始基带信号 $m(t)$ 。然而实际中由于通信信道的随机波动,相位误差 ϕ 是随机取值的,结果导致在检测器的输出端,乘积因子 $\cos\phi$ 也随着时间随机变化,这种变化是无法预知的。因此,必须使接收机本地振荡器与发射机中产生的 DSB-SC 已调信号的载波在频率和相位上完全同步。载波发射功率的节省是以增大系统的复杂度为代价的。

2.3.3 科斯塔接收机

一种适于解调 DSB-SC 信号的实际同步接收系统采用了如图 2.9 所示的科斯塔(Costas)接收机^[2]。该接收机由使用同一输入信号的两个相干检测器组成,输入信号为 DSB-SC 信号 $A_c \cos(2\pi f_c t) m(t)$ 。但两个检测器各自的本地振荡器相位是相互正交的。调节本地振荡器的频率使之与载频 f_c 相等,这里假设载频已知。上方支路称为同相相干检测器或 I 信道,下方支路称为正交相干检测器或 Q 信道。两个检测器一起组成一个负反馈系统,这样设计的目的是为了保持本地振荡器与载波的同步。

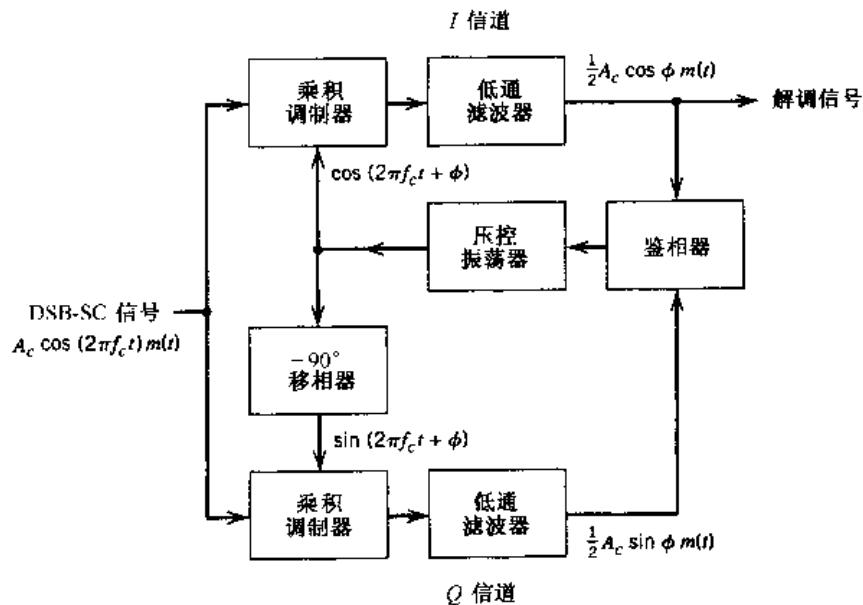


图 2.9 科斯塔接收机

为了理解接收机的工作原理,假定本地振荡器信号与产生输入 DSB-SC 信号的载波 $A_c \cos(2\pi f_c t)$ 是同相的,则 I 信道的输出端包含有用的解调信号 $m(t)$,而 Q 信道由于正交零化现象的影响而输出为 0。若本地振荡器的相位以小角度 ϕ 偏离,则 I 信道的输出保持不变,但此时 Q 信道将输出与 $\sin\phi \approx \phi$ (ϕ 很小时)成正比的信号。当本地振荡器的相位往某一方向偏离时, Q 信道与 I 信道的输出同极性;当本地振荡器的相位向该方向的反向偏离时, Q 信道与 I 信道的输出极性相反。因此,将 I 信道与 Q 信道的输出信号同时送至鉴相器(由乘法器和低通滤波器串联而成,如图 2.9),得到的直流控制信号就能够自动更正压控振荡器中的本地相位误差。

科斯塔接收机的相位控制随调制结束而停止,当调制再次出现时要重新建立锁相。在接收语音传输信号时,这并不是个困难的问题,因为锁相过程非常快,一般察觉不到任何的失真。

2.3.4 正交载波复用

在正交载波复用或正交幅度调制(QAM)中,可有效地利用相干检测中的正交零化效应。在这种方法中,两个DSB-SC已调信号(由两个物理上独立的消息信号分别产生)占用同一信道带宽,但可以在接收机的输出端分离出这两个信号。因此称该方法为带宽保持方案。

正交载波复用系统框图见图2.10。系统发射机部分如图2.10(a)所示,它包括两个分开的乘积调制器,调制器使用的载波是同频的,但存在 -90° 的相位差。传输信号 $s(t)$ 为两个乘积调制器的输出信号之和,即

$$s(t) = A_c m_1(t) \cos(2\pi f_c t) + A_c m_2(t) \sin(2\pi f_c t) \quad (2.12)$$

式中 $m_1(t)$ 和 $m_2(t)$ 表示分别送入两个乘积调制器输入端的不同消息信号。 $s(t)$ 占据以载频 f_c 为中心的 $2W$ 的信道带宽, W 为 $m_1(t)$ 或 $m_2(t)$ 的带宽。根据式(2.12),可将 $A_c m_1(t)$ 看成复用带通信号 $s(t)$ 的同相分量,而将 $-A_c m_2(t)$ 看成 $s(t)$ 的正交分量。

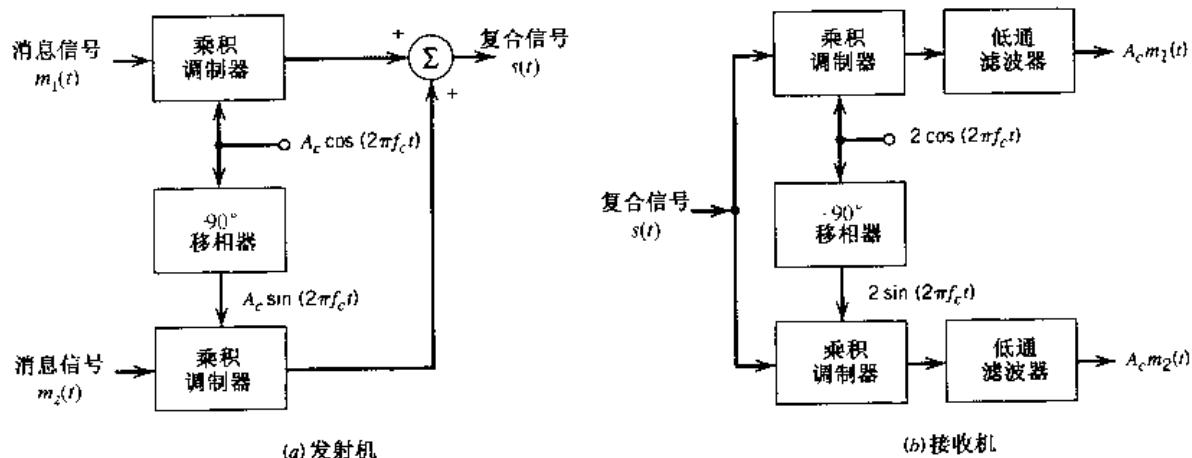


图2.10 正交载波复用系统

系统的接收机部分见图2.10(b)。复用信号 $s(t)$ 同时送至两个分开的相干检测器的输入端,这两个相干检测器采用相位差为 -90° 的同频载波,上面的检测器的输出为 $A_c m_1(t)$,下面检测器的输出为 $A_c m_2(t)$ 。为使系统正常工作,保持发射机与接收机的本地振荡器之间正确的相位和频率关系极为重要。

为了保持同步,可先传送已调信号通带外的控制信号。控制信号为一低功率的正弦音,正弦音的频率和相位与载波 $c(t)$ 有关,在接收机中,通过正确的电路调节后,将其搬到相干检测器使用的正确频率上就可以提取出该控制信号。

2.3.5 SSB 调制

在SSB调制中只传送上边带或下边带,采用频带分离的方法产生已调信号。调制过程如下:

- 第一步,由乘积调制器产生 DSB-SC 已调信号。
- 第二步,使用带通滤波器,通过已调信号的一个边带抑制另一个边带。

从实际角度看,使用频带分离方法产生 SSB 主要是由于无用边带的存在。通过将消息(调制)信号的最低频分量扩大一倍,就可以从有用边带中分离出无用边带的最近频率的分量。为了产生 SSB 已调信号,消息频谱必须使能量间隙的中心在原点,如图 2.11(a)所示。语音信号能够很好地满足这个条件,其能量间隙宽约为 600 Hz(即从 -300 Hz 到 300 Hz)。因此,若保留上边带,则 SSB 已调信号的频谱如图 2.11(b)。

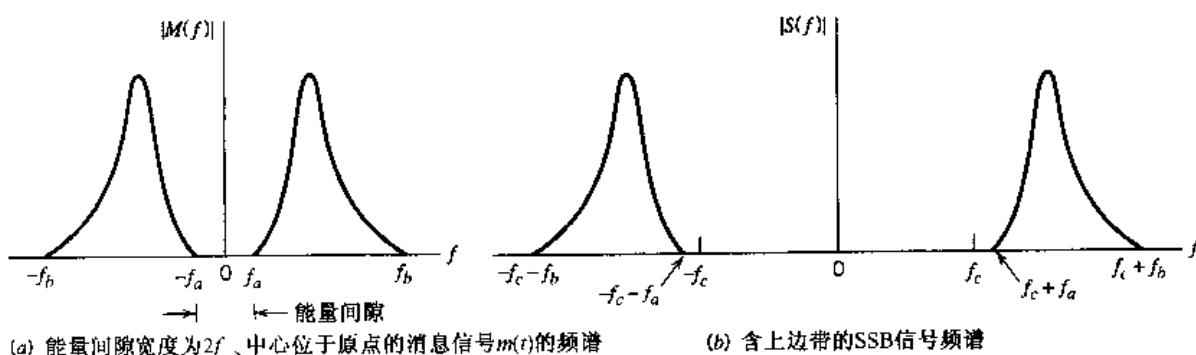


图 2.11 消息频谱与 SSB 已调信号频谱

在对频带分离器(用来产生 SSB 已调信号)中的带通滤波器进行设计时,需满足三个基本要求:

- 有用边带位于滤波器的通带内。
- 无用边带位于滤波器的阻带内。
- 将通带与阻带隔开的滤波器的过渡带是消息信号最低频率分量的两倍。

这种频带分离器经常需要使用选择性很好的滤波器,实际应用中只能选用晶体振荡器。

在解调 SSB 已调信号 $s(t)$ 时,可采用相干检测器,将本地产生的载波乘以 $s(t)$,然后对乘积进行低通滤波。这种解调方法要求相干检测器中的振荡器和发射端产生载波的振荡器之间完全同步。使用下面两种方法之一就可达到这个要求:

- 除了发送选择边带外还要发送低功率导频载波。
- 在接收机中使用与载波同频的高稳定性振荡器。

在第二种方法中,本地振荡器的输出与产生输入 SSB 已调信号的载波间不可避免地会存在相位误差 ϕ 。相位误差的存在使解调信号发生相位失真,即解调得到的原始消息信号的每一个频率分量均产生了恒定的相位偏移 ϕ 。在语音通信中,这种相位失真是可以接受的,因为人耳对相位失真不很敏感。特别地,相位失真可能会引起 Donald Duck 语音效应。但在发送音乐和图像信号时,这种波形失真的存在就绝对不可接受了。

2.3.6 VSB 调制

VSB 调制的做法是,部分地抑制其中一个边带,而发送另一边带的残留部分来补偿抑制掉的那部分边带。产生 VSB 调制信号的一般方法是频带分离法。首先,产生 DSB-SC 已调信号,并使其通过如图 2.12 所示的带通滤波器。带通滤波器经过了特殊的设计,可将 VSB 调制与

SSB 调制区分开来。假设发送下边带的残留部分，则带通滤波器的频率响应 $H(f)$ 如图 2.13 所示。为简化起见，图 2.13 只画出了频率响应中的正频率部分。这里的频率响应已进行了归一化，因此在载频处有 $|H(f_c)| = 1/2$ 。由图 2.13 可得出频率响应的一个重要特点：在载频 f_c 附近，频率响应的截止部分呈奇对称，即频率响应在过渡间隔 $f_c - f_v \leq |f| \leq f_c + f_v$ 内满足下面两个条件：

1. 与 f_c 等距离的任意两个频率点的幅度响应 $|H(f)|$ 之和为 1。
2. 相位响应 $\arg(H(f))$ 是线性的，即 $H(f)$ 满足条件

$$H(f - f_c) + H(f + f_c) = 1 \quad -W \leq f \leq W \quad (2.13)$$

同时需要注意的是，在有用频带（即 $|f| > f_c + W$ ）外，频率响应 $H(f)$ 有可能为任意值。因此 VSB 调制的传输带宽为

$$B_T = W + f_v \quad (2.14)$$

式中， W 为消息带宽， f_v 为残留边带宽度。

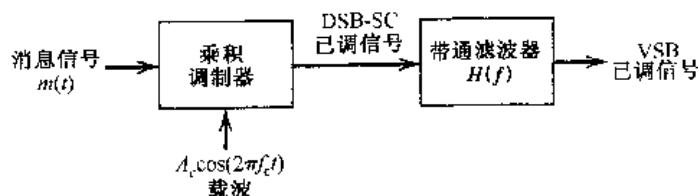


图 2.12 由滤波法产生 VSB 已调信号

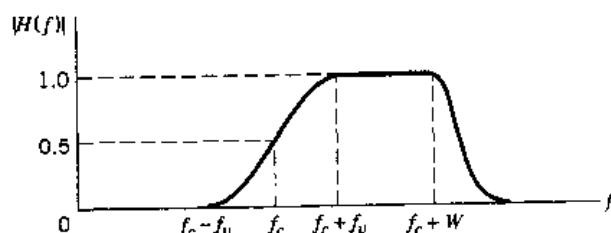


图 2.13 VSB 滤波器的幅度响应(只画出了正频率部分)

根据表 2.1，VSB 已调信号的时域表达式为

$$s(t) = \frac{1}{2} A_c m(t) \cos(2\pi f_c t) \pm \frac{1}{2} A_c m'(t) \sin(2\pi f_c t) \quad (2.15)$$

发送上边带的残留部分时取“+”号，发送下边带的残留部分时取“-”号。将消息信号 $m(t)$ 通过频率响应 $H_Q(f)$ 满足如下要求（见习题 2.20）的滤波器，从而在 $s(t)$ 的正交分量中得到信号 $m'(t)$ ：

$$H_Q(f) = j[H(f - f_c) - H(f + f_c)] \quad -W \leq f \leq W \quad (2.16)$$

图 2.14 的频率响应 $H_Q(f)$ 中，纵轴的单位刻度为 $1/j$ 。由 $H_Q(f)$ 决定的正交分量的作用是干预式（2.15）中的同相分量，这样可以部分降低已调信号 $s(t)$ 一个边带的功率，并保留另一个边带有用的残留部分。

SSB 调制可看成是 VSB 调制的一个特例。特别是当残留边带降为 0 时（即置 $f_v = 0$ ），式（2.15）中的已调信号 $s(t)$ 的极限形式即为单边带已调信号的表达式。

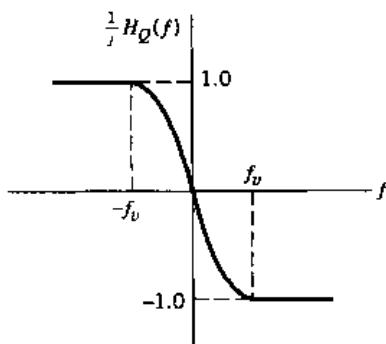


图 2.14 产生 VSB 已调信号正交分量的滤波器的频率响应

2.3.7 电视信号

为完整起见, 我们不得不提到残留边带调制在商用 TV(电视)广播中的作用。TV 系统发送视频信号时采用的调制方式具有一些特殊的细节问题, 这些问题受以下两个因素影响:

1. 视频信号占用很大的带宽, 并具有重要的低频成分, 因此应采用 VSB 调制。
2. 接收机中解调电路应简单而廉价, 所以应采用包络检波, 这需要在 VSB 已调信号中附加载波。

然而, 就第一点而言, 需要强调的是尽管基本目标是保留带宽, 但在商用 TV 广播中, 发送的信号并不完全是 VSB 调制。原因是发射机的功率很大, 要严格地控制边带滤波将耗费巨大的成本。解决办法是将 VSB 滤波器插入到每一个接收机中, 因为接收机的功率较低。这样做的整体效果与传统的残留边带调制相同, 只是浪费了一些功率和带宽, 如图 2.15 所示。图 2.15(a) 为发送的 TV 信号的理想频谱, 发射信号包含了上边带、25% 的下边带以及图像载波。接收机中用于形成所需频谱的 VSB 滤波器的频率响应如图 2.15(b) 所示。

北美用于发送 TV 广播的信道带宽为 6 MHz, 如图 2.15(b) 所示。这个信道带宽不仅包括对视频信号进行 VSB 调制所需的带宽, 还包括相伴随的声音信号所需的带宽。声音信号有自己的调制载波。图 2.15(a) 和图 2.15(b) 中标出的频率值属于一个专门的 TV 信道。图 2.15 中的图像载频为 55.25 MHz, 声音载频为 59.75 MHz。TV 信号的信息内容应位于基带频谱内, 基带频谱的范围为低于图像载波 1.25 MHz 到高于图像载波 4.5 MHz。

至于第二点, 采用包络检波(用于含载波的 VSB 已调信号)将导致检波器输出端恢复的视频信号发生波形失真。这种失真是由 VSB 已调信号的正交分量引起的, 接下来对这个问题进行讨论。

式(2.15)中使用的时域描述, 使我们可以确定包络检波器引起的波形失真。特别地, 将式(2.15)的 VSB 已调信号乘以因子 k_a , 再加上载波分量 $A_c \cos(2\pi f_c t)$, 则修正的包络检波器的输入已调信号为

$$s(t) = A_c \left[1 + \frac{1}{2} k_a m(t) \right] \cos(2\pi f_c t) \pm \frac{1}{2} k_a A_c m'(t) \sin(2\pi f_c t) \quad (2.17)$$

其中, 常数 k_a 决定了调制百分比。用 $a(t)$ 表示的包络检波器的输出为

$$\begin{aligned}
 a(t) &= A_c \left\{ \left[1 + \frac{1}{2} k_a m(t) \right]^2 + \left[\frac{1}{2} k_a m'(t) \right]^2 \right\}^{1/2} \\
 &= A_c \left[1 + \frac{1}{2} k_a m(t) \right] \left\{ 1 + \left[\frac{\frac{1}{2} k_a m'(t)}{1 + \frac{1}{2} k_a m(t)} \right]^2 \right\}^{1/2}
 \end{aligned} \quad (2.18)$$

由式(2.18)可以看出失真是由 $m'(t)$ 引起的, $m'(t)$ 为输入 VSB 已调信号的正交分量。采用以下两种方法可减小失真:

- 减少调制百分比以降低调幅灵敏度 k_a 。
- 通过增加残留边带的带宽以减小 $m'(t)$ 。

在实际应用中,以上两种方法均被采用了。在商用 TV 广播中,通过选择残留边带的宽度(约为 0.75 MHz, 或整个边带宽度的 1/6),使得调制百分比接近 100%,此时由 $m'(t)$ 造成的失真可保持在容许的范围内。

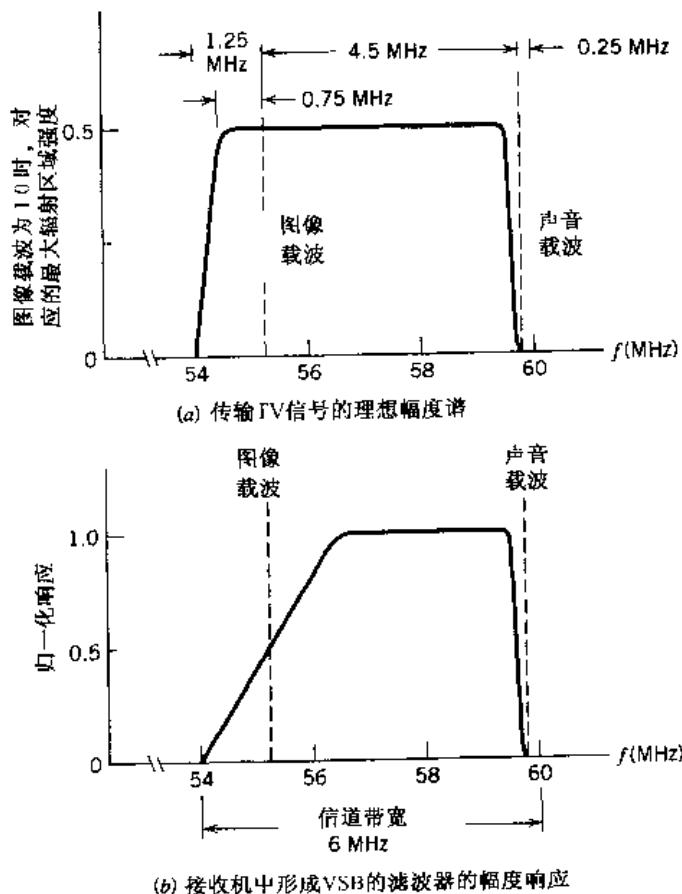


图 2.15 将 VSB 滤波器插入接收机

2.4 频率搬移

SSB 调制中的基本操作实际上就是一种频率搬移,这就是有时会将 SSB 调制称为变频、混

频或外差的原因。将图 2.11(a)中原始消息信号的频谱与图 2.11(b)中信号的频谱进行比较，可以清楚地看到这种操作的效果。图 2.11(a)中消息频谱的正频率部分占用的频带范围从 f_a 到 f_b ，而在图 2.11(b)中该频带被向上搬移了 f_c ，而负频率部分的消息频谱以对称的方式向下搬移了 f_c 。

这里描述的频谱搬移可按如下方法完成。假定有已调信号 $s_1(t)$ ，其频谱中心位于载频 f_1 ，要将它向上搬移，使其中心载频从 f_1 变为 f_2 ，可使用图 2.16 所示的混频器来实现。混频器由乘积调制器和带通滤波器串联而成。

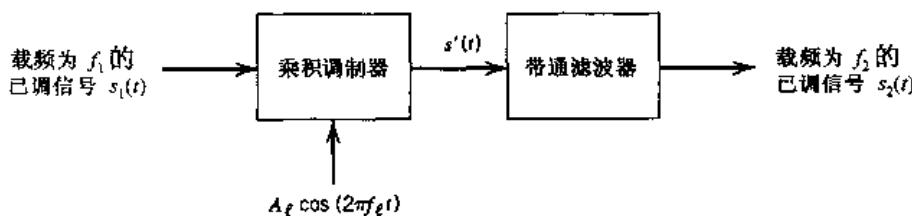


图 2.16 混频器框图

为了解释混频器的工作原理，我们考虑图 2.17 所示的情况。为了便于说明，假定混频器的输入 $s_1(t)$ 是载频为 f_1 ，带宽为 $2W$ 的 AM 信号。图 2.17(a)为 AM 信号的频谱 $S_1(f)$ ，这里假设 $f_1 > W$ 。图 2.17(b)为乘积调制器的输出信号 $s'(t)$ 的频谱 $S'(f)$ 。

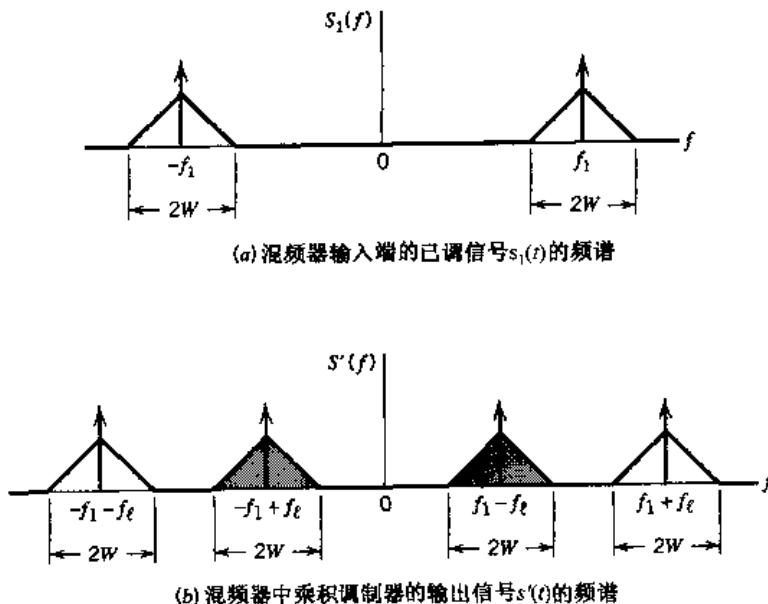


图 2.17 混频器的工作原理

信号 $s'(t)$ 可看成是两个已调分量之和。图 2.17(b)中的阴影部分的频谱为一个分量，非阴影部分的频谱为另一分量。根据输入载频 f_1 是向上搬移还是向下搬移，可定义两种不同的情况：

上变换。此时的载频 f_2 大于输入载频 f_1 ，所需本地振荡器的频率 f_l 定义为

$$f_l = f_2 + f_1$$

或

$$f_l = f_2 - f_1$$

将图 2.17(b) 中非阴影部分的频谱定义为有用的已调信号 $s_2(t)$, 将阴影部分的频谱定义为与 $s_2(t)$ 相关联的镜像信号。显而易见, 这种情况下的混频器称为上变频器。下变换。此时的搬移载频 f_2 小于输入载频 f_1 , 所需的振荡器频率 f_l 定义如下:

$$f_2 = f_1 - f_l$$

或

$$f_l = f_1 - f_2$$

此时得到的频谱图与上变换的相应频谱图相反。将图 2.17(b) 中阴影部分的频谱定义为有用已调信号 $s_2(t)$, 而非阴影部分的频谱定义为与 $s_2(t)$ 相关联的镜像信号。此时称该混频器为下变频器。注意, 为了避免边带重叠, 这种情况下的搬移载频 f_2 必须大于 W (即已调信号带宽的一半)。

图 2.16 所示的混频器中带通滤波器的作用是通过有用信号 $s_2(t)$, 抑制相关联的镜像信号。通过调节滤波器的中心频率使之与搬移载频 f_2 相等, 并使滤波器的带宽等于输入已调信号 $s_1(t)$ 的带宽, 即可实现这一目的。

需要注意的是, 混频是一种线性操作。因此输入已调信号的边带与载波的关系在混频器的输出端被完整地保留下来。

2.5 频分复用

另一个重要的信号处理过程是复用, 即将许多相互独立的信号复合成适于在同一公共信道中传送的信号。例如, 电话系统可传送 300 Hz 到 3100 Hz 范围内的语音频率。为了在同一信道中传送大量这样的信号, 必须将这些信号分开, 使它们之间互不干扰, 这样才能在接收端将这些信号区分开来。可以利用频率或时间区分信号, 从而达到复用的目的。利用频率来区分信号的技术被称为频分复用(FDM), 而利用时间来区分信号的技术则称为时分复用(TDM)。本节讨论 FDM 系统, TDM 系统将在第 3 章中进行讨论。

FDM 系统框图如图 2.18 所示。假定输入消息信号为低通信号, 但它们的频谱在降为零频率之前未必都是非零值。每一输入信号都通过一个低通滤波器, 其目的是滤掉高频分量, 这些高频分量对信号作用不大, 而且还会干扰同一信道中的其他消息信号。只有当原始输入信号为充分带限时才可省去这些低通滤波器。

将经过滤波的信号送入调制器, 调制器改变了信号的频率范围, 使信号占用互不重叠的频带。完成频谱搬移所需的载频由载波发生器产生。该框图中的调制可用本章前几节讲到的任意一种调制方法。然而, 频分复用中采用最多的调制方法还是 SSB 调制。语音信号采用 SSB 调制时需要的带宽近似等于原始语音信号的带宽。在实际应用中, 给每一个语音输入端分配 4 kHz 的带宽。调制器后的带通滤波器将每个已调信号限制在分配的带宽范围内。接着将带通滤波器的输出并联在一起, 作为公共信道的输入。在接收端, 使用输入端并联的若干个带通滤波器, 以便利用频率分离消息信号。最后, 通过带通滤波器各自的解调器恢复出原始消息信号。

注意的是,图 2.18 的 FDM 系统是单向工作的。如果要完成双向传输(如电话),就必须再完整地复制一套复用设备,其中的各个部件以相反的方向排列,信号流向从右往左。

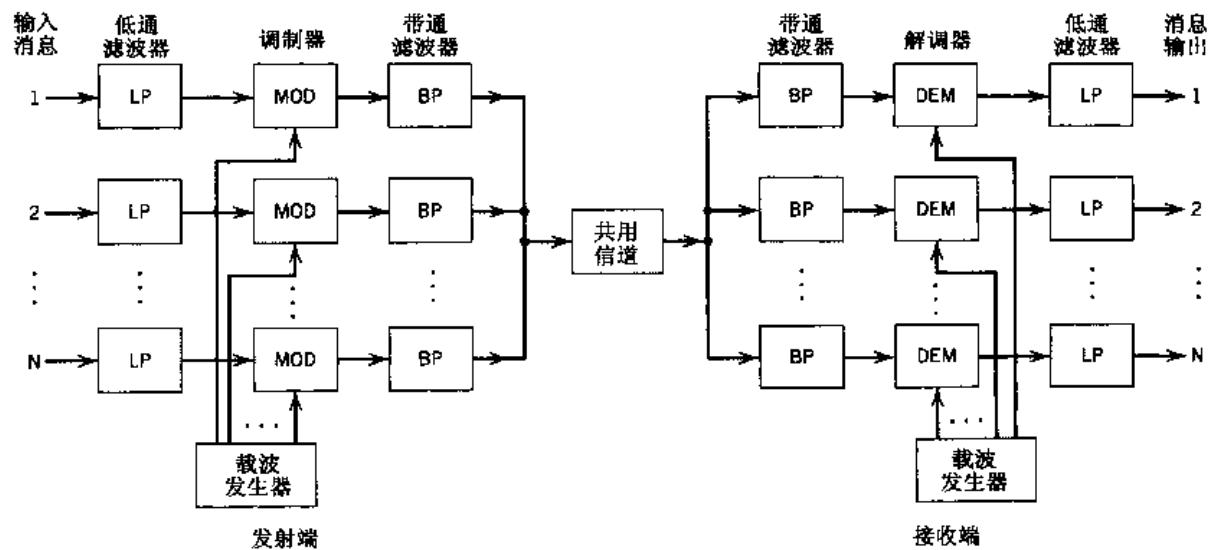


图 2.18 FDM 系统框图

例 2.1

实际的 FDM 系统通常包括多步调制和解调,如图 2.19 所示。第一步复用是将 12 个语音输入合成一个基群,合成的准则是将第 n 个输入信号调制到频率 $f_c = 60 + 4n$ kHz 的载波上,其中 $n = 1, 2, \dots, 12$ 。用带通滤波器选择下边带,并将其复合为含有 12 个下边带(每一边带对应一个语音输出)的基群。基群占用的频带为 60 ~ 108 kHz。下一步复用是将 5 个基群合成一个超群,合成准则是将第 n 群信号调制到频率为 $f_c = 372 + 48n$ kHz 的载波上,其中 $n = 1, 2, \dots, 5$ 。通过滤波器选择下边带,并将下边带组成一起超群,超群占用的带宽为 312 ~ 552 kHz。

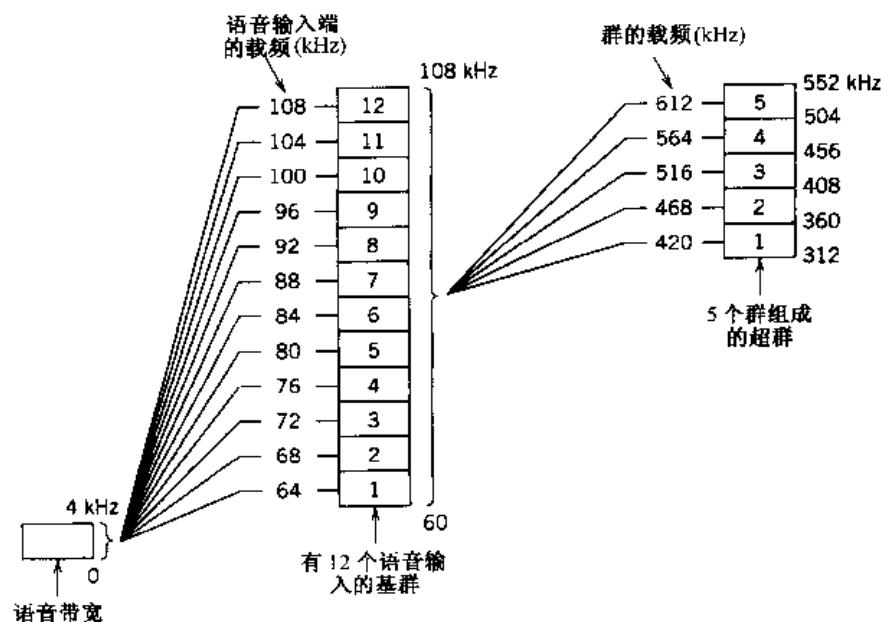


图 2.19 FDM 系统的调制步骤

所以,一个超群包括60个相互独立的语音输出。采用这种形式合成超群的原因是,具备所需特性的较为经济的滤波器只可以工作在有限频率范围内。采用类似的方法,还可将超群复合为主群,主群复合为超主群。

2.6 角度调制

在本章的前几节里,我们研究了使正弦载波的幅度随基带(承载消息)信号缓慢变化的调制方法。调制正弦载波还有另外一种方法,称为角度调制。在角度调制中,正弦载波的角度随基带信号的变化而变化。在这种调制方法中,载波的幅度不变。角度调制的重要特点是它的抗噪能力强于幅度调制,这一点在后面的2.7节中可以看到。但抗噪能力的改善是以增加传输带宽为代价的,即角度调制提供了一种实用的方法:用信道带宽来换取噪声性能的改善。而幅度调制无论采用何种形式,都无法获得这样的抗噪性能。

2.6.1 基本定义

令 $\theta_i(t)$ 表示已调正弦载波的角度,它是消息信号的函数。可将角度调制信号表示为

$$s(t) = A_c \cos[\theta_i(t)] \quad (2.19)$$

其中, A_c 为载波幅度。当 $\theta_i(t)$ 改变了 2π 弧度时,完成一次振荡。若 $\theta_i(t)$ 仅随时间单调增加,那么在时间间隔 t 到 $t + \Delta t$ 内, $\theta_i(t)$ 的平均频率(以Hz记)为

$$f_{\Delta t}(t) = \frac{\theta_i(t + \Delta t) - \theta_i(t)}{2\pi \Delta t} \quad (2.20)$$

可将角度调制信号 $s(t)$ 的瞬时角频率定义为

$$\begin{aligned} f_i(t) &= \lim_{\Delta t \rightarrow 0} f_{\Delta t}(t) \\ &= \lim_{\Delta t \rightarrow 0} \left[\frac{\theta_i(t + \Delta t) - \theta_i(t)}{2\pi \Delta t} \right] \\ &= \frac{1}{2\pi} \frac{d\theta_i(t)}{dt} \end{aligned} \quad (2.21)$$

因此,根据式(2.19),可将角度调制信号 $s(t)$ 理解为长度为 A_c ,角度为 $\theta_i(t)$ 的旋转矢量。根据式(2.21),矢量的角速度等于 $d\theta_i(t)/dt$,单位为弧度每秒。对于未调载波,角度 $\theta_i(t)$ 为

$$\theta_i(t) = 2\pi f_c t + \phi_c$$

相应的矢量以恒定角速度 $2\pi f_c$ 旋转,常量 ϕ_c 为 $\theta_i(t)$ 在 $t = 0$ 时的取值。

角度 $\theta_i(t)$ 随消息(基带)信号的变化可有若干种方式。这里只考虑两种常用的方式,即相位调制(PM)和频率调制(FM),其定义如下:

1. PM是角度 $\theta_i(t)$ 随消息信号 $m(t)$ 呈线性变化的一种角度调制,即

$$\theta_i(t) = 2\pi f_c t + k_p m(t) \quad (2.22)$$

其中 $2\pi f_c t$ 表示未调载波的角度;常量 k_p 表示调制器的相位灵敏度,若 $m(t)$ 的波形用电压表示,则 k_p 的单位为rad/V(弧度每伏)。为了简便,假定在式(2.22)中,当 $t = 0$

时,未调载波的角度为 0。因此角度调制信号 $s(t)$ 的时域表达式为

$$s(t) = A_c \cos[2\pi f_c t + k_p m(t)] \quad (2.23)$$

2. FM 是瞬时频率 $f_i(t)$ 随消息信号 $m(t)$ 呈线性变化的一种角度调制。即

$$f_i(t) = f_c + k_f m(t) \quad (2.24)$$

其中 f_c 表示未调载波的频率;常量 k_f 表示调制器的频率灵敏度,若 $m(t)$ 的波形用电压表示,则 k_f 的单位是 Hz/V(赫每伏)。将式(2.24)按照时间积分,并将结果乘以 2π ,可得

$$\theta_i(t) = 2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \quad (2.25)$$

为简便起见,假定未调载波的角度在 $t=0$ 时为 0。因此 FM 信号的时域表达式为

$$s(t) = A_c \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right] \quad (2.26)$$

在式(2.22)或其积分形式(2.25)中,角度 $\theta_i(t)$ 随消息信号 $m(t)$ 变化的结果为,PM 信号或 FM 信号的过零点的分布不再具有较强的规律性。过零点是指波形从负值变到正值或相反过程所对应的时刻。这是 PM 和 FM 信号与 AM 信号的一个重要区别。另一个重要的区别在于 PM 或 FM 信号的包络是恒定的(等于载波幅度),而 AM 信号的包络是随消息信号而变化的。

比较式(2.23)和式(2.26)可以看出,可将 FM 信号视为用 $\int_0^t m(\tau) d\tau$ 代替 $m(t)$ 作为调制信号的一类 PM 信号。这说明 FM 信号可以用这样的方法产生:首先对 $m(t)$ 积分,然后将积分结果作为相位调制器的输入信号,如图 2.20(a) 所示。相反,产生 PM 信号时应先对 $m(t)$ 微分,然后作为频率调制器的输入信号,如图 2.20(b) 所示。因此可从 FM 信号归纳出 PM 信号的所有特点,反之亦然。因此,我们将重点放在 FM 信号的讨论上。

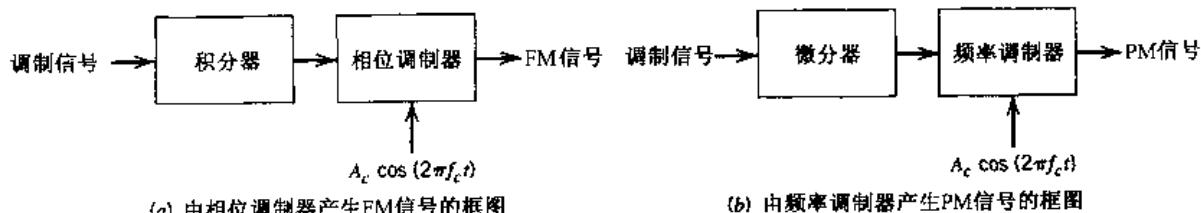


图 2.20 FM 和 PM 之间的关系示意图

2.7 频率调制

式(2.26)定义的 FM 信号 $s(t)$ 是调制信号 $m(t)$ 的非线性函数,这说明频率调制是一个非线性调制过程。因此,与幅度调制不同,FM 信号的频谱与调制信号的频谱之间不再是一种简单的关系,FM 信号的频谱分析要比 AM 信号的频谱分析复杂得多。

如何着手分析 FM 信号的频谱呢?可用以下的经验方法进行分析:

- 尽可能只考虑最简单的情况,即用单音频调制信号产生窄带 FM 信号。

- 然后考察较为普遍的情况,也是单音频调制,但此时 FM 信号为宽带信号。

当然,也可以继续讨论多音频 FM 信号这一复杂的情况。但由于这里的研究目标是建立 FM 信号的传输带宽和消息带宽之间的一种经验关系,因此不需要深入讨论。随后读者可看到,上述两步频谱分析已经足以解决这个问题。

接下来考虑一个定义如下的正弦调制信号:

$$m(t) = A_m \cos(2\pi f_m t) \quad (2.27)$$

则生成的 FM 信号的瞬时角频率等于

$$\begin{aligned} f_i(t) &= f_c + k_f A_m \cos(2\pi f_m t) \\ &= f_c + \Delta f \cos(2\pi f_m t) \end{aligned} \quad (2.28)$$

其中

$$\Delta f = k_f A_m \quad (2.29)$$

称为频偏,代表 FM 信号的瞬时频率偏离载频 f_c 的最大值。FM 信号的一个基本特点是频偏 Δf 与调制信号的幅度成正比,而与调制频率无关。

由式(2.28)得到 FM 信号的角度 $\theta_i(t)$ 如下:

$$\begin{aligned} \theta_i(t) &= 2\pi \int_0^t f_i(\tau) d\tau \\ &= 2\pi f_c t + \frac{\Delta f}{f_m} \sin(2\pi f_m t) \end{aligned} \quad (2.30)$$

频偏 Δf 与调制频率 f_m 的比值通常称为 FM 信号的调制指数。用 β 表示,因而有

$$\beta = \frac{\Delta f}{f_m} \quad (2.31)$$

和

$$\theta_i(t) = 2\pi f_c t + \beta \sin(2\pi f_m t) \quad (2.32)$$

由式(2.32)可见,参数 β 的物理意义为表示 FM 信号的相偏,即角度 $\theta_i(t)$ 偏离未调载波角度 $2\pi f_c t$ 的最大值,因此 β 的单位为弧度。

FM 信号为

$$s(t) = A_c \cos[2\pi f_c t + \beta \sin(2\pi f_m t)] \quad (2.33)$$

调制指数 β 的值可用来区分频率调制的两种情况:

- 窄带 FM, β 小于 1 弧度。
- 宽带 FM, β 大于 1 弧度。

下面依次介绍这两种频率调制。

2.7.1 窄带 FM

式(2.33)定义了由正弦调制信号得到的 FM 信号,展开该式有

$$s(t) = A_c \cos(2\pi f_c t) \cos[\beta \sin(2\pi f_m t)] - A_c \sin(2\pi f_c t) \sin[\beta \sin(2\pi f_m t)] \quad (2.34)$$

若调制指数 β 小于 1 弧度,那么可用以下的近似式

$$\cos[\beta \sin(2\pi f_m t)] \approx 1$$

和

$$\sin[\beta \sin(2\pi f_m t)] \approx \beta \sin(2\pi f_m t)$$

因此式(2.34)可简化为

$$s(t) \approx A_c \cos(2\pi f_c t) - \beta A_c \sin(2\pi f_c t) \sin(2\pi f_m t) \quad (2.35)$$

式(2.35)定义了由正弦调制信号 $A_m \cos(2\pi f_m t)$ 得到的窄带 FM 信号的近似形式。从该表达式可以推出图 2.21 所示的调制器框图。这个调制器将载波 $A_c \cos(2\pi f_c t)$ 分为两路。一路是直接路径；另一路包括一个 -90° 的移相网络和一个乘积调制器，乘积调制器可以产生 DSB-SC 已调信号。两路信号的差值即为窄带 FM 信号，但存在一些失真。

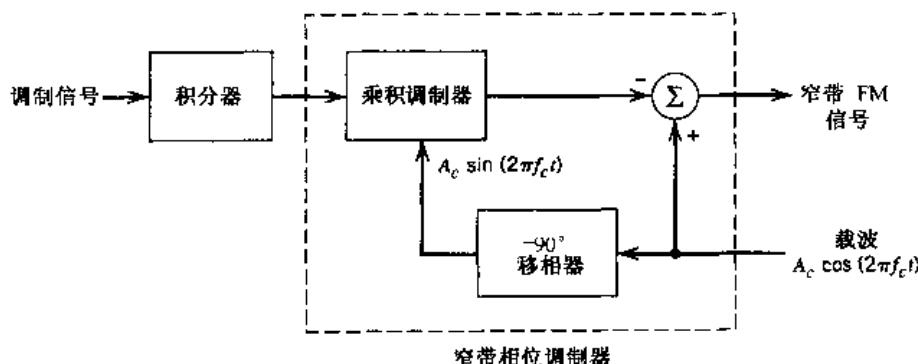


图 2.21 窄带 FM 信号的产生框图

理想情况下，FM 信号具有恒定包络，若正弦调制信号的频率为 f_m ，则角度 $\theta_i(t)$ 也为正弦变化，频率也为 f_m 。但是，由图 2.21 所示窄带调制器产生的已调信号与理想情况的 FM 信号并不完全吻合，二者主要存在两方面的区别：

1. 包络包含残余幅度调制，因此随时间的变化而变化。
2. 对于正弦调制信号，角度 $\theta_i(t)$ 存在调制频率 f_m 的三次和更高次的谐波失真。

但当把调制指数限制在 $\beta \leq 0.3$ 弧度范围内时，就可忽略残留 AM 和谐波 PM 造成的影响了。

再看式(2.35)，可对其做如下展开：

$$s(t) \approx A_c \cos(2\pi f_c t) + \frac{1}{2} \beta A_c [\cos[2\pi(f_c + f_m)t] - \cos[2\pi(f_c - f_m)t]] \quad (2.36)$$

该表达式与如下 AM 信号的表达式有些类似：

$$s_{AM}(t) = A_c \cos(2\pi f_c t) + \frac{1}{2} \mu A_c [\cos[2\pi(f_c + f_m)t] + \cos[2\pi(f_c - f_m)t]] \quad (2.37)$$

其中， μ 为 AM 信号的调制因子。比较式(2.36)和式(2.37)可以看出，在正弦调制中，AM 信号与窄带 FM 信号之间基本的不同在于窄带 FM 信号中的下边频的代数符号与 AM 信号中相反。因此，窄带 FM 信号需要与 AM 信号同样的传输带宽(即 $2f_m$)。

我们可以用图 2.22(a)中的矢量图表示窄带 FM 信号，其中以载波矢量为参考。由该图可见，两个边频矢量的合成总是位于载波矢量的右侧。这使得表示窄带 FM 信号的合成矢量幅

度与载波矢量的幅度近似相等,但相位并不相等。将此矢量图与图 2.22(b)中表示的 AM 信号的矢量图相比较,由图 2.22(b)可以看到,表示 AM 的矢量幅度与载波矢量幅度不相等,但它们的相位是相同的。

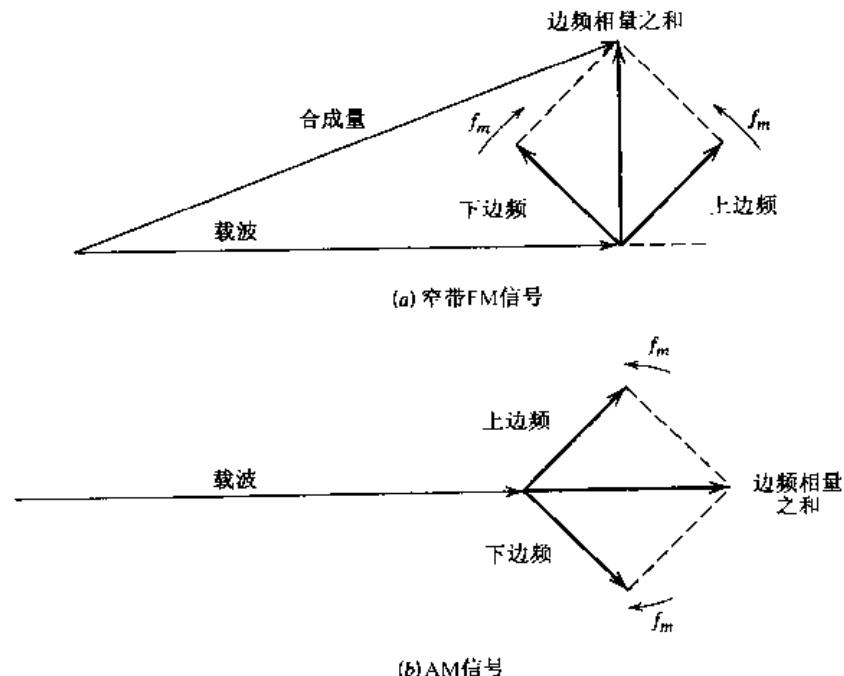


图 2.22 采用正弦调制时,AM 信号与窄带 FM 信号之间的相量比较

2.7.2 宽带 FM

下面在调制指数 β 取任意值的情况下,确定式(2.33)中的单音频 FM 信号的频谱。总的来说,由一正弦调制信号得到的 FM 信号(见式(2.33)),其自身是非周期的,除非载频 f_c 是调制频率 f_m 的整数倍。但可用附录 2 中带通信号的复数表达式来适当简化问题。具体地,假定载频 f_c 足够大(与 FM 信号的带宽相比),则可以将式(2.33)写成

$$\begin{aligned} s(t) &= \operatorname{Re}[A_c \exp(j2\pi f_c t + j\beta \sin(2\pi f_m t))] \\ &= \operatorname{Re}[\tilde{s}(t) \exp(j2\pi f_c t)] \end{aligned} \quad (2.38)$$

其中 $\tilde{s}(t)$ 为 FM 信号 $s(t)$ 的复包络,其定义为

$$\tilde{s}(t) = A_c \exp[j\beta \sin(2\pi f_m t)] \quad (2.39)$$

所以,和原始的 FM 信号 $s(t)$ 不同的是,复包络 $\tilde{s}(t)$ 是一个关于时间的周期函数,基频等于调制频率 f_m 。对 $\tilde{s}(t)$ 做傅里叶级数展开如下:

$$\tilde{s}(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi n f_m t) \quad (2.40)$$

其中,傅里叶系数 c_n 的定义为

$$\begin{aligned} c_n &= f_m \int_{-1/2f_m}^{1/2f_m} \tilde{s}(t) \exp(-j2\pi n f_m t) dt \\ &= f_m A_c \int_{-1/2f_m}^{1/2f_m} \exp[j\beta \sin(2\pi f_m t) - j2\pi n f_m t] dt \end{aligned} \quad (2.41)$$

引入新变量:

$$x = 2\pi f_m t \quad (2.42)$$

因此,以新的形式重写式(2.41)为

$$c_n = \frac{A_c}{2\pi} \int_{-\pi}^{\pi} \exp[j(\beta \sin x - nx)] dx \quad (2.43)$$

式(2.43)右侧的积分式去掉比例因子后,可看成是第一类 n 阶贝塞尔函数^[3],其自变量为 β ,通常用 $J_n(\beta)$ 表示贝塞尔函数:

$$J_n(\beta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp[j(\beta \sin x - nx)] dx \quad (2.44)$$

由此,可将式(2.43)简化为

$$c_n = A_c J_n(\beta) \quad (2.45)$$

将式(2.45)代入式(2.40),则 FM 信号复包络的展开式可以用贝塞尔函数 $J_n(\beta)$ 表示为

$$\tilde{s}(t) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \exp(j2\pi n f_m t) \quad (2.46)$$

接下来,将式(2.46)代入式(2.38),则有

$$s(t) = A_c \cdot \operatorname{Re} \left[\sum_{n=-\infty}^{\infty} J_n(\beta) \exp[j2\pi(f_c + n f_m)t] \right] \quad (2.47)$$

交换式(2.47)中求和的顺序,并取实部,有

$$s(t) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta) \cos[2\pi(f_c + n f_m)t] \quad (2.48)$$

这就是 β 取任意值时单音频 FM 信号 $s(t)$ 的傅里叶级数表达式。对式(2.48)两侧取傅里叶变换可求出 $s(t)$ 的离散频谱为

$$S(f) = \frac{A_c}{2} \sum_{n=-\infty}^{\infty} J_n(\beta) [\delta(f - f_c - n f_m) + \delta(f + f_c + n f_m)] \quad (2.49)$$

图 2.23 为 n 取不同的正整数值时,贝塞尔函数 $J_n(\beta)$ 对调制指数 β 的曲线。利用下面的性质(细节见附录 3),可进一步了解贝塞尔函数 $J_n(\beta)$ 的应用:

$$1. J_n(\beta) = (-1)^n J_{-n}(\beta) \quad \text{对于所有 } n \text{ (正或负)} \quad (2.50)$$

2. 当调制指数 β 很小时,有

$$\left. \begin{aligned} J_0(\beta) &\approx 1 \\ J_1(\beta) &\approx \frac{\beta}{2} \\ J_n(\beta) &\approx 0, \quad n > 2 \end{aligned} \right\} \quad (2.51)$$

$$3. \sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1 \quad (2.52)$$

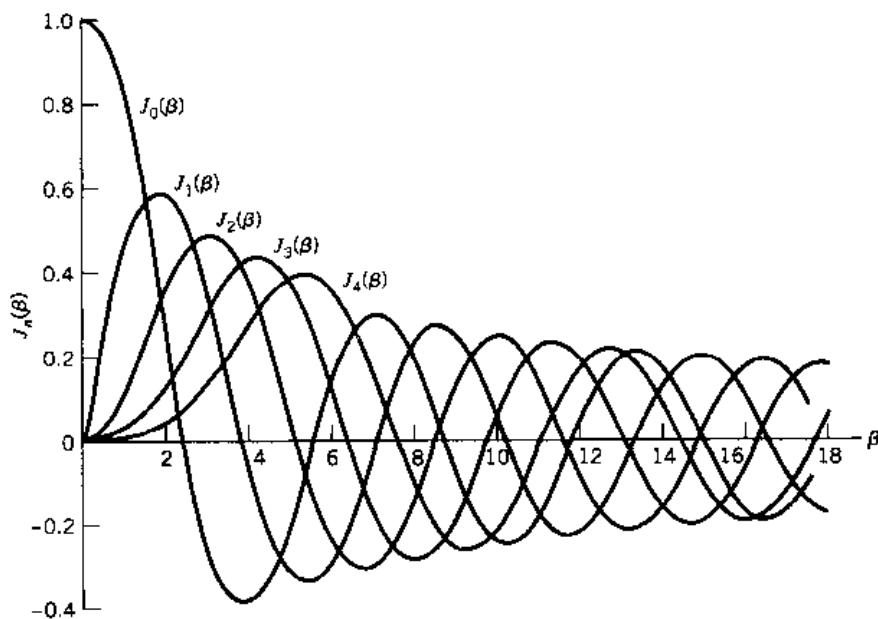


图 2.23 第一类各阶贝塞尔函数曲线

根据式(2.49)~式(2.52)以及图 2.23 中的曲线,可以得到以下结论:

1. FM 信号的频谱含有载波分量和无穷多个边频分量 $f_m, 2f_m, 3f_m, \dots$, 边频分量以载频为中心, 呈对称分布。在这点上, 它不同于 AM 系统, 在 AM 系统中正弦调制信号只产生一对边频。
2. 对于 β 小于 1 的特殊情况, 贝塞尔系数中只有 $J_0(\beta)$ 和 $J_1(\beta)$ 有确定值, 因此 FM 信号由一个载频和位于 $f_c \pm f_m$ 处的一对边频组成。这相当于前面介绍过的窄带 FM 信号的特殊情况。
3. 载波分量的幅度随 $J_0(\beta)$ 中的 β 值变化。即 FM 信号中载波分量的幅度由调制指数 β 决定, 这是不同于 AM 信号的。这个性质的物理含义是: FM 信号是恒包络的。因此, FM 信号在 1Ω 的电阻上产生的平均功率为恒定值, 即

$$P = \frac{1}{2} A_c^2 \quad (2.53)$$

当载波经调制产生 FM 信号后, FM 信号中的边频将损耗部分载波功率, 因此载波分量的幅度与 β 值有关。注意, FM 信号的平均功率也可由式(2.48)得到, 即

$$P = \frac{1}{2} A_c^2 \sum_{n=-\infty}^{\infty} J_n^2(\beta) \quad (2.54)$$

将式(2.52)代入式(2.54)就可将平均功率 P 的表达式简化为式(2.53)。

例 2.2

在此例中, 将研究正弦调制信号幅度和频率的变化对 FM 信号频谱的影响。首先, 考虑调制信号的频率固定不变而幅度发生变化时, 相应的频偏 Δf 发生的变化。保持调制频率

f_m 不变，则 FM 信号的幅谱曲线如图 2.24，其中 β 分别取值 1, 2, 5。在图 2.24 中，我们对幅度谱按未调载波的幅度进行了归一化。

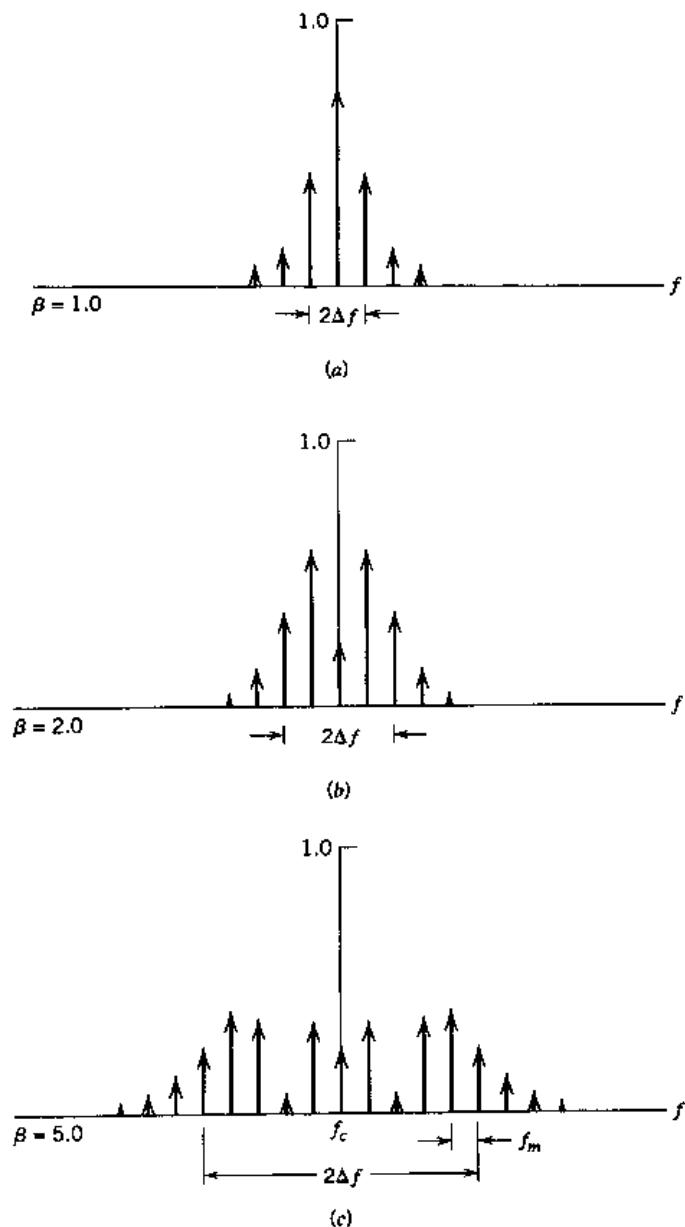


图 2.24 FM 信号的离散幅度谱(图中对载波幅度进行了归一化。其中正弦调制的载频不变，幅度发生改变。并且只画出了频谱的正频率部分)

下面考虑调制信号幅度不变，即频偏 Δf 为常量而调制频率 f_m 变化时的情况。这种情况下的 FM 信号幅度谱曲线如图 2.25 所示， β 取值为 1, 2, 5。由曲线可看出当 Δf 不变， β 增大时，固定频率间隔 $f_c - \Delta f < |f| < f_c + \Delta f$ 内的谱线条数增加了。即当 β 趋于无穷大时，FM 的带宽达到极限值 $2\Delta f$ ，在以后的讨论中要牢记这一要点。

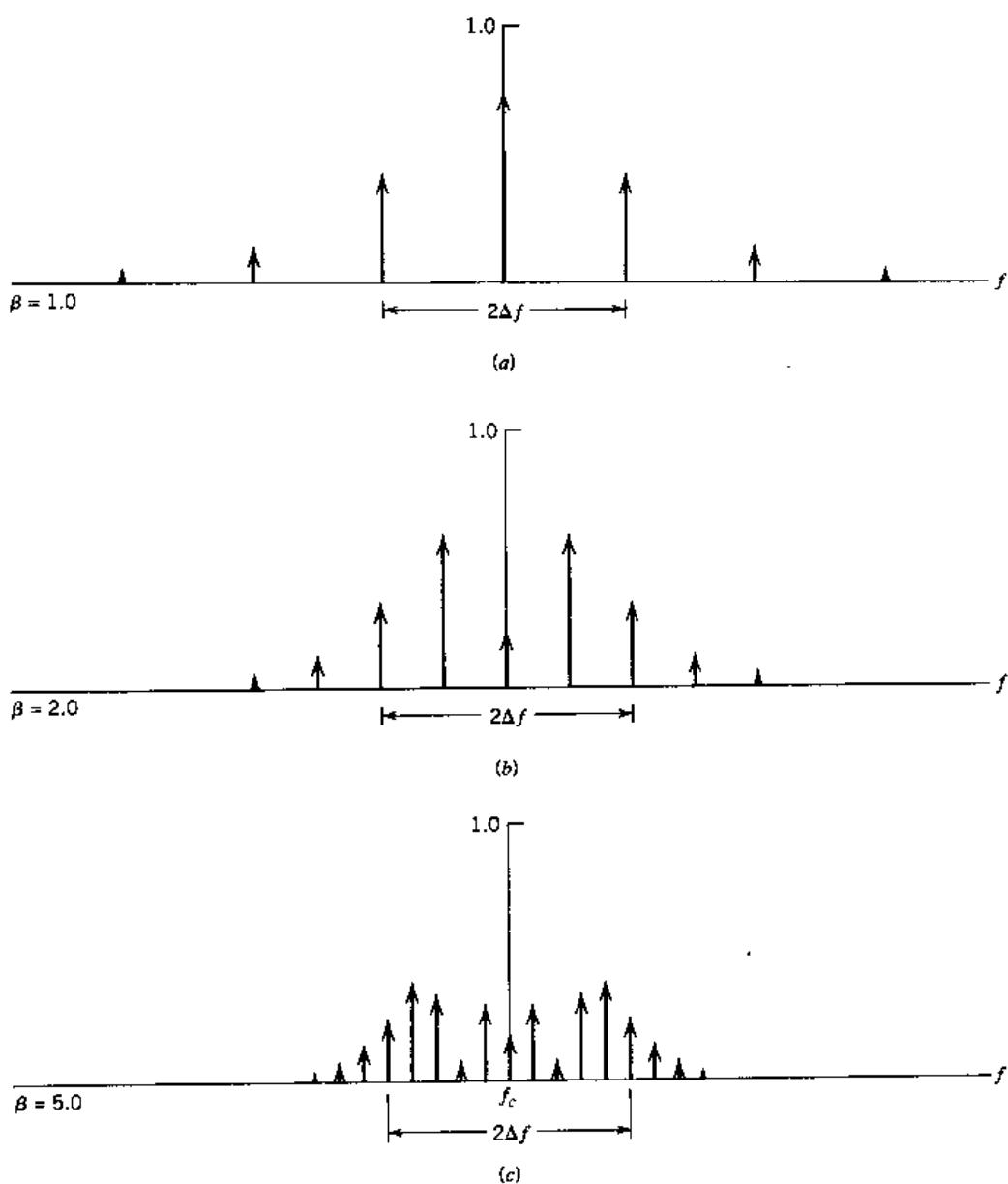


图 2.25 FM 信号的离散幅度谱(图中对载波幅度进行了归一化。其中正弦调制的幅度不变、载频变化。并且只画出了正频率部分)

2.7.3 FM 信号的传输带宽

理论上 FM 信号含有无穷多个边频,因而传输这样一个信号所需的带宽从某种程度上说是趋于无穷大的。但在实际应用中,我们发现如果允许存在一定的失真,FM 信号可被限制在有限的主要边频内。由此可计算出传输这样的 FM 信号所需的有效带宽。首先,考虑频率为 f_m 的单音频调制信号产生的 FM 信号。此 FM 信号中,当边频与载频 f_c 之间的距离大于频偏 Δf 时,边频的幅度将迅速降为 0。因此带宽通常大于整体的频偏量,但仍是有限的。特别地,当调制指数 β 很大时,带宽只需要稍大于整体频偏 $2\Delta f$,如图 2.25 所示。另一方面,当调制指数 β 很小时,FM 信号的频谱主要集中在载频 f_c 和一对边频 $f_c \pm f_m$ 处,因此 FM 带宽近似为 $2f_m$ 。因

此,对频率为 f_m 的单音频调制信号产生的 FM 信号的传输带宽可以定义如下近似规则:

$$B_T \approx 2\Delta f + 2f_m = 2\Delta f \left(1 + \frac{1}{\beta}\right) \quad (2.55)$$

该经验关系为卡逊准则^[4](Carson's rule)。

还有一种计算 FM 信号带宽需求的方法,即定义的带宽必须尽量保留最大数量的主要边频数,这些边频的幅度超过某一指定值。该幅度值可以方便地选取为未调载波幅度的 1%。从而可将 FM 信号的传输带宽定义为两个频率点之间的间隔,在这两个频率点外,所有边频的幅度均小于未调载波幅度的 1%。因此定义传输带宽为 $2n_{max}f_m$, f_m 为调制频率, n_{max} 为满足条件 $|J_n(\beta)| > 0.01$ 的最大整数 n 。 n_{max} 值随调制指数 β 变化,并可从贝塞尔函数表 $J_n(\beta)$ 直接读取。表 2.2 为不同 β 值对应的主要的边频数量(包括上边频和下边频),这里的计算采用上述的 1% 值为基准。采用这种方法算出的传输带宽 B_T 值可用标准曲线表示,即将 B_T 按频偏 Δf 归一化后,作关于 β 的曲线。由表 2.2 中的一系列点经过最佳拟合后可得到如图 2.26 所示的曲线。由图 2.26 可知,当调制指数 β 增大时,主要边频所占的带宽向实际频偏值方向减小。这说明调制指数 β 很小时占用的传输带宽大于 β 很大时占用的传输带宽。

表 2.2 不同调制指数下宽带 FM 信号的主要边频数量

调制指数 β	主要边频数量 $2n_{max}$
0.1	2
0.3	4
0.5	4
1.0	6
2.0	8
5.0	16
10.0	28
20.0	50
30.0	70

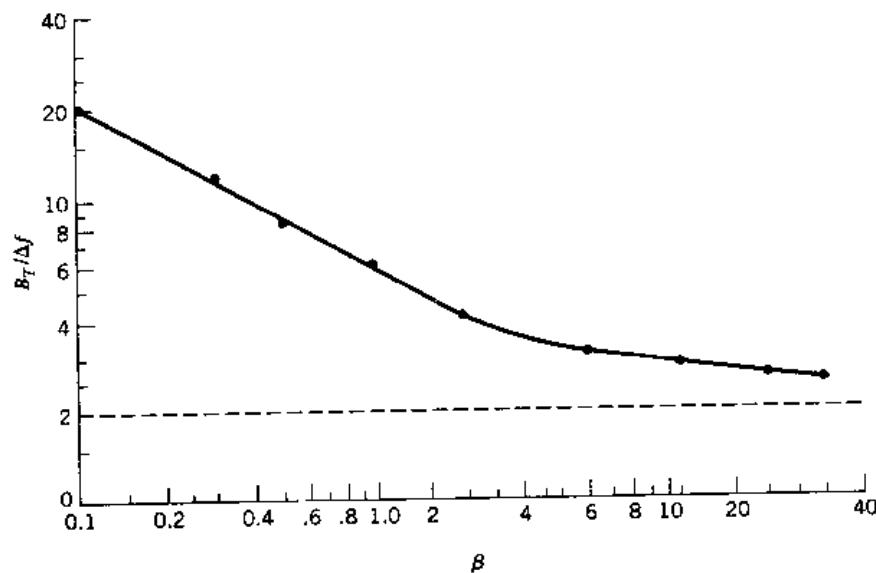


图 2.26 计算 FM 信号带宽的一般曲线(边频幅度大于载频幅度的 1%)

接下来,考虑更为普遍的最高频率分量为 W 的任意调制信号 $m(t)$ 。使用最大单频调制分析来估算该调制信号产生的 FM 信号的传输带宽。首先,将频偏 Δf 与最高调制频率 W 的比值定义为偏移率 D ,并确定 D 的值,其中 Δf 对应于调制信号 $m(t)$ 的最大可能幅度。这些条件代表了可能存在的极限情况。在非正弦调制中偏移率 D 的作用与正弦调制中的调制指数 β 相同。接下来用 D 代替 β , W 代替 f_m ,则可用式(2.55)的卡逊准则或图 2.26 的一般曲线来计算 FM 信号的传输带宽。从实际看,由卡逊准则得到的带宽小于 FM 系统需要的带宽,而由图 2.26 的标准曲线得到的值为保守值(即大于 FM 系统所需带宽)。因此,大多数的实际应用可以接受的传输带宽是位于由这两个准则得到的边界之间的值。

例 2.3

在北美,对于商用 FM 无线广播,其频偏 Δf 最大值为 75 kHz,若取调制频率 $W = 15$ kHz,也就是 FM 传输中音频的“最大值”,那么相应的偏移率为

$$D = \frac{75}{15} = 5$$

根据式(2.55)的卡逊准则,用 D 代替 β , W 代替 f_m ,得到 FM 信号的传输带宽的近似值为

$$B_T = 2 \times (75 + 15) = 180 \text{ kHz}$$

另一方面,由图 2.26 的曲线可得到 FM 信号的传输带宽为

$$B_T = 3.2\Delta f = 3.2 \times 75 = 240 \text{ kHz}$$

在实际应用中,分配给每个 FM 发射机的带宽为 200 kHz。卡逊准则得出的带宽比该带宽小了 10%,而由图 2.26 中的标准曲线得到的带宽比该值大了 20%。

2.7.4 FM 信号的产生

产生调频信号有两种基本的方法,即直接 FM 和间接 FM。在直接 FM 中,载频直接随输入基带信号的变化而变化,这由压控振荡器即可实现。在间接 FM 里,首先用调制信号产生窄带 FM 信号,然后经过倍频达到期望的频偏值。当主要考虑的是载频的稳定性时(如在商用无线广播中),间接 FM 方法是比较好的选择。下面就讨论间接 FM。

间接 FM^[5]

间接 FM 系统的简化框图如图 2.27 所示。首先对消息(基带)信号 $m(t)$ 进行积分,然后用它来对一个晶体振荡器进行相位调制,使用晶体控制的目的是提高频率稳定性。为了减小相位调制器的失真,相偏的最大值或调制指数 β 应取较小值,从而得到窄带 FM 信号。至于窄带相位调制器的实现,可参考图 2.21 中的配置。接下来,通过倍频器对窄带 FM 信号进行倍频,从而得到所希望的宽带 FM 信号。

倍频器由一个非线性器件和带通滤波器组成,如图 2.28 所示。无记忆非线性器件的含义是没有存储能量的元件。于是一个器件的输入 - 输出关系可一般地表示为

$$v(t) = a_1 s(t) + a_2 s^2(t) + \cdots + a_n s^n(t) \quad (2.56)$$

式中, a_1, a_2, \dots, a_n 是由器件的工作点决定的系数, n 为非线性的最高阶数,换言之,无记忆非线性器件是 n 次方器件,输入 FM 信号 $s(t)$ 的表达式为

$$s(t) = A_c \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right]$$

该信号的瞬时频率为

$$f_i(t) = f_c + k_f m(t) \quad (2.57)$$

图 2.28 中带通滤波器的中心频率为 nf_c , f_c 为输入 FM 信号 $s(t)$ 的载频。带通滤波器的带宽等于 $s(t)$ 传输带宽的 n 倍。在 2.8 节讨论 FM 系统的非线性影响时, 我们将这类非线性项对频谱的影响表示成了式(2.56)的输入 - 输出关系式中的二阶和三阶项。可见非线性器件输出 $v(t)$ 经过带通滤波后, 得到一个新的 FM 信号, 表达式为

$$s'(t) = A'_c \cos \left[2\pi n f_c t + 2\pi n k_f \int_0^t m(\tau) d\tau \right] \quad (2.58)$$

该信号的瞬时频率为

$$f'_i(t) = nf_c + nk_f m(t) \quad (2.59)$$

比较式(2.59)和式(2.57)可看到, 图 2.28 中的非线性电路起到了倍频的作用, 倍频倍数由式(2.56)的输入 - 输出关系中的最高次幂 n 决定。倍频倍数也是无记忆非线性器件的一个特性。

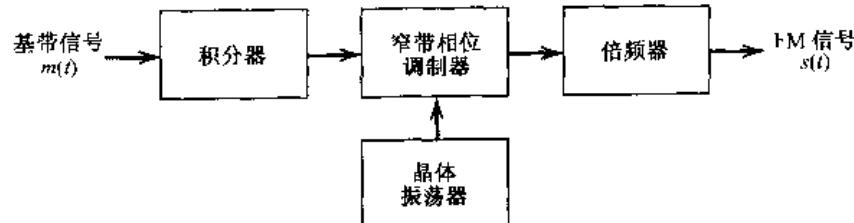


图 2.27 间接产生宽带 FM 信号的框图

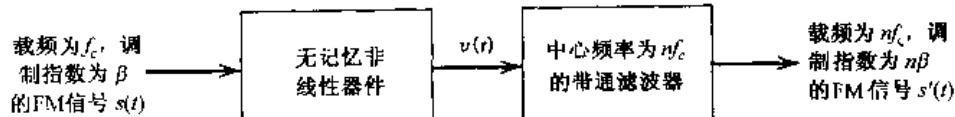


图 2.28 倍频器框图

2.7.5 FM 信号的解调

频率解调是从 FM 信号中恢复出原始调制信号的过程。解调的目的是产生与频率调制器相反的效果。解调可分为直接和间接两种。这里介绍一种直接解调方法, 该方法采用了鉴频器, 鉴频器的瞬时输出幅度与输入 FM 信号的瞬时频率成正比。在 2.14 节中将介绍另一种应用锁相环的间接解调方法。

一般来说, 鉴频器由微分电路和包络检波器级联而成。理想的微分电路的特点是, 它的频率响应是纯虚数的, 在指定的频率间隔内, 响应与频率呈线性关系变化。考虑图 2.29(a)中的频率响应, 其定义为

$$H_1(f) = \begin{cases} j2\pi a \left(f - f_c + \frac{B_T}{2} \right), & f_c - \frac{B_T}{2} \leq f \leq f_c + \frac{B_T}{2} \\ j2\pi a \left(f + f_c - \frac{B_T}{2} \right), & -f_c - \frac{B_T}{2} \leq f \leq -f_c + \frac{B_T}{2} \\ 0, & \text{其他} \end{cases} \quad (2.60)$$

式中, a 为常数。当微分电路的输入为载频 f_c , 传输带宽为 B_T 的 FM 信号 $s(t)$, 我们希望得到的是该微分电路的输出响应 $s_1(t)$ 。假定频率间隔 $f_c - B_T/2 \leq |f| \leq f_c + B_T/2$ 以外的 $s(t)$ 的频谱值为零。

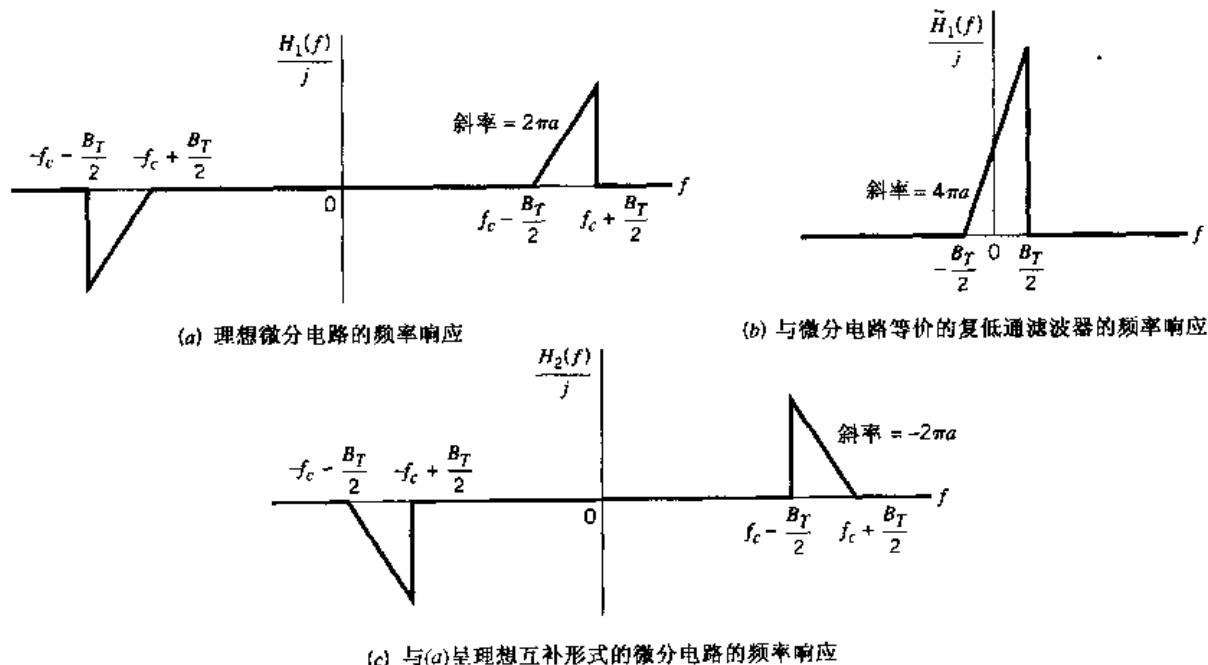


图 2.29 三种频率响应

可以利用实值带通滤波器和相应复值低通滤波器之间的同构性来简化对鉴频器的分析。附录 2 中讨论了同构性。根据附录所述, 可用频率响应为 $\tilde{H}_1(f)$ 的低通滤波器代替频率响应为 $H_1(f)$ 的带通滤波器, 这可以通过以下两步完成:

1. 将 $\tilde{H}_1(f)$ 向右搬移 f_c , f_c 为带通滤波器的中心频率。搬移后低通滤波器的频率响应和带通滤波的频率响应相等。
2. 在 $f > 0$ 时, 令 $\tilde{H}_1(f - f_c) = 2H_1(f)$ 。

即对于当前的问题, 有 $\tilde{H}_1(f - f_c) = 2H_1(f), f > 0$ (2.61)

由式(2.60)和式(2.61)可得

$$\tilde{H}_1(f) = \begin{cases} j4\pi a \left(f + \frac{B_T}{2} \right), & -\frac{B_T}{2} \leq f \leq \frac{B_T}{2} \\ 0, & \text{其他} \end{cases} \quad (2.62)$$

对应曲线如图 2.29(b)所示。

输入 FM 信号 $s(t)$ 由式(2.26)定义,为便于阅读,重写如下:

$$s(t) = A_c \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right]$$

假设载频 f_c 高于 FM 信号 $s(t)$ 的传输带宽,则 $s(t)$ 的复包络为

$$\tilde{s}(t) = A_c \exp \left[j2\pi k_f \int_0^t m(\tau) d\tau \right] \quad (2.63)$$

用 $\tilde{s}_1(t)$ 代表图 2.29(a)所示微分电路中关于 $\tilde{s}(t)$ 响应的复包络。根据附录 2 所述, $\tilde{s}_1(t)$ 的傅里叶变换表达式如下:

$$\begin{aligned} \tilde{S}_1(f) &= \frac{1}{2} \tilde{H}_1(f) \tilde{S}(f) \\ &= \begin{cases} j2\pi a \left(f + \frac{B_T}{2} \right) \tilde{S}(f), & -\frac{B_T}{2} \leq f \leq \frac{B_T}{2} \\ 0, & \text{其他} \end{cases} \end{aligned} \quad (2.64)$$

其中, $\tilde{S}(f)$ 为 $\tilde{s}(t)$ 的傅里叶变换。由傅里叶分析可知,将信号经傅里叶变换后的结果乘以 $j2\pi f$,就相当于将信号在时域中进行了平移,见表 A6.2 的第 8 项。因此,由式(2.64)可推导出

$$\tilde{s}_1(t) = a \left[\frac{d\tilde{s}(t)}{dt} + j\pi B_T \tilde{s}(t) \right] \quad (2.65)$$

将式(2.63)代入(2.65)有

$$\tilde{s}_1(t) = j\pi B_T a A_c \left[1 + \frac{2k_f}{B_T} m(t) \right] \exp \left[j2\pi k_f \int_0^t m(\tau) d\tau \right] \quad (2.66)$$

于是微分电路的响应为

$$\begin{aligned} s_1(t) &= \operatorname{Re}[\tilde{s}_1(t) \exp(j2\pi f_c t)] \\ &= \pi B_T a A_c \left[1 + \frac{2k_f}{B_T} m(t) \right] \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau + \frac{\pi}{2} \right] \end{aligned} \quad (2.67)$$

信号 $s_1(t)$ 是一个载波幅度和频率均随消息信号 $m(t)$ 变化的复合调制信号。然而,若选择

$$\left| \frac{2k_f}{B_T} m(t) \right| < 1 \quad \text{对于所有 } t$$

则可用包络检波器恢复出幅度的变化量,进而恢复出原始消息信号,但存在偏差项。包络检波器的输出为

$$|\tilde{s}_1(t)| = \pi B_T a A_c \left[1 + \frac{2k_f}{B_T} m(t) \right] \quad (2.68)$$

上式右端的偏差项 $\pi B_T a A_c$ 与微分电路的转移函数的斜率 a 成正比。

这表明将第一个包络检波器的输出 $|\tilde{s}_1(t)|$ 与第二个包络检波器的输出相减,就可去掉该偏差项。位于第二个包络检波器之前的电路是频率响应为 $H_2(f)$ 的如图2.29(c)所示的互补微分电路。这两个微分电路之间的关系式为

$$\tilde{H}_2(f) = \tilde{H}_1(-f) \quad (2.69)$$

令 $s_2(t)$ 表示输入FM信号 $s(t)$ 的互补微分电路的响应。与前面同理,有

$$|\tilde{s}_2(t)| = \pi B_T a A_c \left[1 - \frac{2k_f}{B_T} m(t) \right] \quad (2.70)$$

其中, $\tilde{s}_2(t)$ 为信号 $s_2(t)$ 的复包络。式(2.68)和式(2.70)的包络之差为

$$\begin{aligned} s_o(t) &= |\tilde{s}_1(t)| - |\tilde{s}_2(t)| \\ &= 4\pi k_f a A_c m(t) \end{aligned} \quad (2.71)$$

该式与原始的消息信号 $m(t)$ 成正比且无偏差项。

从而得到了理想鉴频器的模型如图2.30所示,它由一对微分电路、包络检波器以及最后的加法器组成。这对微分电路的复转移函数之间的关系式为式(2.69)。该方案称为平衡式鉴频器。

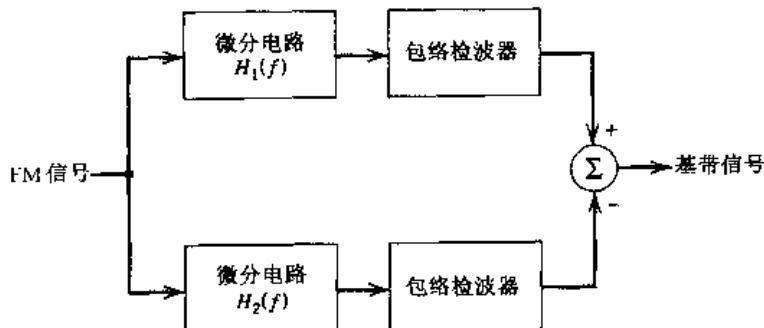


图2.30 鉴频器框图

2.7.6 FM立体声复用^[6]

立体声复用是使用同一载波来传输两个独立信号的一种频分复用(FDM)技术。在FM无线广播中,它被广泛地用来发送一个节目的两个不同部分(例如管弦乐队的两部分:一个小提琴部分和一个伴奏部分),这样会给接收端听众带来一种空间感。

FM立体声技术标准受两个因素影响:

1. 必须在分配的FM广播信道中传输。
2. 必须与单声道无线接收机兼容。

第一个要求设定了允许的频率参数值,包括频偏。第二个要求限制了传输信号的配置方式。

图2.31(a)为FM立体声发射机中复用系统的框图。令 $m_l(t)$ 和 $m_r(t)$ 分别表示系统发射端左侧和右侧的麦克风产生的信号。

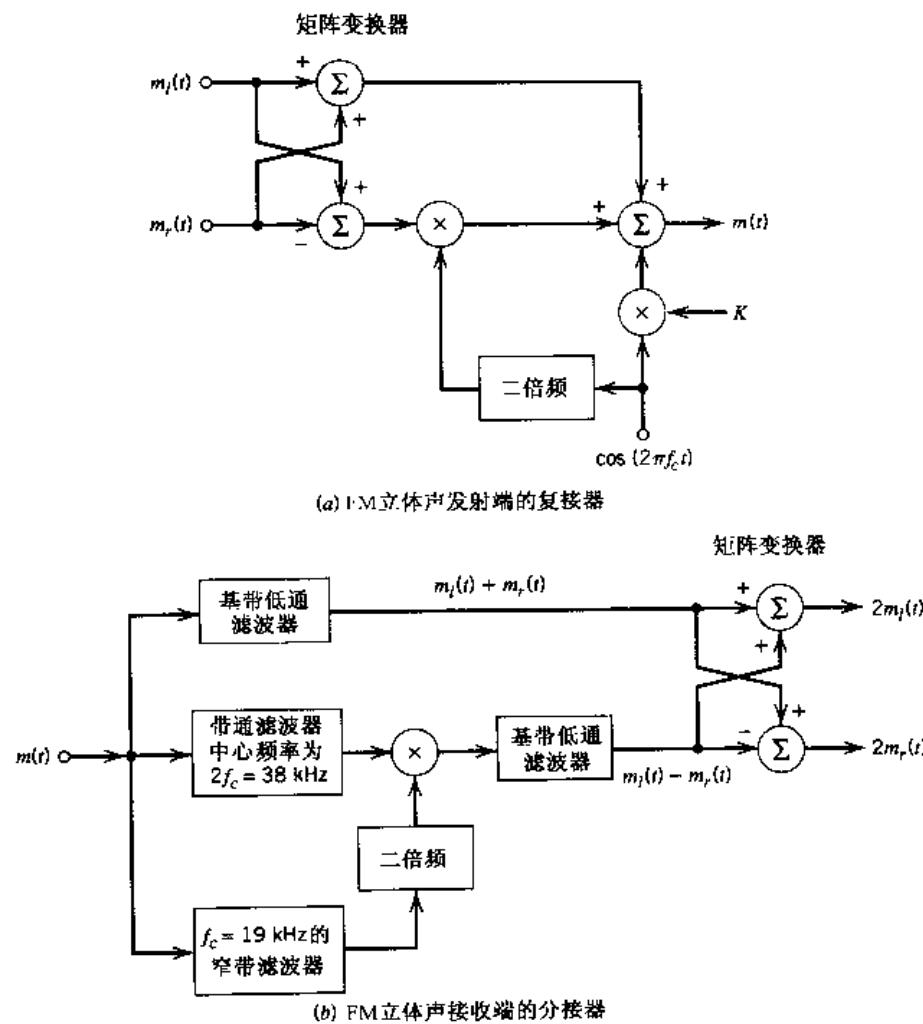


图 2.31 FM 立体声复用

将这两个信号通过简单的矩阵变换电路, 可产生和信号 $m_l(t) + m_r(t)$ 和差信号 $m_l(t) - m_r(t)$ 。和信号仍是一种基带形式, 这种形式适于单声道接收。将差信号与 38 kHz 的副载波(由 19 kHz 的晶体振荡器经倍频得到)通过乘积调制器, 以产生 DSB-SC 已调信号。除了和信号与 DSB-SC 已调信号, 复用信号 $m(t)$ 也包含 19 kHz 的导频信号, 从而为立体声接收机对差信号的相干检测提供了参考频率。因而复用信号可表示为

$$m(t) = [m_l(t) + m_r(t)] + [m_l(t) - m_r(t)] \cos(4\pi f_c t) + K \cos(2\pi f_c t) \quad (2.72)$$

其中 $f_c = 19 \text{ kHz}$, K 为导频的幅度。复用信号 $m(t)$ 对主载波进行频率调制后, 产生了所要传输的信号。导频信号应位于频偏峰值的 8% 与 10% 之间, 式(2.72)中幅度 K 的选择必须符合这个要求。

在立体声接收机中对输入 FM 信号进行频率解调, 恢复出复用信号 $m(t)$ 。 $m(t)$ 被送入图 2.31(b)所示的分离系统, 通过三次适当的滤波得到复用信号 $m(t)$ 的各个频率分量。恢复出的导频信号(用 19 kHz 的窄带滤波器)经一次倍频后得到 38 kHz 的副载波, 用于 DSB-SC 已调信号的相干检波, 从而恢复出差信号 $m_l(t) - m_r(t)$ 。图 2.31(b)上支路的基带低通滤波器可使和信号 $m_l(t) + m_r(t)$ 通过。最后由简化的矩阵变换电路重构出存在比例关系的左侧信号 $m_l(t)$ 和右侧信号 $m_r(t)$, 并将 $m_l(t)$ 和 $m_r(t)$ 分别送至各自的扬声器。

2.8 FM系统中的非线性影响

前两节里我们研究了频率调制的原理以及FM信号的产生和解调的方法。为完整起见，下面我们对FM系统中非线性的影响进行讨论。

非线性在所有的通信网络中均存在，这里考虑非线性的两种基本形式：

1. 由于某些特殊应用而有意引入并加以控制的非线性。这种非线性一般会很强。强非线性的例子包括平方律调制器、硬限幅器和倍频器。
2. 很多情况下设计者希望保持系统的线性。但由于系统的不完善，容易引起寄生的非线性。这时的非线性通常较弱。系统中的弱非线性将对有用信号造成干扰，因此在设计时应尽量避免。

本节讨论频率调制中弱非线性的影响。

考虑一个传输特性由如下非线性输入-输出关系定义的通信信道：

$$v_o(t) = a_1 v_i(t) + a_2 v_i^2(t) + a_3 v_i^3(t) \quad (2.73)$$

其中， $v_i(t)$ 和 $v_o(t)$ 分别为输入和输出信号， a_1 ， a_2 和 a_3 为常量。上式为频率调制中用到的式(2.56)的截短式，其描述的信道是无记忆的，因为输出信号 $v_o(t)$ 是输入信号 $v_i(t)$ 的瞬时函数（即描述中没有能量存储单元）。此处用这样的信道来传输 FM 信号。下面讨论 FM 系统中的非线性影响。FM 信号的定义为

$$v_i(t) = A_c \cos[2\pi f_c t + \phi(t)]$$

其中

$$\phi(t) = 2\pi k_f \int_0^t m(\tau) d\tau$$

对于该输入信号，由式(2.73)得

$$\begin{aligned} v_o(t) &= a_1 A_c \cos[2\pi f_c t + \phi(t)] + a_2 A_c^2 \cos^2[2\pi f_c t + \phi(t)] \\ &\quad + a_3 A_c^3 \cos^3[2\pi f_c t + \phi(t)] \end{aligned} \quad (2.74)$$

将式(2.74)中余弦函数的平方项和立方项展开，并合并同类项，得

$$\begin{aligned} v_o(t) &= \frac{1}{2} a_2 A_c^2 + \left(a_1 A_c + \frac{3}{4} a_3 A_c^3 \right) \cos[2\pi f_c t + \phi(t)] \\ &\quad + \frac{1}{2} a_2 A_c^2 \cos[4\pi f_c t + 2\phi(t)] \\ &\quad + \frac{1}{4} a_3 A_c^3 \cos[6\pi f_c t + 3\phi(t)] \end{aligned} \quad (2.75)$$

因此，信道输出由一个直流分量和三个载频分别为 f_c ， $2f_c$ 和 $3f_c$ 的调频信号分量组成，这三个调频信号分量分别由式(2.73)中的线性项、二次项和三次项产生。

为了从信道输出 $v_o(t)$ 中得到 FM 信号（即载频为 f_c 的特殊分量），有必要将 FM 信号中载

频为 f_c 的分量与最接近载频 $2f_c$ 的分量进行分离。令 Δf 表示输入 FM 信号 $v_i(t)$ 的频偏, W 表示消息信号 $m(t)$ 的最高频率分量。应用卡逊准则(需注意载波二次谐波的频偏应为原来的 2 倍), 则将载频分别为 f_c 和 $2f_c$ 的 FM 信号进行分离的必要条件就是

$$2f_c - (2\Delta f + W) > f_c + \Delta f + W$$

或

$$f_c > 3\Delta f + 2W \quad (2.76)$$

于是经过带宽为 $2\Delta f + 2W$ 、中心频率为 f_c 的带通滤波器后, 信道的输出信号简化为

$$v_o'(t) = \left(a_1 A_c + \frac{3}{4} a_3 A_c^3 \right) \cos[2\pi f_c t + \phi(t)] \quad (2.77)$$

由此可见, FM 信号经过幅度非线性的信道与适当的滤波器后, 受到的惟一影响就是信号幅度发生了改变。即不同于幅度调制, 频率调制经过幅度非线性的信道传输后并没有发生失真现象。正是由于这个原因, 微波无线系统采用了频率调制, 允许使用高阶非线性放大器和功率发射机。这一点对于需要大功率输出的无线信道来说尤为重要。

从直观上看, FM 系统对于相位非线性极其敏感。微波无线系统中存在一种普遍的相位非线性, 称为 AM-PM 转换。造成这种非线性的原因是系统内中继器或放大器的相位依赖于输入信号的瞬时幅度。在实际应用中, AM-PM 转换用常量 K 表示, 单位为 $(^\circ)/dB$ (度每分贝), 其含义为输入端的包络发生 $1 dB$ 的改变时, 对应的输出相位的最大改变量。当 FM 信号经微波无线链路传输时, 由于传输过程中存在噪声和干扰, FM 信号的幅度发生了变化, 这样的 FM 信号通过具有 AM-PM 转换功能的中继器后, 将使输出端包含无用的相位调制和合成失真。因此, 将 AM-PM 转换值保持在较低的水平是极其重要的。例如, 对于较好的微波中继器, AM-PM 转换常数 K 的值应小于 $2^\circ/dB$ 。

2.9 超外差接收机^[7]

在广播系统中, 无论采用的是幅度调制还是频率调制, 接收机都要完成对输入已调信号的解调工作, 以及其他一些系统功能如:

- 载频调谐, 这是为了选择期望信号(即期望的无线广播电台或电视台)。
- 滤波, 通过滤波可将期望的信号与其他已调信号区分开。
- 放大, 能够补偿传输过程中的功率损失。

超外差接收机是具备上述三种功能的一种特殊类型的接收机, 它能够简单方便地完成前两种功能。特别地, 在这种接收机中不需要构造可调谐的、具有高选择性的可变滤波器。实际上, 现在所有的无线广播和电视接收机采用的都是这种超外差接收机。

超外差接收机的基本组成部分包括射频(RF)部分、混频器和本地振荡器、中频(IF)部分、解调器和功率放大器。商用 AM 和 FM 无线接收机的典型频率参数见表 2.3。图 2.32 为超外差接收机的框图, 其中采用包络检波来解调 AM 信号。

输入 AM 信号由接收天线接收, 再经射频部分放大, 其中射频部分已调谐到输入信号载频上。混频器与本地振荡器(可调频)共同完成外差功能, 因此可将输入信号转换到预先给定的

中频上,这个中频通常低于输入载频。这样的频率搬移不会破坏载波与边带间的关系(见2.4节)。外差的结果是得到定义如下的中频信号:

$$f_{IF} = f_{LO} - f_{RF} \quad (2.78)$$

其中, f_{LO} 为本地振荡器的频率, f_{RF} 为输入射频信号的载频。之所以称 f_{IF} 为中频(IF)是因为 f_{IF} 既不在原始输入频率上,也不在最终的基频上。混频器与本地振荡器有时合称为第一检测器,此时将解调器称为第二检测器。

表 2.3 AM 和 FM 无线接收机的典型频率参数

	AM 无线接收机	FM 无线接收机
射频载波范围	0.535 ~ 1.605 MHz	88 ~ 108 MHz
中频部分的中心频率	0.455 MHz	10.7 MHz
中频带宽	10 kHz	200 kHz

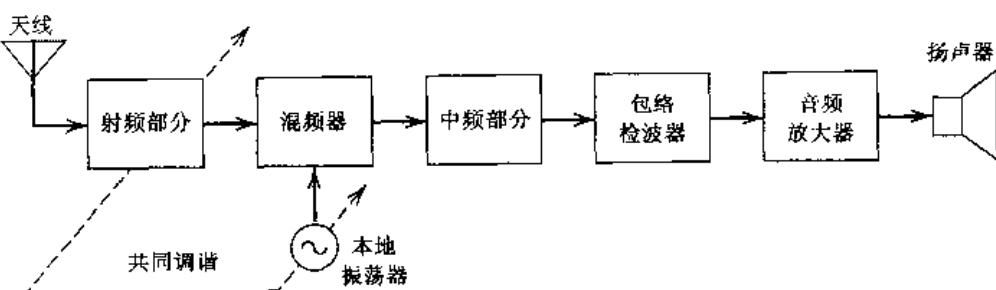


图 2.32 超外差 AM 无线接收机的基本组成

中频部分由一级或多级调谐式放大器组成,带宽等于接收机相应调制类型所需的带宽。中频部分完成接收机里大部分的放大和选择工作。将中频部分的输出作为解调器的输入,从而恢复出基带信号。若采用相干检测,接收机必须提供相干信号源。接收机最后的工作是对已恢复的消息信号进行功率放大。

当输入信号频率与本地振荡器频率之差等于中频时,超外差接收机中的混频器将产生中频信号输出。也就是说,两个输入频率 $|f_{LO} \pm f_{IF}|$ 将使混频器输出为 f_{IF} 。这使得同时接收的两个信号的频差有可能等于中频的两倍。例如,将接收机调至 0.65 MHz, 中频为 0.455 MHz, 则镜像干扰频率为 1.56 MHz。实际上,任何中频为 0.455 MHz 的接收机,当调谐在任一电台时,都会受到频率比该电台频率高 0.910 MHz 的镜像干扰。由于混频器的功能是得到两个频率的差频,因此它不能将有用信号与镜像信号区分开,而只能取有用信号与镜像信号中的一个作为中频输出信号。避免镜像干扰的惟一方法是在射频部分(即天线与混频器之间)置入高选择性器件,达到使有用信号通过并抑制无用或镜像信号的目的。接收机抑制无用或镜像信号的能力随射频部分选择阶数的增加或中频与信号频率比值的增大而提高。

AM 和 FM 超外差接收机基本的不同之处在于后者使用了 FM 解调器,如限幅鉴频器。FM 系统通过正弦载波瞬时频率的变化传输消息信号,正弦载波幅度恒定不变。因而接收机输入端载波信号的幅度波动的原因必定是噪声或干扰。在中频部分之后置入限幅器,通过将中频部分输出的已调信号幅度限制在零轴附近,可消除幅度波动。通过带通滤波器抑制载频的谐波可达到对输出矩形波进行整形的目的。此时滤波器的输出为正弦信号,其幅度与接收机输入载波的幅度无关(见习题 2.42)。

2.10 连续波调制系统中的噪声

前面部分所进行的讨论完全是从定性角度来研究连续波调制技术的特点。本章后面的几节将讨论信道噪声对连续波已调信号接收的影响,从而使读者对模拟通信过程有更深入的理解。

按照惯例,在讨论的开始我们先给出两个模型:

1. **信道模型**。假定一种通信信道是无失真的,但存在加性高斯白噪声(AWGN)的干扰。
2. **接收机模型**。假定接收机由理想的带通滤波器和理想的解调器级联而成。带通滤波器的作用是减小信道噪声的影响。

这些简化性假设的目的是让我们对噪声如何影响接收机性能有一个基本的理解。这些假设同时也为比较各种连续波调制-解调方案提供了框架。

图 2.33 为基于以上两个假设的噪声接收机模型。图中 $s(t)$ 表示输入已调信号, $w(t)$ 表示信道噪声。因此接收信号为 $s(t)$ 和 $w(t)$ 之和。图 2.33 所示模型中的带通滤波器代替调谐放大器起到了滤波的作用, 调谐放大器的目的是在解调之前将信号放大。带通滤波器的带宽刚好可以使已调信号 $s(t)$ 无失真地通过。图 2.33 模型中解调器的具体细节要根据所采用的调制类型来确定。

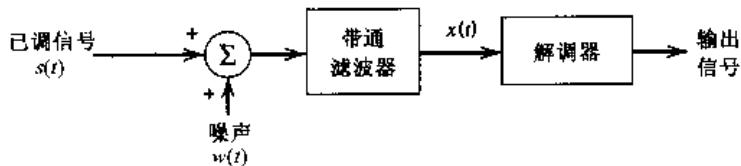


图 2.33 接收机模型

2.10.1 信噪比: 基本定义

定义噪声 $w(t)$ 的功率谱密度在正频和负频上均为 $N_0/2$, 其中 N_0 等于接收机前端测得的单位带宽内的平均噪声功率。假定图 2.33 接收机模型中的带通滤波器是理想的, 其带宽等于已调信号 $s(t)$ 的传输带宽 B_T , 中心频率等于载频 f_c 。该假定只适用于抑制载波的双边带调制(DSB-SC), 标准幅度调制(AM)和频率调制(FM); 对于单边带(SSB)调制和残留边带(VSB)调制需要进行特殊分析。取带通滤波器的中心频率值为载频 f_c , 画出噪声 $n(t)$ 的功率密度谱 $S_N(f)$, 如图 2.34 所示, 其中 $n(t)$ 为白噪声 $w(t)$ 通过滤波器后的噪声。

通常载频 f_c 大于传输带宽 B_T , 因此可将滤波后的噪声 $n(t)$ 看成窄带噪声, 其典型表达式为

$$n(t) = n_I(t) \cos(2\pi f_c t) + n_Q(t) \sin(2\pi f_c t) \quad (2.79)$$

其中, $n_I(t)$ 为噪声的同相分量, $n_Q(t)$ 为噪声的正交分量, 这两个分量的单位与载波 $A_c \cos(2\pi f_c t)$ 相同。滤波后用于解调的信号 $x(t)$ 定义如下:

$$x(t) = s(t) + n(t) \quad (2.80)$$

其中 $s(t)$ 由采用的调制类型决定。在任何情况下,解调器输入端的平均噪声功率均等于功率谱密度曲线 $S_N(f)$ 下方的总面积。从图 2.34 可以很容易地看出,平均噪声功率等于 $N_0 B_T$ 。只要给定了 $s(t)$ 的形式,就可以算出解调器输入信号的平均功率。根据式(2.80),已解调信号 $s(t)$ 与滤波后的噪声 $n(t)$ 在解调器输入端是加性的,因此可定义输入信噪比(SNR)_i,
(SNR)_i 等于已调信号 $s(t)$ 的平均功率与滤波后的噪声 $n(t)$ 的平均功率之比。

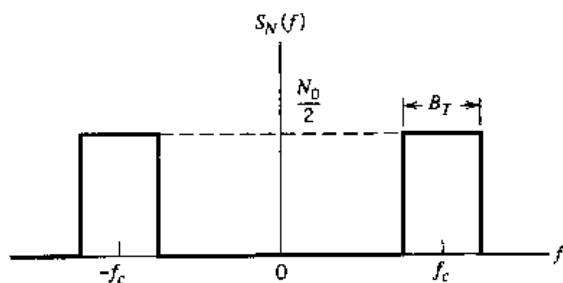


图 2.34 带通滤波后噪声的理想特性

另一个衡量噪声性能的有用参量是输出信噪比(SNR)_o,其定义为接收机输出端测得的解调后的消息信号的平均功率与噪声平均功率之比。输出信噪比能够直观地衡量解调过程的逼真度。为了准确定义信噪比,恢复出来的消息信号与噪声分量在解调器输出端必须是加性的。当接收机采用相干检测时,这一条件完全满足。此外,当接收机在标准 AM 中采用包络检波或在 FM 中采用鉴频器时,必须假定滤波后噪声 $n(t)$ 的平均功率较小,才能用输出信噪比来衡量接收机性能的好坏。

在其他因素中,输出信噪比由发射机采用的调制类型和接收机采用的解调类型决定。因此,比较各种调制-解调系统的输出信噪比是很有意义的,但必须在下述的相同基础上进行比较:

- 每一系统传输的已调信号 $s(t)$ 具有相同的平均功率。
- 在消息带宽 W 内,信道噪声 $w(t)$ 有相同的平均功率。

相应地,理论上可以定义信道信噪比(SNR)_c 为接收机输入端测得的,已调信号的平均功率与消息带宽内信道噪声的平均功率之比。该定义说明见图 2.35。

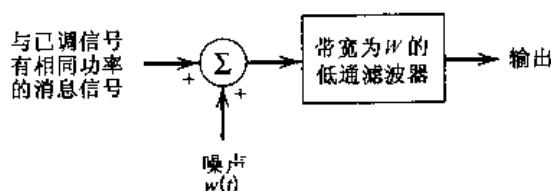


图 2.35 消息信号带宽为 W 的基带传输模型,用于计算信道信噪比

为了比较不同连续波调制系统,要对接收机性能进行归一化,即用输出信噪比除以信道信噪比。定义接收机的解调增益如下

$$\text{解调增益} = \frac{(\text{SNR})_o}{(\text{SNR})_c} \quad (2.81)$$

显然,解调增益越高,说明接收机的抗噪声性能越好。对于不同的调制类型,解调增益可能等于 1,小于 1 或者大于 1,下面的讨论会说明这一点。

2.11 相干检测线性接收机中的噪声

在 2.2 节和 2.3 节中我们了解到 AM 信号的解调方法应根据是否抑制了载波而决定。当抑制载波时, 应采用相干检测, 此时接收机是线性的。而当幅度调制中包含了载波时, 解调只需采用包络检波即可, 此时接收机为非线性的。本节研究噪声对线性接收机的影响。非线性接收机情况见 2.12 节。

图 2.36 是采用相干检测的 DSB-SC 接收机模型。相干检测需将滤波后的信号 $x(t)$ 与本地产生的正弦信号 $\cos(2\pi f_c t)$ 相乘, 然后经过低通滤波器。为简化分析, 假定本地产生的正弦信号幅值为 1。为了达到令人满意的解调效果, 本地振荡器与发射机中产生的载波必须为同频同相的。这里假定已实现这种同步。

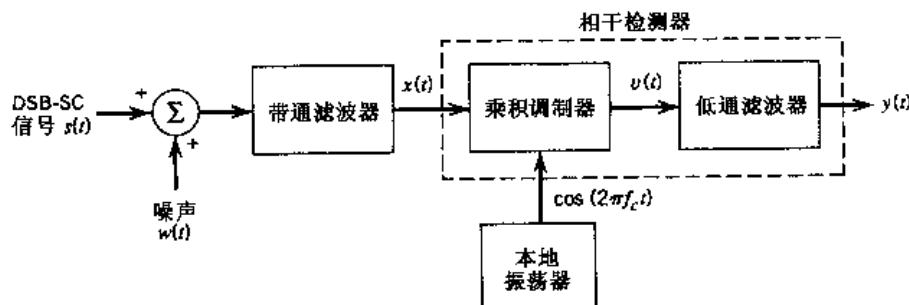


图 2.36 采用相干检测的 DSB-SC 接收机模型

滤波后信号 $x(t)$ 的 DSB-SC 分量表达式为

$$s(t) = CA_c \cos(2\pi f_c t)m(t) \quad (2.82)$$

其中, $A_c \cos(2\pi f_c t)$ 为正弦载波, $m(t)$ 为消息信号。式(2.82)中 C 为系统的比例因子, 其目的是保证信号分量 $s(t)$ 与加性噪声分量 $n(t)$ 的单位相同。

假定 $m(t)$ 是零均值平稳过程的样本函数, $m(t)$ 的功率谱密度 $S_M(f)$ 被限制在最高频率 W 以内, W 为消息带宽。消息信号的平均功率 P 等于功率谱密度曲线下的总面积, 表达式为

$$P = \int_{-W}^W S_M(f) df \quad (2.83)$$

载波与消息信号之间是统计独立的。载波应包含一个在 2π 弧度范围内均匀分布的随机相位。在 $s(t)$ 的定义式中, 为便于表达, 省略了随机相位。应用第 1 章例 1.7 的结果, 对于随机调制过程, 可将 DSB-SC 已调信号分量 $s(t)$ 的平均功率记为 $C^2 A_c^2 P / 2$ 。若噪声功率谱密度为 $N_0/2$, 则消息带宽 W 内的平均噪声功率等于 WN_0 。因而, DSB-SC 调制系统中信道信噪比为

$$(SNR)_{C, DSB} = \frac{C^2 A_c^2 P}{2 W N_0} \quad (2.84)$$

分子中的常量 C^2 确保了该比例式是无量纲的。

接下来确定系统的输出信噪比。利用滤波后噪声 $n(t)$ 的窄带表达式, 可将相干检测器的输入信号表示成

$$\begin{aligned} x(t) &= s(t) + n(t) \\ &= CA_c \cos(2\pi f_c t)m(t) + n_I(t) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t) \end{aligned} \quad (2.85)$$

其中, $n_I(t)$ 和 $n_Q(t)$ 分别为 $n(t)$ 的同相和正交分量。则相干检测器中乘积调制器的输出为

$$\begin{aligned} v(t) &= x(t) \cos(2\pi f_c t) \\ &= \frac{1}{2}CA_c m(t) + \frac{1}{2}n_I(t) \\ &\quad + \frac{1}{2}[CA_c m(t) + n_I(t)] \cos(4\pi f_c t) - \frac{1}{2}n_Q(t) \sin(4\pi f_c t) \end{aligned}$$

图 2.86 所示相干检测器中的低通滤波器滤掉了 $v(t)$ 的高频分量, 从而接收机的输出为

$$y(t) = \frac{1}{2}CA_c m(t) + \frac{1}{2}n_I(t) \quad (2.86)$$

式(2.86)表明:

1. 在接收机输出端, 消息信号 $m(t)$ 和滤波后噪声 $n(t)$ 的同相分量 $n_I(t)$ 是加性的。
2. 相干检测器完全去掉了噪声 $n(t)$ 的正交分量 $n_Q(t)$ 。

这两个结论与输入信噪比无关。因此, 相干检测区别于其他解调技术的一个重要特点是: 无论输入信噪比如何, 输出消息分量总是非乘性的, 噪声分量总是加性的。

接收机输出的消息信号分量为 $CA_c m(t)/2$ 。该分量的平均功率可表示为 $C^2 A_c^2 P/4$, 其中, P 为原始消息信号 $m(t)$ 的平均功率, C 为系统的比例因子。

在 DSB-SC 调制里, 图 2.36 中带通滤波器的带宽 B_T 等于 $2W$, 是为了使已调信号 $s(t)$ 的上边带和下边带均通过。因而, 滤波后噪声 $n(t)$ 的平均功率等于 $2WN_0$ 。从 1.11 节对窄带噪声的讨论可知, (低通) 同相噪声分量 $n_I(t)$ 的平均功率与(带通) 滤波后噪声 $n(t)$ 的平均功率相等。结合式(2.86)得到的接收机输出噪声分量 $n_I(t)/2$, 则接收机输出噪声的平均功率应为

$$(\frac{1}{2})^2 2WN_0 = \frac{1}{2}WN_0$$

对于采用相干检测的 DSB-SC 接收机, 其输出信噪比为

$$\begin{aligned} (\text{SNR})_{O, \text{DSB-SC}} &= \frac{C^2 A_c^2 P / 4}{WN_0 / 2} \\ &= \frac{C^2 A_c^2 P}{2WN_0} \end{aligned} \quad (2.87)$$

由式(2.84)和式(2.87), 得解调增益为

$$\left. \frac{(\text{SNR})_O}{(\text{SNR})_C} \right|_{\text{DSB-SC}} = 1 \quad (2.88)$$

其中输出信噪比和信道信噪比均包含因子 C^2 , 所以在计算过程中被约掉了。

对 SSB 相干检测的噪声分析表明, 尽管它与 DSB-SC 相干检测之间有根本性的差别, 但它们的解调增益却是相同的(见习题 2.49)。

由本节和习题 2.49 的讨论, 可得两条重要结论:

1. 在平均发送功率或已调信号功率与消息带宽内的平均噪声功率相等的情况下, 相干

SSB 接收机与相干 DSB-SC 接收机有着相等的输出信噪比。

2. 在上述两种情况下,接收机的噪声性能与相同信道噪声下只发送消息信号时的性能相同。调制过程对接收机性能唯一的影响是,为了便于信号在带通信道中传输,将消息信号搬移到了另一个不同的频带上。

DSB-SC 调制和 SSB 调制均无法在提高噪声性能与增加信道带宽之间提供一种折中的办法。当需要高质量的接收时,这就成了一个重要的问题。

2.12 包络检波 AM 接收机的噪声

接下来对 AM 系统中采用包络检波的接收机进行噪声分析,模型如图 2.37 所示。在标准 AM 信号里,两个边带和载波均被发送,表达式为

$$s(t) = A_c[1 + k_a m(t)] \cos(2\pi f_c t) \quad (2.89)$$

其中, $A_c \cos(2\pi f_c t)$ 为载波, $m(t)$ 为消息信号, k_a 为决定调制百分比的常量。由上式给出的调幅信号分量 $s(t)$ 的表达式可见,并不一定需要使用比例因子,因为可以假设载波幅度 A_c 与加性噪声分量的量纲相同。

在 AM 信号 $s(t)$ 中,载波分量的平均功率为 $A_c^2/2$ 。载有信息的分量 $A_c k_a m(t) \cos(2\pi f_c t)$ 的平均功率为 $A_c^2 k_a^2 P/2$,其中 P 为消息信号 $m(t)$ 的平均功率。因此,标准 AM 信号 $s(t)$ 的平均功率等于 $A_c^2(1 + k_a^2 P)/2$ 。对于 DSB-SC 系统,消息带宽内噪声的平均功率为 WN_0 。则 AM 系统的信道信噪比为

$$(SNR)_{C,AM} = \frac{A_c^2(1 + k_a^2 P)}{2WN_0} \quad (2.90)$$

为了计算输出信噪比,首先将滤波后噪声 $n(t)$ 表示为同相分量和正交分量的形式。从而可将图 2.37 所示接收机模型中加在包络检波器上的滤波后信号 $x(t)$ 定义如下:

$$\begin{aligned} x(t) &= s(t) + n(t) \\ &= [A_c + A_c k_a m(t) + n_I(t)] \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t). \end{aligned} \quad (2.91)$$

用矢量表示组成信号 $x(t)$ 的分量,如图 2.38(a)所示。由该矢量图可容易地得到接收机输出为

$$\begin{aligned} y(t) &= x(t) \text{ 的包络} \\ &= \{[A_c + A_c k_a m(t) + n_I(t)]^2 + n_Q^2(t)\}^{1/2} \end{aligned} \quad (2.92)$$

信号 $y(t)$ 为理想包络检波器的输出信号。 $x(t)$ 的相位是多少并无意义,因为理想的包络检波器对 $x(t)$ 的相位波动并不敏感。

$y(t)$ 的表达式比较复杂,需进行适当简化,以推导出有意义的结果。特别地,我们希望将输出 $y(t)$ 近似表示成消息项与噪声项之和。一般说来,要做到这一点很难。但当载波的平均功率大于噪声的平均功率时(此时接收机可正常工作),绝大部分时间内,信号项 $A_c[1 + k_a m(t)]$ 大于噪声项 $n_I(t)$ 和 $n_Q(t)$ 。由此可得输出 $y(t)$ 的近似表达式为(见习题 2.51):

$$y(t) \approx A_c + A_c k_a m(t) + n_I(t) \quad (2.93)$$

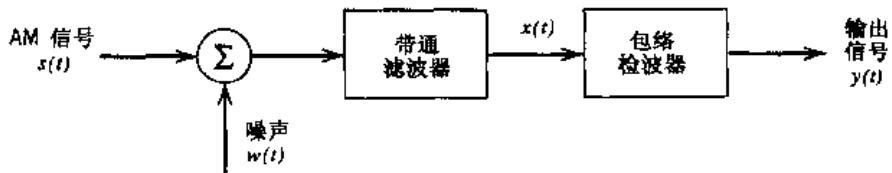


图 2.37 AM 接收机模型

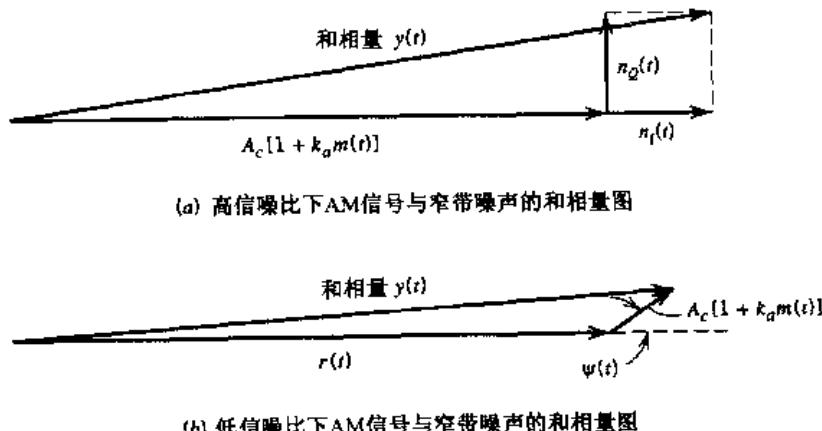


图 2.38 AM 信号与窄带噪声的和相量图

上式所示包络检波器输出 $y(t)$ 中出现的直流项(或常数项) A_0 是由于对传输的载波进行解调而产生的。该项与消息信号 $m(t)$ 无关,因此可以忽略。且在任何情况下,均可通过直流电容器去掉该项。如果忽略式(2.93)中的直流项 A_0 ,就会发现剩余项除了比例因子外,与采用相干检测的 DSB-SC 接收机输出信号的形式相类似。相应地,采用包络检波的 AM 接收机的输出信噪比近似为

$$(\text{SNR})_{\text{O,AM}} \approx \frac{A_c^2 k_a P}{2 W N_0} \quad (2.94)$$

式(2.94)只有在满足以下两个条件时才有效:

1. 在包络检波器的输入端, 平均噪声功率小于平均载波功率。
 2. 调幅灵敏度 k_a 调整到调制百分比小于或等于 100%。

由式(2.90)和式(2.94),得到幅度调制的解调增益为

$$\left. \frac{(\text{SNR})_O}{(\text{SNR})_C} \right|_{\text{AM}} \approx \frac{k_a^2 P}{1 + k_a^2 P} \quad (2.95)$$

因此,采用相干检波的 DSB-SC 接收机或 SSB 接收机的解调增益总是 1,而采用包络检波的 AM 接收机的解调增益通常小于 1。换言之,标准 AM 接收机的噪声性能总是次于 DSB-SC 接收机,原因是 AM 信号中包含载波分量,浪费了发射功率。

例 2.4 单音频调制

考虑将频率为 f_m , 幅度为 A_m 的正弦信号作为调制信号的例子, 即

$$m(t) = A_m \cos(2\pi f_m t)$$

相应的 AM 信号为

$$s(t) = A_c [1 + \mu \cos(2\pi f_m t)] \cos(2\pi f_c t)$$

其中 $\mu = k_a A_m$ 为调制因子。调制信号 $m(t)$ (假定电阻为 1Ω) 的平均功率为

$$P = \frac{1}{2} A_m^2$$

因此, 由式(2.95)可得

$$\begin{aligned} \left. \frac{(\text{SNR})_O}{(\text{SNR})_C} \right|_{\text{AM}} &= \frac{\frac{1}{2} k_a^2 A_m^2}{1 + \frac{1}{2} k_a^2 A_m^2} \\ &= \frac{\mu^2}{2 + \mu^2} \end{aligned} \quad (2.96)$$

当 $\mu = 1$, 即对应的调制百分比为 100% 时, 得出解调增益等于 $1/3$ 。这说明其他因素均相同的情况下, 若想达到相同的噪声性能, AM 系统(采用包络检波)发送的平均功率应为抑制载波系统(采用相干检波)的 3 倍。

2.12.1 门限效应

当信噪比小于 1 时, 以噪声项为主, 包络检波的性能将完全不同于以上的描述。在这种情况下, 将窄带噪声 $n(t)$ 用包络 $r(t)$ 和相位 $\psi(t)$ 进行表示更为方便, 即

$$n(t) = r(t) \cos[2\pi f_c t + \psi(t)] \quad (2.97)$$

检测器输入 $x(t) = s(t) + n(t)$ 的矢量图如图 2.38(b)所示, 此时以噪声为主, 因此将噪声包络作为参考。在噪声矢量 $r(t)$ 上叠加一个表示信号项 $A_c [1 + k_a m(t)]$ 的矢量, 两者之间的夹角等于噪声 $n(t)$ 的相位 $\psi(t)$ 。图 2.38(b)中假定信噪比非常低, 使得大部分时间内载波幅度 A_c 小于噪声包络 $r(t)$ 。忽略信号与噪声正交的分量, 则图 2.38(b)所示包络检波器的输出为

$$y(t) \approx r(t) + A_c \cos[\psi(t)] + A_c k_a m(t) \cos[\psi(t)] \quad (2.98)$$

该关系式说明当载噪比很低时, 检测器没有与消息信号 $m(t)$ 严格成正比的输出分量。 $y(t)$ 表达式的最后一项含有消息信号 $m(t)$ 与噪声的乘积项, 其中噪声以 $\cos[\psi(t)]$ 形式存在。由 1.11 节可知, 窄带噪声 $n(t)$ 的相位 $\psi(t)$ 服从 2π 弧度范围内的均匀分布。因此, 在检测器输出端并不包含消息信号 $m(t)$, 即信息量完全损失了。低载噪比的情况下包络检波器中消息的丢失称为门限效应。^[8] 门限是一个载噪比值, 当载噪比低于该值时, 检测器的噪声性能将迅速恶化。每一个非线性检测器(例如包络检波器)都存在门限效应, 这一点很重要。但在相干检测器中, 并不存在门限效应。

对于普遍情况下的 AM 信号门限效应的严格数学分析, 超出了本书的范围。在下一节中, 为了简化问题, 我们只考虑未调制的载波, 但这并不影响我们对包络检波器中门限效应的理解。

2.12.2 包络检波中 $(SNR)_o$ 的通式^①

考虑一个包络检波器的输入信号为

$$x(t) = A_c \cos(2\pi f_c t) + n(t) \quad (2.99)$$

其中, $A_c \cos(2\pi f_c t)$ 为未调制载波, $n(t)$ 为零均值、带限高斯白噪声 $N(t)$ 的样本函数。 $N(t)$ 的功率谱密度为

$$S_N(f) = \begin{cases} \frac{N_0}{2} & |f - f_c| \leq W \\ 0 & \text{其他} \end{cases} \quad (2.100)$$

将窄带噪声 $n(t)$ 用它的同相分量 $n_I(t)$ 和正交分量 $n_Q(t)$ 表示, 则检测器输入的含噪声的信号可表示为

$$x(t) = (A_c + n_I(t)) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t) \quad (2.101)$$

噪声分量 $n_I(t)$ 和 $n_Q(t)$ 均为零均值、联合高斯分布、相互独立的低通随机过程, 它们的功率谱密度相等(见式(1.101)):

$$S_{N_I}(f) = S_{N_Q}(f) = \begin{cases} S_N(f - f_c) + S_N(f + f_c) & |f| \leq W \\ 0 & \text{其他} \end{cases} \quad (2.102)$$

输入信号由平均功率等于 $A_c^2/2$ 的未调制载波组成。检测器输入的平均噪声功率为

$$\sigma_N^2 = 2WN_0 \quad (2.103)$$

因而载噪比定义为

$$\begin{aligned} \rho &= \frac{A_c^2/2}{\sigma_N^2} \\ &= \frac{A_c^2}{4WN_0} \end{aligned} \quad (2.104)$$

在这个问题中可将 ρ 视为输入信噪比值。

但输出信噪比的确定比较困难, 这是由于包络检波器的输出

$$y(t) = \sqrt{(A_c + n_I(t))^2 + n_Q^2(t)} \quad (2.105)$$

是信号和噪声项的非线性组合。在检测器输出 $y(t)$ 中, 信号与噪声没有严格分开。那么, 如何将信号 $s(t)$ 从含有噪声 $n(t)$ 的 $y(t)$ 中分离出来呢? 为了解决这个问题, 我们采用一种基于信号平均的近似方法。首先引入以下两个概念:

1. 平均输出信号 s_o , 表示信号和噪声均存在时的 $y(t)$ 的期望值与只含噪声的 $y(t)$ 的期望值之差, 即

$$s_o = E[y(t)] - E[y_o(t)] \quad (2.106)$$

式中, $y(t)$ 由式(2.105)所定义, $y_o(t)$ 的定义式为

^① 不关心噪声对 AM 信号包络检测影响数学细节的读者可直接跳至式(2.124), 并阅读该式的两种极限情况。

$$y_o(t) = \sqrt{n_I^2(t) + n_Q^2(t)} \quad (2.107)$$

2. 平均输出噪声功率, 等于检测器输出 $y(t)$ 的均方值与 $y(t)$ 均值平方之差, 即

$$\text{var}[y(t)] = E[y^2(t)] - (E[y(t)])^2 \quad (2.108)$$

在此基础上, 定义输出信噪比为

$$(\text{SNR})_o = \frac{s_o^2}{\text{var}[y(t)]} \quad (2.109)$$

由 1.12 节可知, 当只含噪声时, 包络检波器的输出服从瑞利分布, 即

$$f_{Y_o}(y) = \begin{cases} \frac{y}{\sigma_N^2} \exp\left(-\frac{y^2}{2\sigma_N^2}\right), & y \geq 0 \\ 0 & \text{其他} \end{cases} \quad (2.110)$$

因此, $y_o(t)$ 的期望值为

$$\begin{aligned} E[y_o(t)] &= \int_{-\infty}^{\infty} y f_{Y_o}(y) dy \\ &= \int_0^{\infty} \frac{y^2}{\sigma_N^2} \exp\left(-\frac{y^2}{2\sigma_N^2}\right) dy \end{aligned} \quad (2.111)$$

由自变量 x 取正值时 γ 函数的定义可得

$$\Gamma(x) = \int_0^{\infty} z^{x-1} \exp(-z) dz \quad (2.112)$$

因此, 可将式(2.111)记为

$$\begin{aligned} E[y_o(t)] &= \sqrt{2} \sigma_N \Gamma\left(\frac{3}{2}\right) \\ &= \sqrt{\frac{\pi}{2}} \sigma_N \end{aligned} \quad (2.113)$$

此处用到了 $\Gamma(3/2) = \sqrt{\pi/2}$ 。为计算检测器输出信号 s_o 的均值, 需要知道 $y(t)$ 的期望值。由于 $y(t)$ 是信号与噪声之和, 由 1.13 节可知 $y(t)$ 服从莱斯分布, 即

$$f_Y(y) = \begin{cases} \frac{y}{\sigma_N^2} \exp\left(-\frac{y^2 + A_c^2}{2\sigma_N^2}\right) I_0\left(\frac{A_c y}{\sigma_N^2}\right) & y \geq 0 \\ 0 & \text{其他} \end{cases} \quad (2.114)$$

其中 $I_0(\cdot)$ 为修正的第一类零阶贝塞尔函数(见附录 3)。因此

$$E[y(t)] = \int_0^{\infty} \frac{y^2}{\sigma_N^2} \exp\left(-\frac{y^2 + A_c^2}{2\sigma_N^2}\right) I_0\left(\frac{A_c y}{\sigma_N^2}\right) dy \quad (2.115)$$

令 $A_c y / \sigma_N^2 = u$, 结合 $\rho = A_c^2 / 2\sigma_N^2$, 可将期望值表示为

$$E[y(t)] = \frac{\sigma_N}{(2\rho)^{3/2}} \exp(-\rho) \int_0^{\infty} u^2 \exp\left(-\frac{u^2}{4\rho}\right) I_0(u) du \quad (2.116)$$

可用汇合型超几何函数(见附录 4)将式(2.116)中的积分写为简单形式。特别地, 利用积分式

$$\int_0^{\infty} u^{m-1} \exp(-b^2 u^2) I_0(u) du = \frac{\Gamma(m/2)}{2b^m} \left({}_1F_1\left(\frac{m}{2}; 1; \frac{1}{4b^2}\right) \right) \quad (2.117)$$

取 $m = 3$, $\Gamma(m/2) = \sqrt{\pi/2}$, $b^2 = 1/4\rho$, 由汇合型超几何函数 ${}_1F_1(3/2; 1; \rho)$, 可将 $y(t)$ 的期望值表示为

$$E[y(t)] = \sqrt{\frac{\pi}{2}} \sigma_N \exp(-\rho) \left({}_1F_1\left(\frac{3}{2}; 1; \rho\right) \right) \quad (2.118)$$

利用下边的恒等式可进一步简化问题:

$$\exp(-u)({}_1F_1(\alpha; \beta; u)) = {}_1F_1(\beta - \alpha; \beta; -u) \quad (2.119)$$

最终将 $y(t)$ 的期望值简单地表示为

$$E[y(t)] = \sqrt{\frac{\pi}{2}} \sigma_N \left({}_1F_1\left(-\frac{1}{2}; 1; -\rho\right) \right) \quad (2.120)$$

在式(2.106)中, 应用式(2.113)和式(2.120), 求得平均输出信号为

$$s_o = \sqrt{\frac{\pi}{2}} \sigma_N \left({}_1F_1\left(-\frac{1}{2}; 1; -\rho\right) - 1 \right) \quad (2.121)$$

输出信号与噪声 $n(t)$ 的标准偏差 σ_N 有关, 这说明检测器输出端得到的是信号与噪声的混合。

按照类似的推导过程, 可将检测器输出信号 $y(t)$ 的均方值表示为

$$\begin{aligned} E[y^2(t)] &= \int_0^{\infty} \frac{y^3}{\sigma_N^2} \exp\left(-\frac{y^2 + A_c^2}{2\sigma_N^2}\right) I_0\left(\frac{A_c y}{\sigma_N^2}\right) dy \\ &= 2\sigma_N^2 ({}_1F_1(-1; 1; -\rho)) \end{aligned} \quad (2.122)$$

在式(2.108)中, 应用式(2.120)和式(2.122), 可求得平均输出噪声功率为

$$\text{var}[y(t)] = 2\sigma_N^2 \left({}_1F_1(-1; 1; -\rho) - \frac{\pi}{4} \left({}_1F_1\left(-\frac{1}{2}; 1; -\rho\right) \right)^2 \right) \quad (2.123)$$

最后, 在式(2.109)中应用式(2.121)和式(2.123), 可得包络检波的输出信噪比为

$$(\text{SNR})_o = \frac{\left({}_1F_1\left(-\frac{1}{2}; 1; -\rho\right) - 1 \right)^2}{(4/\pi) \left({}_1F_1(-1; 1; -\rho) \right) - \left({}_1F_1\left(-\frac{1}{2}; 1; -\rho\right) \right)^2} \quad (2.124)$$

当包络检波器的输入由未调制载波和带限高斯白噪声组成时, 式(2.124)为检测器输出信噪比的通式。该通式的两种极限情况有特殊的意义:

1. 大载噪比。当 ρ 很大时, 可用下面的渐近公式(见附录 4)

$${}_1F_1\left(-\frac{1}{2}; 1; -\rho\right) \approx 2\sqrt{\frac{\rho}{\pi}} \quad \text{当 } \rho \rightarrow \infty \quad (2.125)$$

且有下面的恒等式

$${}_1F_1(-1; 1; -\rho) = 1 + \rho \quad (2.126)$$

上式对于所有 ρ 值都是成立的。将式(2.125)和式(2.126)代入式(2.124)中, 可求得输出信噪比的近似公式如下:

$$(\text{SNR})_o \approx \rho \quad \text{当 } \rho \rightarrow \infty \quad (2.127)$$

由于 ρ 很大时, 与 ρ 相比, $\rho^{1/2}$ 和 ρ^0 对结果的影响很小, 因而将式(2.124)分子中的对应的两项忽略。式(2.127)说明: 在大载噪比 ρ 的情况下, 包络检波器的作用类似于相干检测器, 即输出信噪比与输入信噪比成正比。

2. 小载噪比。当 ρ 很小时有(见附录 4)

$${}_1F_1(a; c; -\rho) \approx 1 - \frac{a}{c} \rho \quad \text{当 } \rho \rightarrow 0 \quad (2.128)$$

利用该渐近式可得, ρ 很小时, 输出信噪比的近似表达式为

$$\begin{aligned} (\text{SNR})_o &\approx \frac{\pi \rho^2}{16 - 4\pi} \\ &\approx 0.91\rho^2 \quad \text{当 } \rho \rightarrow 0 \end{aligned} \quad (2.129)$$

当 ρ 很小时, 与 ρ^0 相比 ρ 和 ρ^2 对结果的影响很小, 因此将该式分母中的 ρ 和 ρ^2 项忽略。式(2.129)说明, 在小载噪比的情况下, 包络检波器的输出信噪比与输入信噪比的平方成正比。

由这两种极限情况得到的结论是, 包络检波器支持强信号而抑制弱信号。这种弱信号被抑制的现象称为弱信号抑制, 是门限效应的表现。

利用式(2.124), 可画出包络检波器输出信噪比 $(\text{SNR})_o$ 对载噪比 ρ 的曲线, 如图 2.39 所示, 这里用到了汇合型超几何函数已有的函数值。该图包括了 ρ 很大和 ρ 很小这两种极限情况。从图中看出, 大约在载噪比为 10 dB(即 $\rho \approx 10$)时, 输出信噪比就开始偏离线性区域了。

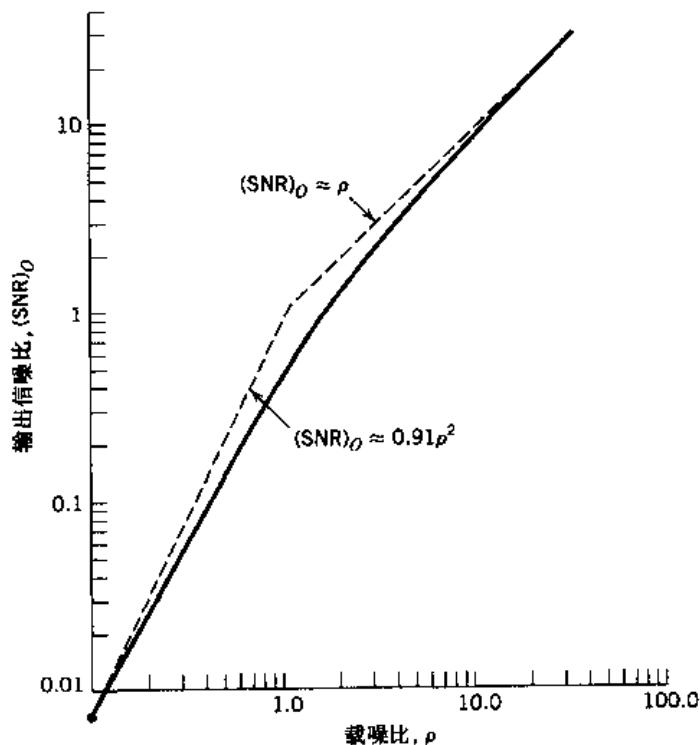


图 2.39 包络检波器的输出信噪比随载噪比的变化曲线

2.13 FM 接收机中的噪声

最后,我们转向 FM 系统中的噪声分析,为此,采用图 2.40 中的接收机模型。如前所述,噪声 $w(t)$ 是零均值且功率密度谱为 $N_0/2$ 的高斯白噪声。接收到的 FM 信号为 $s(t)$,其载频为 f_c ,传输带宽为 B_T ,在正频率部分,频带 $f_c \pm B_T/2$ 以外的功率可以忽略,负频率部分也一样。

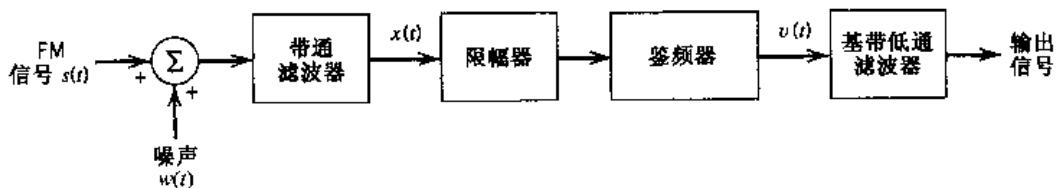


图 2.40 FM 接收机模型

与在 AM 系统中一样,带通滤波器中心频率为 f_c ,带宽为 B_T ,因此可允许 FM 信号无失真地通过。普遍来说, B_T 小于中心频率 f_c ,因此信道噪声 $w(t)$ 经滤波得到的噪声 $n(t)$ 可采用窄带表达式,即用噪声的正交分量和同相分量来表示。

在 FM 系统里,通过正弦载波瞬时频率的变化来传递消息信号,载波的幅度保持不变。因此,接收机端输入载波幅度的波动必定缘于噪声或干扰。图 2.40 接收机模型中带通滤波器之后的限幅器是为了去掉幅度的波动,即通过对滤波器输出的已调信号进行削波处理,使其接近于零轴。产生的方波由另一个限幅器中的带通滤波器进行整流,从而抑制了载频的谐波分量。此时滤波器的输出重新变为正弦信号,其幅度与接收机输入的载波幅度无关。

图 2.40 所示模型中的鉴频器由以下两部分组成:

1. 频率响应为纯虚数,且随频率呈线性变化的微分网络或微分器,其作用是产生幅度和频率均随消息信号变化的复合已调信号。
2. 包络检波器,作用是恢复出幅度的变化量从而重新生成消息信号。

微分网络与包络检波器常被视为一个整体物理单元。

图 2.40 中的基带低通滤波器为一后检波滤波器,其带宽刚好允许消息信号的最高频率分量通过。该滤波器滤掉了鉴频器输出端的带外噪声分量,从而将输出噪声的影响降到了最低。

图 2.40 中,带通滤波器输出的滤波后噪声 $n(t)$ 可用同相分量和正交分量表示为

$$n(t) = n_I(t) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t)$$

也可将 $n(t)$ 用包络和相位表示为

$$n(t) = r(t) \cos[(2\pi f_c t) + \psi(t)] \quad (2.130)$$

其中,包络为

$$r(t) = [n_I^2(t) + n_Q^2(t)]^{1/2} \quad (2.131)$$

而相位为

$$\psi(t) = \tan^{-1} \left[\frac{n_O(t)}{n_I(t)} \right] \quad (2.132)$$

其中, 包络 $r(t)$ 服从瑞利分布, 相位 $\psi(t)$ 服从 2π 弧度内的均匀分布(见 1.12 节)。

输入 FM 信号 $s(t)$ 由下式定义:

$$s(t) = A_c \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right] \quad (2.133)$$

其中, A_c 为载波幅度, f_c 为载频, k_f 为频率灵敏度, $m(t)$ 为消息信号。需注意, 与标准 AM 系统一样, FM 系统里不需要对已调信号 $s(t)$ 引入比例因子, 因为可以假定 $s(t)$ 的幅度 A_c 与加性噪声分量 $n(t)$ 同量纲。接下来定义

$$\phi(t) = 2\pi k_f \int_0^t m(\tau) d\tau \quad (2.134)$$

则可将 $s(t)$ 简单表示为

$$s(t) = A_c \cos[2\pi f_c t + \phi(t)] \quad (2.135)$$

因此, 带通滤波器输出的含噪声信号为

$$\begin{aligned} x(t) &= s(t) + n(t) \\ &= A_c \cos[2\pi f_c t + \phi(t)] + r(t) \cos[2\pi f_c t + \psi(t)] \end{aligned} \quad (2.136)$$

将 $x(t)$ 用图 2.41 所示的矢量图表示是很有意义的。图中以信号相位为参照。由图 2.41 直接得到代表 $x(t)$ 的和矢量相位的 $\theta(t)$ 为

$$\theta(t) = \phi(t) + \tan^{-1} \left\{ \frac{r(t) \sin[\psi(t) - \phi(t)]}{A_c + r(t) \cos[\psi(t) - \phi(t)]} \right\} \quad (2.137)$$

$x(t)$ 的包络没有研究意义, 因为带通滤波器输出端的包络波动可通过限幅器去掉。

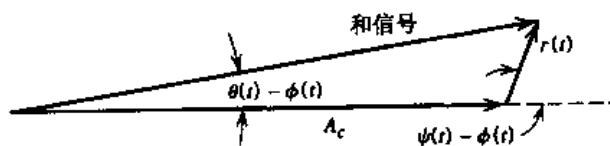


图 2.41 在大载噪比下, FM 信号和窄带噪声的矢量图

我们研究的目的是确定由于滤波后噪声 $n(t)$ 的存在所造成的载波瞬时频率误差。假定鉴频器是理想的, 其输出与 $\theta'(t)/(2\pi)$ 成正比, $\theta'(t)$ 是 $\theta(t)$ 对时间的导数。由于 $\theta(t)$ 的定义式太复杂, 为了得到有用的结论, 需要进行一些简化近似。

假定鉴频器输入端测得的载噪比远大于 1。通过观察(在固定时间内)样本函数 $r(t)$ (由噪声 $n(t)$ 引起)的包络变化, 得到一随机变量 R 。则在大部分时间内, 随机变量 R 小于载波幅度 A_c , 因此, 相位 $\theta(t)$ 的表达式可简化为如下形式:

$$\theta(t) \approx \phi(t) + \frac{r(t)}{A_c} \sin[\psi(t) - \phi(t)] \quad (2.138)$$

或将式(2.134)中的 $\phi(t)$ 代入, 得

$$\theta(t) \approx 2\pi k_f \int_0^t m(\tau) d\tau + \frac{r(t)}{A_c} \sin[\psi(t) - \phi(t)] \quad (2.139)$$

因此, 鉴频器输出为

$$\begin{aligned} v(t) &= \frac{1}{2\pi} \frac{d\theta(t)}{dt} \\ &= k_f m(t) + n_d(t) \end{aligned} \quad (2.140)$$

其中, 噪声项 $n_d(t)$ 的定义为

$$n_d(t) = \frac{1}{2\pi A_c} \frac{d}{dt} \{r(t) \sin[\psi(t) - \phi(t)]\} \quad (2.141)$$

由式(2.140)可见, 若载噪比很高, 则鉴频器输出 $v(t)$ 为原始消息信号 $m(t)$ 乘以常量因子 k_f 后, 与加性噪声分量 $n_d(t)$ 之和。相应地, 可用前面定义过的输出信噪比来衡量 FM 接收机的性能。为研究方便, 我们有必要将噪声 $n_d(t)$ 的定义式进行适当简化。

由图 2.41, 我们注意到窄带噪声相位 $\psi(t)$ 的变化给信号相位 $\phi(t)$ 带来了一定的影响。由于相位 $\psi(t)$ 服从 2π 弧度内的均匀分布, 因此, 我们考虑假设相位差 $\psi(t) - \phi(t)$ 也服从 2π 弧度内的均匀分布。如果假设成立的话, 那么鉴频器输出噪声 $n_d(t)$ 将与调制信号无关, 而只与载波和窄带噪声的性质有关。理论推导证明, 当载噪比很高^[9] 时, 该假设是成立的。从而可将式(2.141)简化为

$$n_d(t) \approx \frac{1}{2\pi A_c} \frac{d}{dt} \{r(t) \sin[\psi(t)]\} \quad (2.142)$$

由 $r(t)$ 和 $\psi(t)$ 的定义式, 结合滤波后噪声 $n(t)$ 的正交分量 $n_Q(t)$:

$$n_Q(t) = r(t) \sin[\psi(t)] \quad (2.143)$$

可将式(2.142)写为

$$n_d(t) \approx \frac{1}{2\pi A_c} \frac{dn_Q(t)}{dt} \quad (2.144)$$

该式说明, 鉴频器输出端的加性噪声 $n_d(t)$ 由载波幅度 A_c 和窄带噪声 $n(t)$ 的正交分量 $n_Q(t)$ 决定。

输出信噪比就是平均输出信号功率与平均输出噪声功率之比。由式(2.140)可见, 鉴频器输出的信号分量, 亦即低通滤波器的输出为 $k_f m(t)$ 。因而, 平均输出信号功率等于 $k_f^2 P$ 。其中, P 为消息信号 $m(t)$ 的平均功率。

鉴频器输出噪声 $n_d(t)$ 与噪声正交分量 $n_Q(t)$ 对时间的导数成正比。由于时域中的函数微分对应在频域中为其傅里叶变换乘以 $j2\pi f$, 故可使 $n_Q(t)$ 通过频率响应 $n_d(t)$ 为

$$\frac{j2\pi f}{2\pi A_c} = \frac{jf}{A_c}$$

的线性滤波器来得到噪声过程 $n_d(t)$ 。这说明噪声 $n_d(t)$ 的功率密度谱 $S_{n_d}(f)$ 与正交噪声分量 $n_Q(t)$ 的功率密度谱 $S_{n_Q}(f)$ 有如下关系:

$$S_{N_d}(f) = \frac{f^2}{A_c^2} S_{N_0}(f) \quad (2.145)$$

图 2.40 所示的接收机模型中的带通滤波器的频率响应是理想的, 其带宽为 B_T , 中心频率为 f_c 。窄带噪声 $n(t)$ 通过这样的滤波器后, 噪声的功率谱密度特性将与滤波器的频率响应相似。这意味着窄带噪声 $n(t)$ 的正交分量 $n_0(t)$ 将具有理想低通特性, 如图 2.42(a)。相应的 $n_d(t)$ 功率谱密度如图 2.42(b) 所示, 即

$$S_{N_d}(f) = \begin{cases} \frac{N_0 f^2}{A_c^2}, & |f| \leq \frac{B_T}{2} \\ 0, & \text{其他} \end{cases} \quad (2.146)$$

在图 2.40 所示的接收机模型中, 鉴频器输出后为低通滤波器, 低通滤波器的带宽等于消息带宽 W 。对于宽带 FM 系统, W 通常要小于 $B_T/2$, B_T 为 FM 信号的传输带宽。这意味着噪声 $n_d(t)$ 的带外分量将被舍弃。因此, 接收机输出噪声 $n_o(t)$ 的功率谱密度 $S_{N_o}(f)$ 为

$$S_{N_o}(f) = \begin{cases} \frac{N_0 f^2}{A_c^2}, & |f| \leq W \\ 0, & \text{其他} \end{cases} \quad (2.147)$$

如图 2.42(c) 所示。平均输出噪声功率等于功率密度谱 $S_{N_o}(f)$ 从 $-W$ 到 W 的积分值。因此有如下结果:

$$\begin{aligned} \text{输出噪声平均功率} &= \frac{N_0}{A_c^2} \int_{-W}^W f^2 df \\ &= \frac{2N_0 W^3}{3A_c^2} \end{aligned} \quad (2.148)$$

注意, 平均输出噪声功率与平均载波功率 $A_c^2/2$ 成反比。因此, 在 FM 系统里, 增加载波功率可起到静噪的作用。

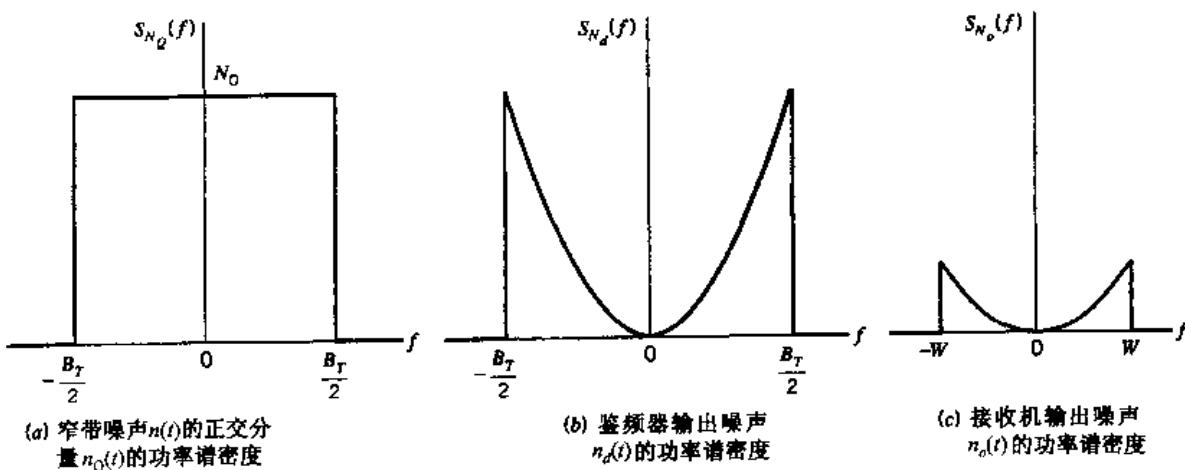


图 2.42 FM 接收机的噪声分析

前面求出了平均输出信号功率为 $k_f^2 P$ 。若载噪比很高,用平均输出信号功率除以式(2.148)所示的平均输出噪声功率,可得输出信噪比为

$$(\text{SNR})_{O,\text{FM}} = \frac{3A_c^2 k_f^2 P}{2N_0 W^3} \quad (2.149)$$

已调信号 $s(t)$ 的平均功率为 $A_c^2/2$,消息带宽内的平均噪声功率等于 WN_0 ,因此信道信噪比为

$$(\text{SNR})_{C,\text{FM}} = \frac{A_c^2}{2WN_0} \quad (2.150)$$

用输出信噪比除以信道信噪比,得到频率调制的解调增益为:

$$\left. \frac{(\text{SNR})_O}{(\text{SNR})_C} \right|_{\text{FM}} = \frac{3k_f^2 P}{W^2} \quad (2.151)$$

由2.7节可知,频偏 Δf 与调制器频率灵敏度 k_f 成正比。由定义,偏移率 D 等于频偏 Δf 除以消息带宽 W 。换言之,偏移率 D 与 $k_f P^{1/2}/W$ 成正比。因此,由式(2.151)可知,宽带 FM 系统的解调增益是偏移率的二次函数。此时,在宽带 FM 系统中,传输带宽 B_T 与偏移率 D 近似成正比。由此可知,当载噪比很高时,输出信噪比或 FM 系统的解调增益将随传输带宽 B_T 的增加以平方律形式增大。由此可得一个重要结论,即在 FM 系统中,可通过增加传输带宽来换取噪声性能的改善,这一点是不同于 AM 系统的。

例 2.5 单音频调制

考虑将频率为 f_m 的正弦信号作为调制信号,假设最大频偏值为 Δf 。则已调 FM 信号定义如下:

$$s(t) = A_c \cos \left[2\pi f_c t + \frac{\Delta f}{f_m} \sin(2\pi f_m t) \right]$$

则有

$$2\pi k_f \int_0^t m(\tau) d\tau = \frac{\Delta f}{f_m} \sin(2\pi f_m t)$$

在上式两边对时间求导,并解出 $m(t)$,可得

$$m(t) = \frac{\Delta f}{k_f} \cos(2\pi f_m t)$$

消息信号 $m(t)$ 通过 1Ω 的电阻后,其平均功率为

$$P = \frac{(\Delta f)^2}{2k_f^2}$$

将结果代入输出信噪比的计算式(2.149)中,有

$$\begin{aligned} (\text{SNR})_{O,\text{FM}} &= \frac{3A_c^2 (\Delta f)^2}{4N_0 W^3} \\ &= \frac{3A_c^2 \beta^2}{4N_0 W} \end{aligned}$$

其中, $\beta = \Delta f / W$ 是调制指数。用式(2.151)计算相应的解调增益, 有

$$\left. \frac{(\text{SNR})_O}{(\text{SNR})_C} \right|_{\text{FM}} = \frac{3}{2} \left(\frac{\Delta f}{W} \right)^2 = \frac{3}{2} \beta^2 \quad (2.152)$$

需注意调制指数 $\beta = \Delta f / W$ 是由检波后的低通滤波器带宽 W 决定的, 与正弦消息频率 f_m 无关。同时, 滤波器的选择应使有用消息的频谱能够通过。对于带宽为 W 的系统, 正弦消息频率 f_m 可位于 0 到 W 区间内的任何位置, 对应的输出信噪比不变。

对 AM 和 FM 系统的噪声性能进行比较有着重要的意义。利用基于正弦调制信号的两个系统的解调增益来进行比较, 是一种有效的方法。在正弦调制信号下, 调制百分比为 100% 的 AM 系统的解调增益为(由例 2.4):

$$\left. \frac{(\text{SNR})_O}{(\text{SNR})_C} \right|_{\text{AM}} = \frac{1}{3}$$

将这个结果与式(2.152)所示的 FM 系统的相应解调增益进行比较可发现, 当满足下述条件时, 采用频率调制能提供比幅度调制更好的抗噪声性能:

$$\frac{3}{2} \beta^2 > \frac{1}{3}$$

即

$$\beta > \frac{\sqrt{2}}{3} = 0.471$$

可粗略地将 $\beta = 0.5$ 作为窄带 FM 和宽带 FM 的分界线。这个观点将噪声考虑在内, 进一步证实了 2.7 节中研究 FM 信号带宽时得出的类似结论。

2.13.1 捕获效应

FM 系统能将无用信号(例如刚刚讨论的噪声)的影响降至最低。即使其他调频信号的频率分量接近于有用 FM 信号的载频, 从而对有用信号产生干扰, FM 系统也能将这种干扰降至最低。但只有当干扰弱于有用的 FM 输入信号时, FM 接收机才能有效地抑制干扰。如果干扰信号较强, 则接收机将处理强信号, 而抑制有用的 FM 输入信号。当这两种信号强度接近时, 接收机将在两个信号间来回切换工作, 这种现象称为捕获效应。捕获效应是 FM 系统中又一个独特的性质。

2.13.2 FM 门限效应

式(2.149)定义了 FM 接收机的输出信噪比。只有当鉴频器输入端测得的载噪比远大于 1 时, 该式才有效。实验表明, 随着输入噪声功率的增大, 载噪比减小, FM 接收机将失效。最初的现象是在接收机的输出端听到咔哒声。随着载噪比的进一步减小, 咔哒声迅速变为爆裂声或劈啪声。在接近失效点时, 预测的输出信噪比将大于实际值, 式(2.149)开始失效。这种现象称为门限效应^[10]。在低噪声功率下, 接收机性能不会发生恶化, 仍可由通常的公式计算信噪比, 这种情况下的最小载噪比值被称为门限值。

为了对 FM 门限效应进行定性的讨论, 首先考虑没有信号, 只有未调制载波时的情况。此

时, 鉴频器输入端的组合信号为

$$x(t) = [A_c + n_I(t)] \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t) \quad (2.153)$$

其中, $n_I(t)$ 和 $n_Q(t)$ 分别为窄带噪声关于载波的同相分量和正交分量。图 2.43 的矢量图给出了式(2.153)里 $x(t)$ 不同分量间的相位关系。随着 $n_I(t)$ 和 $n_Q(t)$ 的幅值及相位随时间随机变化, P_1 点(代表 $x(t)$ 的矢量箭头处)绕 P_2 点(代表载波的矢量箭头处)转动。当载噪比很大时, $n_I(t)$ 和 $n_Q(t)$ 的幅度通常远小于载波幅度 A_c , 因此在绝大部分时间里, 图 2.43 中动点 P_1 与 P_2 点非常接近。因此, 角度 $\theta(t)$ 近似地位于 $n_Q(t)/A_c$ 到 2π 的整数倍之间。另一方面, 当载噪比很小时, 动点 P_1 偶尔扫过原点, $\theta(t)$ 则相应地增大或减小 2π 弧度。图 2.44 粗略绘制了当 $\theta(t)$ 如图 2.44(a)发生变化时, 由 $\theta'(t) = d\theta/dt$ 得到的脉冲形分量。鉴频器的输出 $v(t)$ 等于 $\theta'(t)/(2\pi)$ 。脉冲形分量的高度由动点 P_1 与原点 O 之间的距离决定, 但所有脉冲形的面积均接近 $\pm 2\pi$ 弧度, 如图 2.44(b) 所示。当图 2.44(b) 中的信号检波后通过低通滤波器时, 相应地, 在接收机输出端将产生更宽的脉冲形分量, 此时, 就产生了咔嗒声。只有当 $\theta(t)$ 变化了 $\pm 2\pi$ 弧度时, 才会产生咔嗒声。

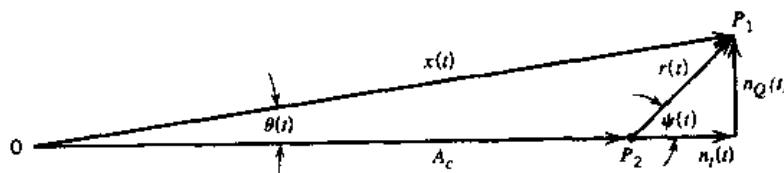


图 2.43 式(2.153)的相量图解释

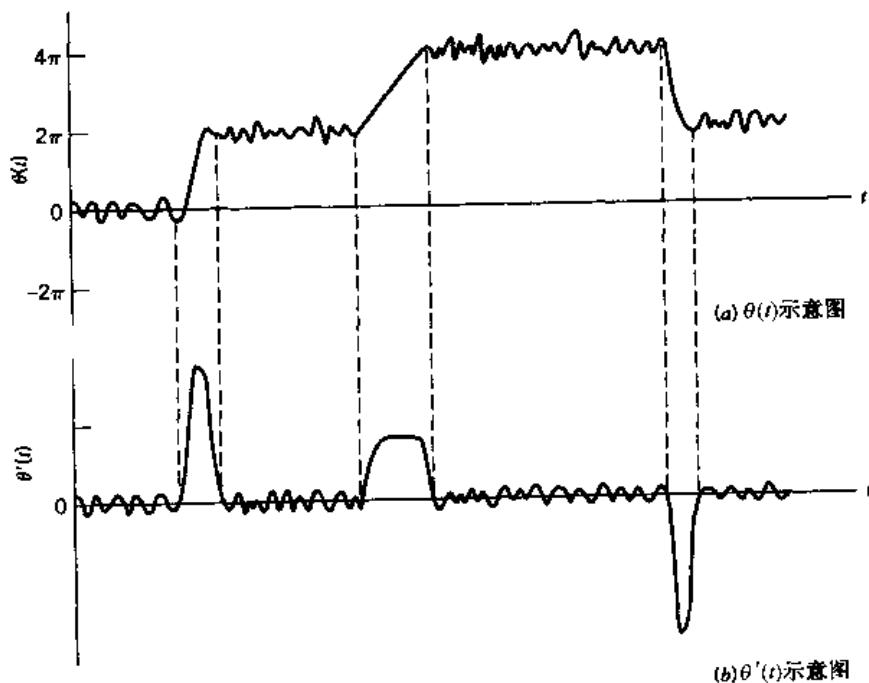


图 2.44 $\theta(t)$ 变化了 2π , 使 $\theta'(t) = d\theta(t)/dt$, 从而产生脉冲形分量的示意图

由图 2.43 所示的矢量图, 可推导出产生咔嗒声的条件。当窄带噪声 $n(t)$ 的包络 $r(t)$ 和相位 $\psi(t)$ 满足以下条件时, 出现正向咔嗒声:

$$\begin{aligned} r(t) &> A_c \\ \psi(t) &< \pi \leq \psi(t) + d\psi(t) \\ \frac{d\psi(t)}{dt} &> 0 \end{aligned}$$

这些条件保证了在时间增量 dt 内组合矢量 $x(t)$ 的相位 $\theta(t)$ 改变了 2π 弧度, 而窄带噪声的相位增加了为 $d\psi(t)$ 。同理, 产生负向咔嗒声的条件如下:

$$\begin{aligned} r(t) &> A_c \\ \psi(t) &> -\pi > \psi(t) + d\psi(t) \\ \frac{d\psi(t)}{dt} &< 0 \end{aligned}$$

这些条件保证了在时间增量 dt 内, $\theta(t)$ 改变了 -2π 。

载噪比的定义为

$$\rho = \frac{A_c^2}{2B_T N_0} \quad (2.154)$$

随着 ρ 的不断减小, 单位时间内咔嗒声的平均次数逐渐增大。当次数很大时, 就会发生门限效应。

输出信噪比的计算过程如下所述:

1. 在接收机输出端测出无噪时的输出信号。假定为正弦信号调制, 频偏 Δf 等于 $B_T/2$, 载波在整个输入频带内来回变化, 计算该情况下的平均输出信号功率。
2. 计算无信号时的平均输出噪声功率。即载波未经过调制, 对载噪比 ρ 值没有限制。

在此基础上, 图 2.45 给出了比值 $B_T/2W$ 等于 5 时, 根据理论^[1]得出的输出信噪比随载噪比的变化曲线, 即曲线 I。该曲线表明, 当载噪比 ρ 小于 10 dB 时, 输出信噪比将明显偏离 ρ 的线性函数。图 2.45 中曲线 II 说明了当调制信号(假定是正弦信号)和噪声同时存在时, 调制对输出信噪比的影响。曲线 II 中平均输出信号功率的取值与曲线 I 中相同。但由于平均输出噪声功率主要由调制信号决定, 造成了曲线 II 与曲线 I 的明显偏离。尤其是当 ρ 趋于负无穷时, 输出信噪比更是严重偏离了 $\rho \approx 11$ dB 时的线性函数。当存在信号时, 载波调制将使每秒的咔嗒声次数增加。实验表明, 当载噪比约为 13 dB 时, 在接收机的输出端就可断续听到咔嗒声, 这个载噪比的值只是稍大于理论值。值得注意的是, 每秒平均咔嗒声次数的增加易造成输出信噪比的急剧下降, 甚至低于调制门限值。

由前面的讨论可总结出, 在大多数实际情况中, 只要载噪比 ρ 等于或大于 20, 即 13 dB, FM 接收机的门限效应是能够避免的。由式(2.154)可知, 如果满足下式, 鉴频器输出端的消息损失是可以忽略的:

$$\frac{A_c^2}{2B_T N_0} \geq 20$$

或者说, 如果平均发射功率 $A_c^2/2$ 满足条件

$$\frac{A_c^2}{2} \geq 20B_T N_0 \quad (2.155)$$

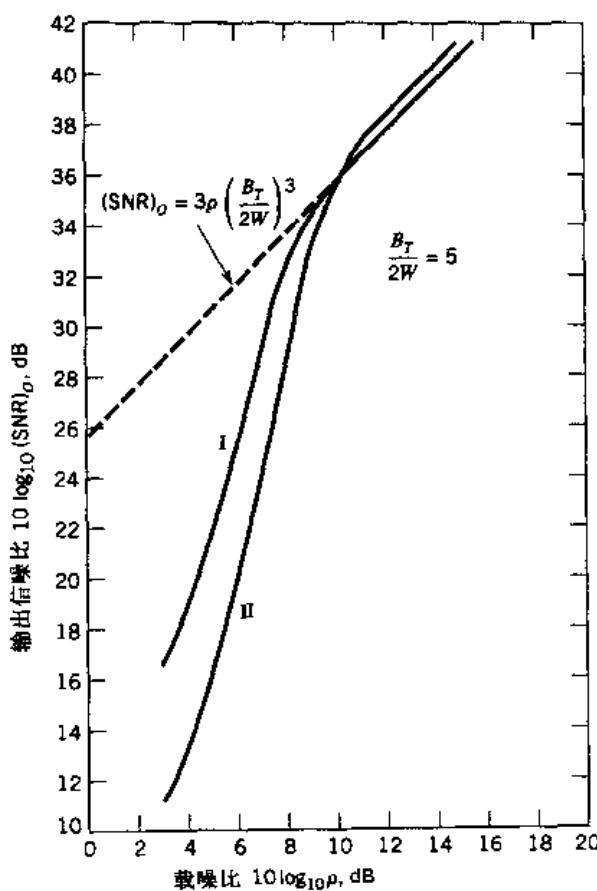


图 2.45 FM 接收机中输出信噪比与载噪比的依赖关系。曲线 I 为载波未经调制的情况下算出的输出噪声平均功率。曲线 II 是在正弦调制载波下算出的输出噪声平均功率。曲线 I 和 II 均为理论值

在应用式(2.155)时,应按下列步骤进行:

1. 对于指定的调制指数 β 和消息带宽 W ,用图 2.26 的标准曲线或卡逊准则来确定 FM 信号的传输带宽 B_T 。
2. 对于每单位带宽的指定平均噪声功率 N_0 ,用式(2.155)来计算工作点高于门限的平均发射功率 $A_e^2/2$ 的最小值。

2.13.3 FM 门限的降低

在采用频率调制的通信系统里,为使接收机在尽可能低的信号功率下正常工作,降低 FM 接收机的门限值有着特殊的意义。FM 接收机门限的降低可通过在 FM 解调器里加负反馈^[12](常称为 FMFB 解调器)实现,或是采用锁相环解调器。这些器件称为扩展门限解调器,其原理见图 2.46。图中的扩展门限值是以标准鉴频器(即无负反馈的鉴频器)为基础测量所得。

FMFB 解调器^[13]的框图如图 2.47 所示。图中,传统的 FM 接收机里的本地振荡器被压控振荡器(VCO)所代替,VCO 输出的瞬时频率由解调后的信号控制。为了理解这种接收机的工作原理,假设某一刻,将 VCO 从电路中移除,则反馈环变为开环。假定接收机的输入为

宽带 FM 信号，在混频器的 VCO 终端输入同一信源产生的第二个 FM 信号，此 FM 信号的调制指数小于接收机输入信号的调制指数。则混频器的输出信号由差频分量组成，因为和频分量被带通滤波器滤掉了。尽管两个输入 FM 信号的频偏均很大，但混频器输出信号的频偏值很小，因为两个信号之间的瞬时频偏差很小。因此，将两个调制指数相减，则混频器输出 FM 信号的调制指数变得更小。将调制指数变小的 FM 信号通过带通滤波器，此滤波器的带宽只是任一宽带 FM 信号所需带宽的一部分，然后对滤波后的 FM 信号进行频率解调。此时可明显看出，要想得到混频器输入端的第二个宽带 FM 信号，需将鉴频器的输出信号反馈回 VCO 中。

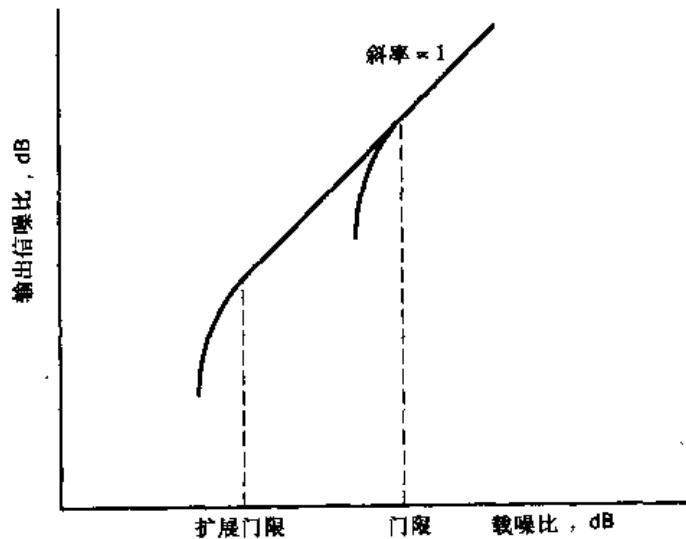


图 2.46 FM 门限扩展

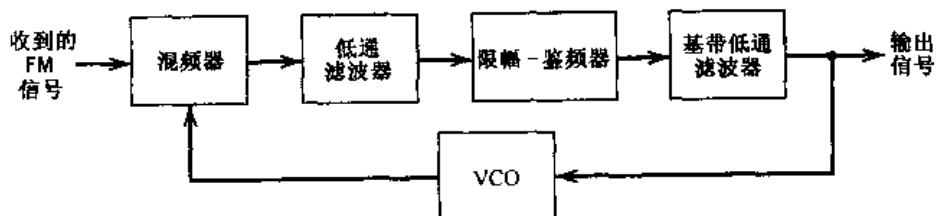


图 2.47 带有负反馈的 FM 解调器

现在可以说，若载噪比足够大，则 FMFB 接收机的信噪比与具有相同输入信号和噪声功率的传统 FM 接收机的信噪比应相等。假定在某一时刻解调器中没有反馈，则未调载波 $A_c \cos(2\pi f_c t)$ 与窄带噪声之和为

$$n(t) = n_I(t) \cos(2\pi f_c t) + n_Q(t) \sin(2\pi f_c t)$$

假设载噪比很高，则限幅-鉴频器输入组合信号 $x(t)$ 的相位近似等于 $n_Q(t)/A_c$ 。 $x(t)$ 的包络没有研究意义，因为限幅器去掉了包络的波动量。因此，鉴频器的输入组合信号由调制指数很小的 PM 组成，该 PM 信号是由与载波相正交的噪声分量 $n_Q(t)$ 调制而得。加上反馈后，VCO 将产生 FM 信号，降低带通滤波器输出信号的调相指数，即降低噪声正交分量 $n_Q(t)$ 的调相指数。由此可见，只要载噪比足够大，FMFB 接收机就不会受噪声同相分量 $n_I(t)$ 的影响。FMFB

接收机将对噪声的正交分量 $n_q(t)$ 进行解调, 解调方式与对信号进行解调的方式完全相同。由于反馈的引入, 信号与正交噪声以相同的比例减小, 其中基带信噪比与反馈无关。在大载噪比的情况下, FMFB 接收机的基带信噪比与传统 FM 接收机相同。

FMFB 接收机能够扩展门限的原因是, 与传统 FM 接收机不同, FMFB 利用了重要的先验信息, 即尽管输入 FM 信号的载频常有很大的频偏, 但其变化率仍等于基带速率。FMFB 解调器实际上是一个跟踪滤波器, 它只能跟踪宽带 FM 信号频率的缓慢变化, 因此它只对中心频率约等于瞬时载频的窄带噪声有响应。FMFB 接收机响应的噪声带宽等于 VCO 跟踪的噪声带宽。最终的结果是 FMFB 接收机能够实现 5~7 dB 范围内的门限扩展, 这表示小功率 FM 系统的设计有了重要的改进。

与 FMFB 解调器相类似, 锁相环(将在后面的 2.14 节中进行讨论)也是跟踪滤波器, 且响应的噪声带宽也等于 VCO 跟踪的噪声带宽。锁相环解调器用相对简单的电路实现了门限扩展。遗憾的是, 现有的任何一种理论都无法预知门限扩展的量值, 该值取决于信号参数。粗略地说, 在典型的应用中, 锁相环可对门限扩展几个(2~3) dB, 但效果不如 FMFB 解调器。

2.13.4 FM 系统中的预加重和去加重

式(2.147)表明, FM 接收机输出噪声的功率谱密度是工作频率的二次函数, 如图 2.48(a)所示。图 2.48(b)所示为典型消息源的功率谱密度。音频和视频信号的频谱均为这种形式。特别地, 这种消息的功率谱密度通常在高频处迅速下降。而另一方面, 输出噪声的功率谱密度却随频率增加而迅速增大。因此, 在 $f = \pm W$ 附近, 消息的频谱密度很低, 相比之下输出噪声的频谱密度却很高。很明显, 消息并没有有效利用所分配的频带。稍微减小检波后的低通滤波器带宽, 就可抑制大部分噪声功率, 且仅失去小部分消息功率, 这似乎是提高系统噪声性能的一种方法。然而, 这种方法常常不尽如人意, 因为由滤波器带宽减小造成的消息失真, 即使很微小, 也可能无法容忍的。例如, 在音频信号中, 虽然高频分量只占全部功率的一小部分, 但一旦发生失真, 就会影响欣赏效果。

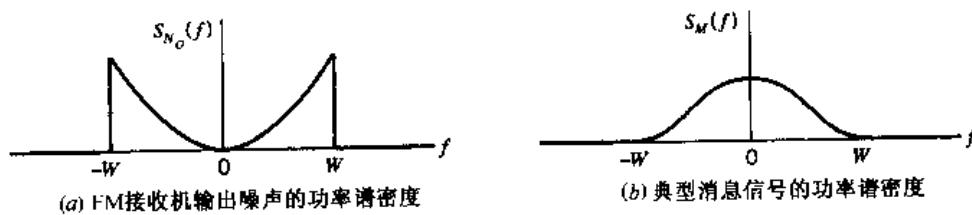


图 2.48 FM 接收机的输出噪声和消息信号的功率谱密度

有效利用所分配频带的一种令人满意的方法是在发射端采用预加重, 在接收端采用去加重, 如图 2.49 所示。

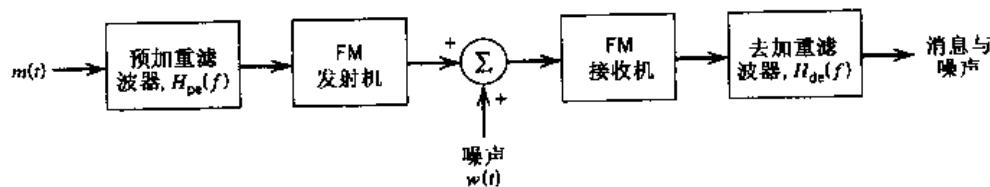


图 2.49 采用预加重和去加重的 FM 系统

在这种方法中,在发射端调制之前,也就是在噪声被引入接收机之前,人为地加重消息信号的高频分量。这使得消息功率谱密度的低频和高频部分能够充分占用所分配频带。那么,在接收机鉴频器的输出端,通过相反的操作进行高频分量的去加重,就可以恢复消息原始的信号功率分配。在这样的过程中,鉴频器输出噪声的高频分量也减小了,因而有效地增大了系统的输出信噪比。这种预加重和去加重过程在商用 FM 无线发射机和接收机中被广泛采用。

为了在接收机的输出端无失真地恢复出原始消息,在理想情况下,发射机中预加重和接收机中去加重的滤波器必须具有完全相反的频率响应。即,如果用 $H_{pe}(f)$ 表示预加重滤波器的频率响应,那么理想的去加重滤波器的频率响应 $H_{de}(f)$ 应当为

$$H_{de}(f) = \frac{1}{H_{pe}(f)}, \quad -W \leq f \leq W \quad (2.156)$$

这种选择频率响应的方法保证了接收机输出信号的平均功率与预加重和去加重过程无关。

由前面对 FM 系统的噪声分析可得,当载噪比很高时,鉴频器输出噪声 $n_d(t)$ 的功率谱密度如式(2.146)所示。因此,去加重滤波器输出噪声的改进功率谱密度为

$$|H_{de}(f)|^2 S_{N_d}(f) = \begin{cases} \frac{N_0 f^2}{A_c^2} |H_{de}(f)|^2, & |f| \leq \frac{B_T}{2} \\ 0, & \text{其他} \end{cases} \quad (2.157)$$

如前所述,一般情况下,检波后低通滤波器的带宽 W 小于 $B_T/2$,则接收机输出的改进噪声平均功率如下:

$$\left(\begin{array}{l} \text{采用去加重的平} \\ \text{均输出噪声功率} \end{array} \right) = \frac{N_0}{A_c^2} \int_{-W}^W f^2 |H_{de}(f)|^2 df \quad (2.158)$$

在理想情况下,因为接收机输出的平均消息功率不受预加重和去加重过程的影响,故由于采用了发射机预加重和接收机去加重,使得输出信噪比得到的改善值为

$$I = \frac{\text{无预加重和去加重的平均输出噪声功率}}{\text{采用预加重和去加重的平均输出噪声功率}} \quad (2.159)$$

I 称为改善因子。在前面已得出,未采用预加重和去加重时,平均输出噪声功率等于 $(2N_0 W^3 / 3A_c^2)$,见式(2.148)。约去公共项后,可将改善因子 I 表示为

$$I = \frac{2W^3}{3 \int_{-W}^W f^2 |H_{de}(f)|^2 df} \quad (2.160)$$

必须强调的是,只有在接收机中鉴频器输入载噪比很大的情况下,上式才成立。

例 2.6

在实际中广泛应用的一种加重高频的简单预加重滤波器的频率响应定义如下:

$$H_{pe}(f) = 1 + \frac{if}{f_0}$$

如果在有用频带内,有 $R \ll r$ 和 $2\pi f Cr \ll 1$,则该式为图 2.50(a) 中 RC 放大器网络对应的频率响应。图 2.50(a) 中的放大器是为了弥补低频时由 RC 网络造成的衰减。频率参

数 f_0 为 $1/(2\pi Cr)$ 。

接收机中相应的去加重滤波器的频率响应为

$$H_{de}(f) = \frac{1}{1 + jf/f_0}$$

这可用图 2.50(b) 所示中的简单 RC 网络实现。

由于采用了图 2.50 所示的预加重和去加重, FM 接收机输出信噪比的改善量为

$$\begin{aligned} I &= \frac{2W^3}{3 \int_{-W}^W \frac{f^2 df}{1 + (f/f_0)^2}} \\ &= \frac{(W/f_0)^3}{3[(W/f_0) - \tan^{-1}(W/f_0)]} \end{aligned} \quad (2.161)$$

在商用 FM 广播中, 通常取 $f_0 = 2.1 \text{ kHz}$, 并合理地假设 $W = 15 \text{ kHz}$ 。由这些值算出 $I = 22$, 相应的接收机输出信噪比改善量为 13 dB。无预加重和去加重时的输出信噪比的典型值为 40~50 dB。因此, 采用图 2.50 中的简单预加重和去加重滤波器后, 大大提高了接收机的抗噪声性能。

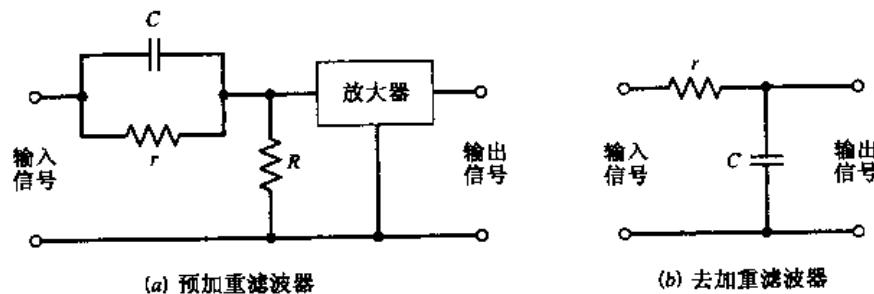


图 2.50 预加重和去加重滤波器

利用系统信号与噪声性质之间的差异, 提高 FM 系统性能的一个例子就是采用简单线性预加重和去加重滤波器。这些简单的滤波器也应用于磁带录音中。特别地, 非线性预加重和去加重技术成功地应用在磁带记录中。尤其在信号功率很低时, 这些技术^[14] (如 Dolby-A, Dolby-B 和 DBX 系统) 利用了组合滤波器和压缩动态范围来降低噪声的影响。

2.14 计算机实验: 锁相环

本节实验主要研究锁相环在 FM 信号的解调中的应用。在实验开始前, 先对锁相环理论做一个简单的阐述。

基本来说, 锁相环由三个主要部件组成: 乘法器、环路滤波器和压控振荡器(VCO), 三者共同组成一个反馈系统, 如图 2.51。VCO 是正弦信号发生器, 其频率由外部源提供的电压控制。事实上, 任何频率调制器都可当成 VCO。假定开始时已对 VCO 进行了适当的调节, 以保证当控制电压为 0 时, 满足以下两点要求:

1. VCO 的频率准确置于未调制载频 f_c 处。
2. 相对于未调制载波, VCO 的输出存在 90° 的相移。

然后,假设锁相环输入的 FM 信号定义式为

$$s(t) = A_c \sin[2\pi f_c t + \phi_1(t)]$$

其中, A_c 为载波幅度,对于调制信号 $m(t)$,角度 $\phi_1(t)$ 与 $m(t)$ 之间存在以下积分关系:

$$\phi_1(t) = 2\pi k_f \int_0^t m(\tau) d\tau$$

其中, k_f 为 VCO 的频率灵敏度。令锁相环里 VCO 的输出为

$$r(t) = A_v \cos[2\pi f_c t + \phi_2(t)]$$

其中, A_v 为幅度。在 VCO 输入端加上控制电压 $v(t)$,则角度 $\phi_2(t)$ 与 $v(t)$ 为以下积分关系:

$$\phi_2(t) = 2\pi k_v \int_0^t v(\tau) d\tau \quad (2.162)$$

其中, k_v 是 VCO 的频率灵敏度,单位为 Hz/V。锁相环的目的是使 VCO 输出 $r(t)$ 的相角与输入信号 $s(t)$ 相同(除了 90°的固定相位差之外)。 $s(t)$ 中随时间变化的相角 $\phi_1(t)$ 与消息信号 $m(t)$ 的调制有关,因此我们希望从 $s(t)$ 中恢复出 $\phi_1(t)$,从而得到 $m(t)$ 的一个估计。在锁相环的其他应用里,输入信号 $s(t)$ 中随时间变化的角度 $\phi_1(t)$ 可能因通信信道的扰动而产生无用的相移。在后面将述及的情况里,为了进行相干检测(同步解调),将通过跟踪 $\phi_1(t)$ 来产生与具有相同相角的信号。

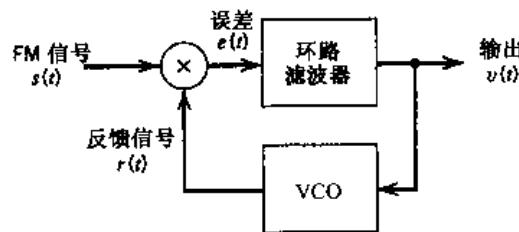


图 2.51 锁相环

2.14.1 锁相环^[15]的模型

为了进一步理解锁相环,有必要建立一个锁相环模型。为简化分析,先建立一个非线性模型,然后将其线性化。根据图 2.51,加在乘法器输入端的输入 FM 信号 $s(t)$ 与 VCO 输出信号 $r(t)$ 产生了两个分量:

1. 高频分量,用倍频项表示:

$$k_m A_c A_v \sin[4\pi f_c t + \phi_1(t) + \phi_2(t)]$$

2. 低频分量,用差频项表示:

$$k_m A_c A_v \sin[\phi_1(t) - \phi_2(t)]$$

其中, k_m 为乘法器增益,单位为 V^{-1} 。

锁相环中的环路滤波器是一个低通滤波器,它能够去除高频分量的响应。同时 VCO 也使得高频分量发生衰减。因此,可舍弃高频分量(即倍频项),将环路滤波器的输入简化为

$$e(t) = k_m A_c A_v \sin[\phi_e(t)] \quad (2.163)$$

其中, $\phi_e(t)$ 为相位误差, 定义如下:

$$\begin{aligned} \phi_e(t) &= \phi_1(t) - \phi_2(t) \\ &= \phi_1(t) - 2\pi k_v \int_0^t v(\tau) d\tau \end{aligned} \quad (2.164)$$

环路滤波器对误差 $e(t)$ 进行处理后, 产生由卷积定义的输出 $v(t)$:

$$v(t) = \int_{-\infty}^{\infty} e(\tau) b(t - \tau) d\tau \quad (2.165)$$

其中, $b(t)$ 为环路滤波器的冲激响应。用式(2.164)和式(2.165)将 $\phi_e(t)$ 和 $\phi_1(t)$ 联系起来, 得到下面的非线性微积分方程, 该方程描述了锁相环的动态特性:

$$\frac{d\phi_e(t)}{dt} = \frac{d\phi_1(t)}{dt} - 2\pi K_0 \int_{-\infty}^{\infty} \sin[\phi_e(\tau)] b(t - \tau) d\tau \quad (2.166)$$

其中, K_0 为环型增益参数, 其定义为

$$K_0 = k_m k_v A_c A_v \quad (2.167)$$

幅度 A_c 和 A_v 的单位均为 V, 乘法器增益 k_m 单位为 V^{-1} , 频率灵敏度 k_v 的单位为 Hz/V 。因此, 由式(2.167)可知, K_0 的量纲为频率。图 2.52 所示的锁相环模型中还建立了 $e(t)$ 和 $v(t)$ 之间的关系, 如式(2.163)和式(2.165)所示。图 2.52 的模型与图 2.51 中的实际框图相类似。锁相环输入端的乘法器被减法器和正弦非线性器件所代替, VCO 则被积分器所代替。

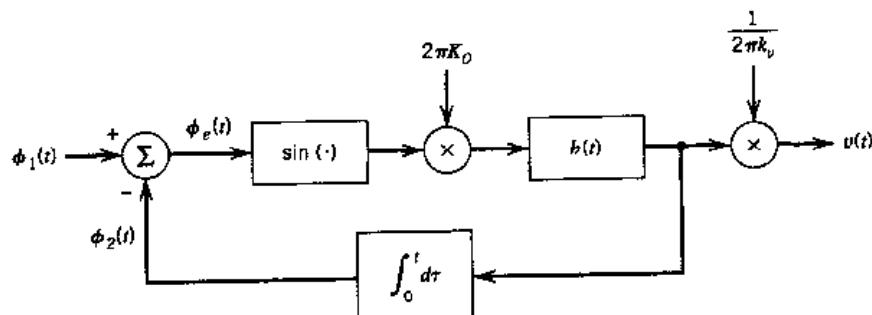


图 2.52 锁相环的非线性模型

图 2.52 模型中的正弦非线性器件使得对锁相环特性的分析工作变得复杂。将此模型线性化有助于简化分析, 同时也可以得到对某些工作模式下锁相环特性的较好的近似描述。当相位误差 $\phi_e(t)$ 为零时, 称锁相环已相位锁定。当 $\phi_e(t)$ 总是小于 1 弧度时, 可用下面的近似式

$$\sin[\phi_e(t)] \approx \phi_e(t)$$

在 $\phi_e(t)$ 小于 0.5 弧度时, 该式的精确度在 4% 以内。在这种情况下, 称锁相环为接近相位锁定, 此时可忽略图 2.52 中的正弦非线性器件。在此条件下, 除了存在比例因子 k_f/k_e 外, $v(t)$ 近似等于 $m(t)$ 。

锁相环的复杂度由环路滤波器的频率响应 $H(f)$ 决定。当 $H(f) = 1$ 时，锁相环为最简形式，即不存在环路滤波器，此时的锁相环称为一阶锁相环。一阶锁相环主要的缺点是，环路增益参数 K_0 同时控制着锁相环的带宽和锁相环的锁定频率范围。锁定频率范围是指锁相环能对输入信号进行相位锁定的频率范围。通过使用频率响应如下的环路滤波器可以克服这个缺点：

$$H(f) = 1 + \frac{a}{jf} \quad (2.168)$$

其中， a 为常量。使用该环路滤波器，并使锁相环以线性模型工作，则由式(2.166)可知锁相环的作用与二阶反馈系统相同，其标准频率响应为

$$\frac{\Phi_e(f)}{\Phi_1(f)} = \frac{(jf/f_n)^2}{1 + 2\zeta(jf/f_n) + (jf/f_n)^2} \quad (2.169)$$

其中， $\Phi_e(f)$ 和 $\Phi_1(f)$ 分别为 $\phi_e(t)$ 和 $\phi_1(t)$ 的傅里叶变换。系统参数为自由振荡频率 f_n ，以及衰减因子 ζ ，其定义分别为

$$f_n = \sqrt{aK_0} \quad (2.170)$$

和

$$\zeta = \sqrt{\frac{K_0}{4a}} \quad (2.171)$$

这里描述的二阶锁相环可由下面的计算机实验来实现。

2.14.2 实验 1：捕获模型

当锁相环用于相干检测(同步解调)时，锁相环必须首先锁定输入信号，然后跟随输入相角随时间的变化而相应变化。使环路进入锁相的过程称为捕获，后继的跟随输入信号角度变化的过程称为跟踪。在捕获模型(也可能是跟踪模型)中，输入信号 $s(t)$ 与 VCO 输出 $r(t)$ 之间的相位误差 $\phi_e(t)$ 将相当大，因此要求采用图 2.52 中的非线性模型。但对捕获过程的非线性分析超出了本书的范围。在本实验中，将通过计算机仿真来研究捕获过程，从而使读者对捕获过程特点的一些认识。

考虑采用了式(2.168)所示环路滤波器的二阶锁相环，并选取如下参数：

$$\text{自由振荡频率 } f_n = \frac{1}{2\pi} \text{ Hz}$$

$$\text{阻尼因子 } \zeta = 0.3, 0.707, 1.0$$

为了适应 ζ 的变化量，滤波器参数 a 应根据下式变化：

$$a = \frac{f_n}{2\zeta}$$

该式由式(2.170)和式(2.171)得到。图 2.53 为阻尼因子 ζ 分别取三个给定值时，对应的相位误差 $\phi_e(t)$ 随时间的变化曲线，设频率间隔为 0.125 Hz。结果表明，阻尼因子 $\zeta = 0.707$ 时，在较快的响应时间和低衰减振荡特性之间有较好的折中效果。

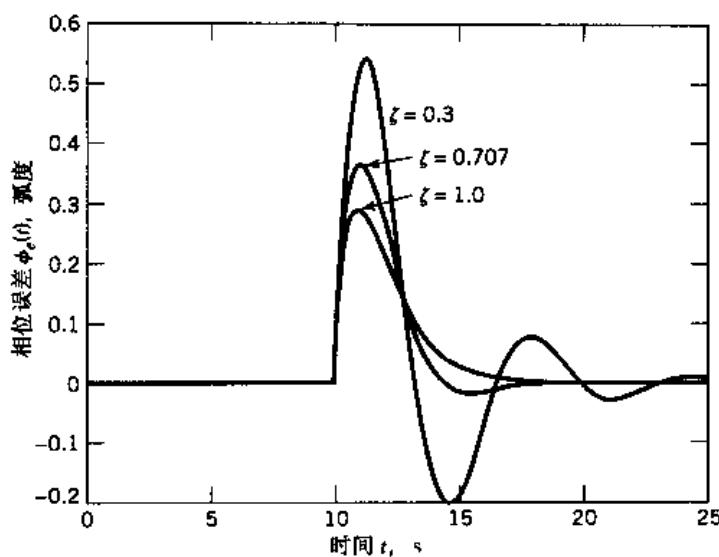


图 2.53 三个不同阻尼因子对应的相位误差变化曲线

2.14.3 实验 2: 相位平面图

相位平面图是一系列轨迹,每一个轨迹对应于式(2.166)的一个解。对于这个实验,画出正弦调制下二阶锁相环的相位平面图,如图 2.54。锁相环的系统参数如下

$$\text{环路增益参数 } K_0 = \frac{50}{2\pi} \text{ Hz}$$

$$\text{环路自由振荡频率 } f_n = \frac{50}{2\sqrt{2}} \text{ Hz}$$

$$\text{正弦调制频率 } f_m = \frac{50}{2\pi\sqrt{2\pi}} \text{ Hz}$$

图 2.54 同时也是临界阻尼因子下锁相环的相位平面图,该图中的轨迹(频率误差对相位误差)对应着不同的起始点。由此图,我们得到以下结论:

1. 对于正弦非线性器件,相位误差 ϕ_e 的相位平面图本身是以 2π 为周期的,但 $d\phi_e/dt$ 的相位平面图是非周期的。
2. 对于初始频率误差

$$\frac{1}{K} \frac{d\phi_e}{dt}$$

其绝对值小于或等于 1,锁相环必在 $(0,0)$ 或 $(2\pi,0)$ 处达到稳定(平衡)点,平衡点呈倍数关系说明相位平面图是周期性的。

3. 对于初始频率误差

$$\frac{1}{K} \frac{d\phi_e}{dt}$$

其绝对值等于 2,点 $(\pi,0)$ 处为鞍点,在此点处加在锁相环上的最轻微扰动都会导致锁相环转到平衡点 $(0,0)$ 或 $(2\pi,0)$ 处。

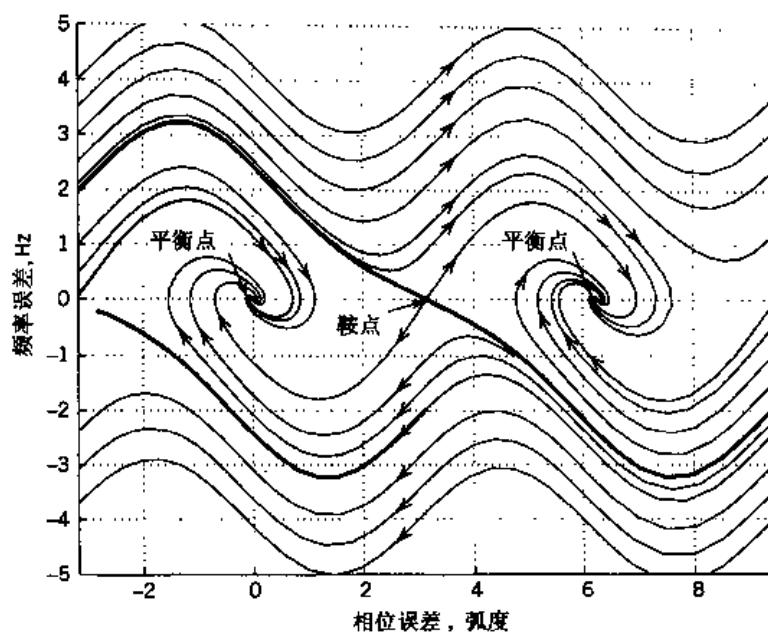


图 2.54 基于临界阻尼因子和正弦调制的相位平面图

2.15 总结与讨论

本章研究了连续波(CW)调制的原理。这种模拟形式的调制采用的是正弦载波，其幅度或角度随消息信号的变化而变化。由此区分两种连续波调制：幅度调制和角度调制。

2.15.1 幅度调制

根据已调信号频谱内容，幅度调制本身可分为四种类型。这四种幅度调制及其优点分别如下：

1. 标准幅度调制(AM)。在这种调制中，上边带和下边带均被发送，并伴有载波。
接收端对 AM 信号的解调相当简单，例如采用包络检波。正因如此，AM 被广泛应用于商用 AM 无线广播中，无线广播系统包括一个功率发射机和很多造价低廉的接收机。
2. 抑制载波的双边带(DSB-SC)调制。在这种调制中，只发送上边带和下边带。
抑制载波意味着发送同样的消息信号时，DSB-SC 调制需要的功率远小于标准 AM 的发射功率，这是 DSB-SC 调制优于标准 AM 的地方，但这同时也增加了接收机的复杂度。因此，DSB-SC 调制适用于只有一个发射机和一个接收机的点对点通信，这种通信形式的发射功率较低，但接收机结构相对复杂。
3. 单边带(SSB)调制。在这种调制中，只发送上边带或下边带。
SSB 调制将消息信号从一点传到另一点时，需要的发射功率和信道带宽均是最小的，从这种意义上说，SSB 调制是最佳的连续波调制形式。但它只能应用于能量间隙中心在零频率的消息信号。

4. 残留边带(VSB)调制。在这种调制中,以互补的方式发送一个边带的几乎所有部分和另一边带的残留部分。

VSB 调制需要的信道带宽介于 SSB 和 DSB-SC 系统之间,如果调制信号带宽很大,那么,节约带宽就具有很重要的意义,例如电视信号和高速率数据信号均采用 VSB 调制。

DSB-SC,SSB 和 VSB 都是线性调制的例子,从严格的意义上说,标准 AM 属于非线性调制。但是,标准 AM 只是稍微背离了线性调制。因此,上述四种幅度调制均可通过傅里叶变换进行频谱分析。

2.15.2 角度调制

角度调制可分为频率调制(FM)和相位调制(PM)。在 FM 中,正弦载波的瞬时频率随消息信号呈正比例变化。而在 PM 中,正弦载波的相位随消息信号呈正比例变化。在相位对时间的导数中去掉比例因子 $1/(2\pi)$,得到的结果称为瞬时频率。因此,FM 和 PM 是相互关联的,知道了其中一个的特性,就能够确定另一个的特性。由于这个原因,同时也因为广播中广泛采用了 FM 技术,因此在本章的角度调制内容中,大部分篇幅都是介绍 FM 的。

与 AM 不同,FM 是非线性调制过程。所以,FM 的频谱分析比 AM 困难得多。通过对单音频 FM 的研究,能够使读者对 FM 的频谱特性有较多的认识。特别地,为了近似地计算 FM 的传输带宽 B_T ,还推导出了卡逊准则这一经验公式。根据该准则, B_T 由一个参数控制:在正弦 FM 中,这个参数为调制指数 β ;在非正弦 FM 中,这个参数则为偏移率 D 。

2.15.3 噪声分析

将各种连续波调制系统的抗噪声性能进行比较,从而对本章的讨论进行总结。这里假设对正弦信号进行调制。为使比较具有一定意义,还同时假定了各种调制系统有相同的信道信噪比。对于以下调制方法,画出了输出信噪比对信道信噪比的变化曲线,如图 2.55 所示:

- 100% 调制的标准 AM
- 相干 DSB-SC,SSB
- $\beta = 2$ 和 $\beta = 5$ 的 FM

图 2.55 还包括了 AM 和 FM 的门限效应。比较中应注意各种调制系统的传输带宽要求。基于这个考虑,采用归一化传输带宽,其定义为

$$B_n = \frac{B_T}{W}$$

其中, B_T 为已调信号的传输带宽, W 为消息带宽。表 2.4 为不同连续波调制方案下的 B_n 值。

由图 2.55 和表 2.4,可得以下结论:

- 在 AM 系统中,从噪声性能和保留带宽看,SSB 调制是最佳的。
- FM 抗噪声性能的改善是以额外的传输带宽为代价的。这里假设 FM 系统工作在能实现抗噪声性能改善的门限之上。

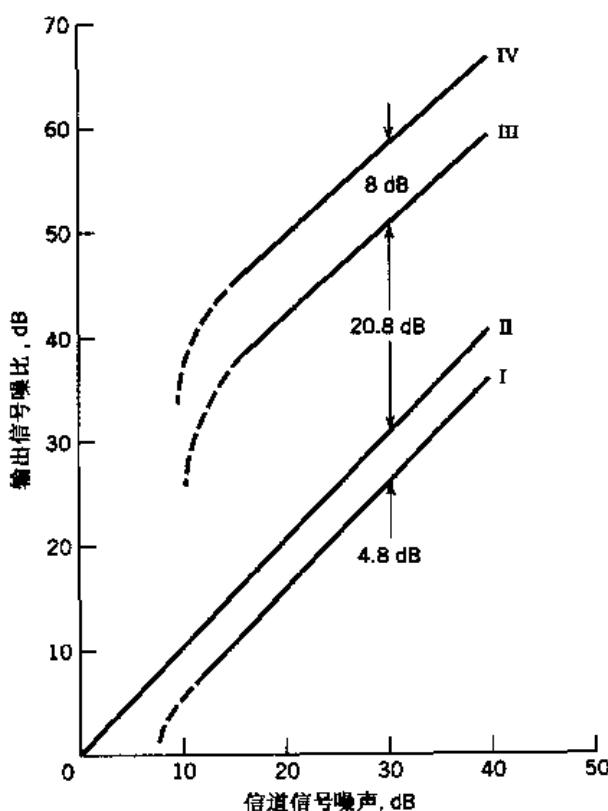


图 2.55 各种连续波调制系统噪声性能的比较。曲线 I: 全幅 AM,
 $\mu = 1$ 。曲线 II: DSB-SC, SSB。曲线 III: FM, $\beta = 2$ 。曲线 IV:
 FM, $\beta = 5$ 。(曲线 III 和 IV 包含 13dB 的预加重和去加重)

表 2.4 各种连续波调制方案的 B_n 值

B_n	AM, DSB-SC		SSB		FM	
					$\beta = 2$	$\beta = 5$
	2		1		8	16

关于连续波调制有一个重要的结论, 只有频率调制可用传输带宽来换取抗噪声性能的改善。这种权衡遵循平方律, 平方律是在连续波调制(即模拟通信)中能达到的最好规律。在第3章里将介绍脉冲编码调制(PCM), 这种调制方法是用数字通信系统来传输承载信息的模拟信号的基础, 具有较好的性能。

注释与参考

- [1] 20世纪初 Reginald Fessenden 首次使用了连续波以及外差这两个术语。
- [2] 科斯塔接收机的命名是为了纪念其发明者, 见 Costas(1956)。
- [3] 贝塞尔函数在模拟和数字通信系统的研究中起着重要的作用。它分为第一类和第二类。附录3讨论了这两类贝塞尔函数的性质和数学细节。第一类贝塞尔函数表见表 A6.5。
- [4] 关于 FM 信号带宽的卡逊准则是为了纪念它的创始者而命名的。Carson 和 Fry

- (1937)撰写了一篇关于频率调制理论的经典论文。
- [5] Armstrong(1936)首次提出了间接产生宽带 FM 信号的方法。频率调制的抗噪特性也是由 Armstrong 最先认识的。
- [6] 立体声复用通常采用无线传输中的频率调制,也可采用习题 2.14 中讨论的幅度调制。更多的细节见 Mennie(1978)。
- [7] 关于超外差接收机的细节描述见 Henney(1958, pp. 19-34 – 19-41)编辑的无线工程手册。
- [8] 这里提出的包络检波中关于门限的定性分析参见 Downing(1964, p. 17)的论文。
- [9] 简化式(2.142)时所做的临界假设见 Rice(1963)的论文。
- [10] 关于 FM 接收机门限效应的详细讨论见 Rice(1963)以及 Schwartz, Bennett 与 Stein (1966, pp. 129 – 163)。
- [11] 图 2.45 来源于 Rice(1963)。实验证明了此图所示理论曲线 II 是有效的,见 Schwartz, Bennett 和 Stein (1966, pp. 153)。关于 FM 中门限现象的早期实验工作见 Crosby (1937)。
- [12] Chaffee(1939)最早提出了在 FM 解调器中采用反馈这一思想。
- [13] 2.13 节提出的 FMFB 解调器是以 Enloe(1962)为基础的,同时也可参阅 Roberts(1977, pp. 166 – 181)。
- [14] 关于 2.13 节后半部分提到的 Dolby 系统的详细讨论见 Stremler(1990, pp. 732 – 734)的论文。
- [15] 关于锁相环的完整非线性分析见 Gardner(1979), Lindsey(1972)和 Viterbi(1966)的论文。

习题

幅度调制

2.1 假设非线性器件的输出电流 i_o 和输入电压 v_i , 满足下面的关系式:

$$i_o = a_1 v_i + a_3 v_i^3$$

其中, a_1 和 a_3 为常数。解释该器件怎样用做:(a) 乘积调制器;(b) 幅度调制器。

2.2 图 P2.2 为平方律调制器的电路图。加在非线性器件上的信号很弱,因而可用二次方表示:

$$v_2(t) = a_1 v_1(t) + a_2 v_1^2(t)$$

其中, a_1 和 a_2 是常数, $v_1(t)$ 是输入电压, $v_2(t)$ 为输出电压。输入电压为

$$v_1(t) = A_c \cos(2\pi f_c t) + m(t)$$

其中, $m(t)$ 是消息信号, $A_c \cos(2\pi f_c t)$ 为载波。

(a) 计算输出电压 $v_2(t)$ 。

(b) 确定图 P2.2 调谐电路的频率响应,使之产生载频为 f_c 的 AM 信号。

(c) 此 AM 信号的调幅灵敏度是多少?

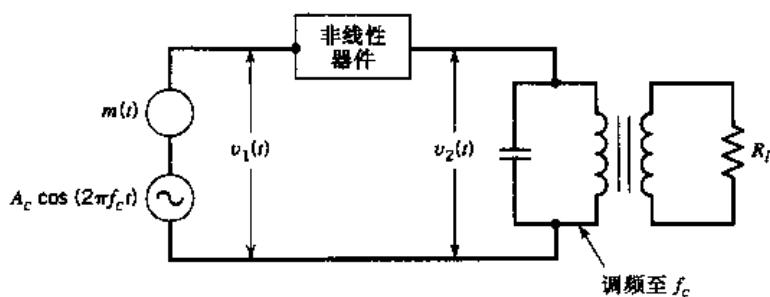


图 P2.2

2.3 图 2.3(a) 为开关调制器的电路图。假设加在二极管上的载波 $c(t)$ 幅值非常大,使得二极管可被视为理想的开关;仅当正偏(即 $c(t) > 0$)时导通。因此可用如下分段线性特性来近似描述二极管 - 负载电阻器的转移特性(见图 P2.3(b)):

$$v_2(t) = \begin{cases} v_1(t), & c(t) > 0 \\ 0, & c(t) < 0 \end{cases}$$

即负载电压 $v_2(t)$ 在 $v_1(t)$ 值与零值之间周期性地变化,变化率等于载频 f_c 。因此可记为

$$v_2(t) \approx [A_c \cos(2\pi f_c t) + m(t)]g_{T_0}(t)$$

其中, $g_{T_0}(t)$ 为周期性脉冲串,其定义如下:

$$g_{T_0}(t) = \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} \cos[2\pi f_c t(2n-1)]$$

- (a) 求输出电压 $v_2(t)$ 所含的 AM 信号分量。
 (b) 指出 $v_2(t)$ 中需要被带通滤波器去掉的无用分量。

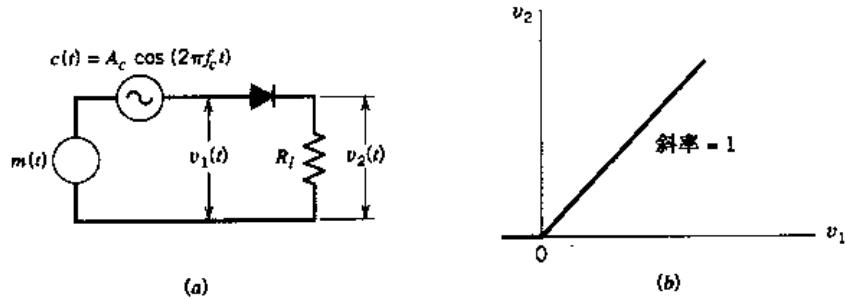


图 P2.3

2.4 考虑由频率为 f_m 的正弦调制信号产生的 AM 信号:

$$s(t) = A_c [1 + \mu \cos(2\pi f_m t)] \cos(2\pi f_c t)$$

设调制因子 $\mu = 2$, 载频 f_c 远大于 f_m 。将 AM 信号 $s(t)$ 加在理想的包络检波器上, 得到输出 $v(t)$ 。

- (a) 确定输出 $v(t)$ 的傅里叶级数。
 (b) $v(t)$ 中二次谐波与基波的幅度之比是多少?

2.5 图 2.5 为包络检波器的电路图。包络检波器由一个二极管和一个简单的电阻 - 电

容(RC)滤波器构成。在输入信号的正半周里,二极管正偏,电容 C 迅速充电达到输入信号的峰值。当输入信号低于峰值时,二极管反偏,电容 C 通过负载电阻 R_l 缓慢放电。放电过程一直持续到下一个正半周为止。周而复始,不断重复着充电-放电的过程。

- 指出当二极管导通时,电容 C 迅速充电并达到输入电压正峰值的条件。
- 电容 C 在载波的正峰值之间缓慢放电,且电容放电的最大速率不会超过调制信号的频率,指出满足这两点要求的负载电阻 R_l 需满足的条件。

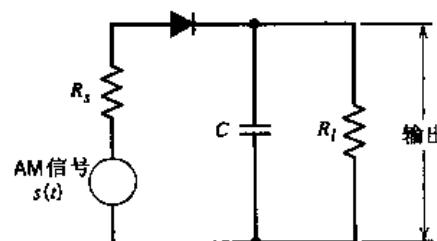


图 P2.5

2.6 采用具有如下转移特性的非线性器件的平方律检测器:

$$v_2(t) = a_1 v_1(t) + a_2 v_1^2(t)$$

其中, a_1 和 a_2 为常量, $v_1(t)$ 为输入, $v_2(t)$ 为输出。输入为 AM 信号

$$v_1(t) = A_c[1 + k_a m(t)] \cos(2\pi f_c t)$$

- 计算输出 $v_2(t)$ 。
- 求可从 $v_2(t)$ 中恢复出消息信号 $m(t)$ 的条件。

2.7 将 AM 信号

$$s(t) = A_c[1 + k_a m(t)] \cos(2\pi f_c t)$$

加在图 P2.7 所示的系统中。假设对于所有 t 值,都有 $|k_a m(t)| < 1$, 消息信号 $m(t)$ 限制在间隔 $-W \leq f \leq W$ 内, 载频 $f_c > 2W$, 说明 $m(t)$ 是怎样从方根器输出 $v_3(t)$ 中恢复出来的。



图 P2.7

2.8 消息信号 $m(t)$ 的频谱如图 P2.8 所示。消息带宽 $W = 1$ kHz。信号和载波 $A_c \cos(2\pi f_c t)$ 一起被加在乘积调制器输入端,产生 DSB-SC 已调信号 $s(t)$ 。得到的已调信号被送入相干检测器中。假定调制器和检测器的载波之间已达到很好的同步,求出满足下面条件时检测器的输出频谱图:

- 载频 $f_c = 1.25$ kHz。
- 载频 $f_c = 0.75$ kHz。

已调信号 $s(t)$ 的每个分量都由 $m(t)$ 惟一确定的最低载频是多少?

- 2.9 图 P2.9 为平衡调制器的电路图。上方 AM 调制器的输入为 $m(t)$, 而下方 AM 调制器的输入是 $-m(t)$ 。这两个调制器的调幅灵敏度相同。证明平衡调制器的输出 $s(t)$ 为 DSB-SC 已调信号。

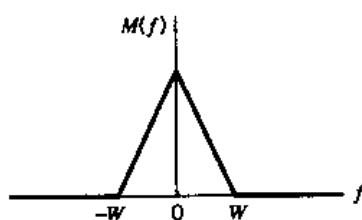


图 P2.8

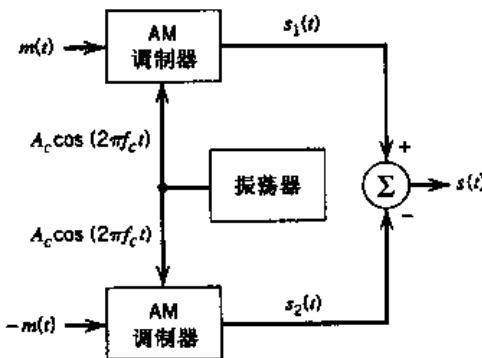


图 P2.9

- 2.10 将 DSB-SC 已调信号送入相干检测器进行解调。

- 当检测器的本地载频与输入 DSB-SC 信号载频存在频率误差 Δf 时, 估算 Δf 会产生怎样的影响。
- 对于正弦调制信号, 证明: 由于上述载频误差, 解调出来的信号在该误差频率上存在差拍现象。并画出解调后的信号草图来加以说明。

- 2.11 有如下 DSB-SC 信号:

$$s(t) = A_c \cos(2\pi f_c t)m(t)$$

其中, $A_c \cos(2\pi f_c t)$ 为载波, $m(t)$ 为消息信号。将该已调信号加在平方律器件上, 该器件的特性为

$$y(t) = s^2(t)$$

然后, 将输出 $y(t)$ 加在通带幅度响应为 1, 中心频率为 $2f_c$, 带宽为 Δf 的窄带滤波器上。假设 Δf 足够小, 使得 $y(t)$ 的频谱在滤波器的通带内为常量。

- 计算平方律器件输出 $y(t)$ 的频谱。
- 证明滤波器的输出 $v(t)$ 近似地为正弦型, 即

$$v(t) \approx \frac{A_c^2}{2} E \Delta f \cos(4\pi f_c t)$$

式中, E 为消息信号 $m(t)$ 的能量。

- 2.12 考虑图 2.10 中的正交载波复用系统。将图 2.10(a) 所示发射机输出的复用信号 $s(t)$ 送入频率响应为 $H(f)$ 的通信信道中。将信道的输出信号加在图 2.10(b) 所示接收机的输入端。证明能够在接收机输出端恢复出消息信号 $m_1(t)$ 和 $m_2(t)$ 的必要条件为

$$H(f_c + f) = H^*(f_c - f), \quad 0 \leq f \leq W$$

其中, f_c 为载频, W 为消息带宽。提示: 计算接收机两个输出的频谱。

- 2.13 假定在图 2.10(b)所示正交载波复用系统的接收机中, 用于解调的本地载波与发射机中载波源之间存在相位误差 ϕ 。并假定发射机和接收机之间的通信信道是无失真的, 证明该相位误差会导致接收机输出端两个解调信号之间的串话。串话是指接收机输出端某一消息信号的一部分包含在了另一个消息信号中, 反之亦然。
- 2.14 在 AM 立体声系统的一种特殊形式中, 采用了正交复用技术。载波 $A_c \cos(2\pi f_c t)$ 用于调制和信号

$$m_1(t) = V_0 + m_e(t) + m_r(t)$$

其中, V_0 为用于发送载波分量的直流偏移量, $m_e(t)$ 为左侧音频信号, $m_r(t)$ 为右侧音频信号。正交载波 $A_c \sin(2\pi f_c t)$ 用于调制差信号

$$m_2(t) = m_e(t) - m_r(t)$$

- (a) 证明可用包络检波器从正交复用信号中恢复出和信号 $m_r(t) + m_e(t)$ 。怎样才能将包络检波造成的信号失真降至最小?
- (b) 证明相干检测器可恢复出差信号 $m_e(t) - m_r(t)$ 。
- (c) 最后怎样获得有用信号 $m_r(t)$ 和 $m_e(t)$?

- 2.15 消息信号为

$$m(t) = \frac{1}{1+t^2}$$

确定并大致画出采用下列调制方法得到的已调信号形:

- (a) 调制指数为 50 的幅度调制(AM)。
- (b) 抑制载波双边带(DSB-SC)调制。
- (c) 只发送上边带的单边带(SSB)调制。
- (d) 只发送下边带的单边带(SSB)调制。
- 2.16 存在傅里叶变换的信号 $m(t)$, 其希尔伯特变换(Hilbert transform)用 $\hat{m}(t)$ 表示, 定义式为(见附录 2):

$$\hat{m}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{m(\tau)}{t - \tau} d\tau$$

在频率范围内, 有

$$\hat{M}(f) = -j \operatorname{sgn}(f) M(f)$$

其中, $m(t) \Leftrightarrow M(f)$, $\hat{m}(t) \Leftrightarrow \hat{M}(f)$, $\operatorname{sgn}(f)$ 为符号函数。

利用希尔伯特变换的定义式, 证明由消息信号 $m(t)$ 和幅度为 1 的载波 $\cos(2\pi f_c t)$ 产生的单边带已调信号, 其表达式如下(见表 2.1):

$$s(t) = \frac{1}{2} m(t) \cos(2\pi f_c t) \pm \frac{1}{2} \hat{m}(t) \sin(2\pi f_c t)$$

其中, 发送上边带时取“-”号, 发送下边带时取“+”号。

- 2.17 用于解调 SSB 信号 $s(t)$ 的本地振荡器与产生 $s(t)$ 的载频 f_c 间存在频率误差 Δf , 因此, 接收端的本地振荡器与发射端载波发射机之间不能达到完全同步。计算以下

两种情况中的解调信号：

- 只含上边带的 SSB 信号 $s(t)$ 。
- 只含下边带的 SSB 信号 $s(t)$ 。

2.18 图 P2.18 为产生 SSB 已调信号的 Weaver 方法的框图。消息(调制)信号 $m(t)$ 限制在频带 $f_a \leq |f| \leq f_b$ 内。加在第一对乘积调制器上的副载波频率为 f_0 , f_0 为频带的中心频率, 即

$$f_0 = \frac{f_a + f_b}{2}$$

同相信道与正交信道中的低通滤波器完全相同, 它们的截止频率均等于 $(f_b - f_a)/2$ 。加在第二对乘积调制器上的载波频率为 f_c , f_c 大于 $(f_b - f_a)/2$ 。大致画出图 P2.18 所示调制器中各点的频谱图, 并由频谱图证明以下几点:

- 对于下边带, 同相信道和正交信道产生了两个极性相反的信号, 在调制器输出端将这两个信号相加, 就抑制了下边带。
- 对于上边带, 同相信道和正交信道产生了两个极性相同的信号, 在调制器输出端将这两个信号相加后, 就发送了上边带。
- 对图 P2.18 应做怎样的改动, 才能使它只发送下边带?

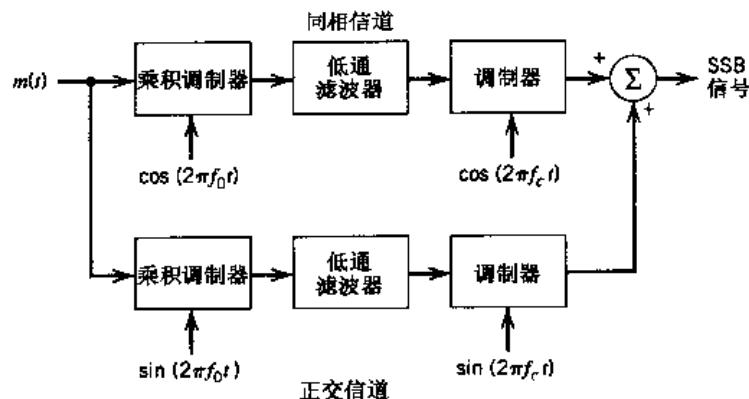


图 P2.18

2.19 语音信号 $m(t)$ 的频谱在间隔 $f_a \leq |f| \leq f_b$ 以外为零。为使通信保密, 将信号加在以下部件级联组成的扰频器上: 乘积调制器, 高通滤波器, 另一个是乘积调制器和低通滤波器。加在第一个乘积调制器上的载频为 f_c , 而加在第二个乘积调制器上的载频为 $f_b + f_c$ 。两个载波的幅度均为 1。高通滤波器和低通滤波器有相同的截止频率 f_c 。假定 $f_c > f_b$ 。

- 推导出扰频器的输出 $s(t)$ 的表达式, 画出其频谱草图。
- 证明通过与扰频器完全相同的矫正器后, 可从 $s(t)$ 中恢复出原始语音信号 $m(t)$ 。

2.20 推导式(2.16)中滤波器频率响应 $H_0(f)$ 的表达式, 此滤波器对消息信号 $m(t)$ 进行处理后, 产生了 VSB 调制所需的 $m'(t)$ 。信号 $m'(t)$ 去掉比例因子后构成了 $s(t)$ 的正交分量。为了进行推导, 将 $s(t)$ 加在图 P2.20 所示的相干检测器上, 用于恢复与原始消息信号 $m(t)$ 成比例的信号。

- (a) 根据图 2.12,画出 VSB 已调信号的框图。根据消息信号 $m(t)$ 的傅里叶变换和图 2.12 所示带通滤波器的频率响应 $H(f)$, 确定图 P2.20 中乘积调制器输出 $v(t)$ 的傅里叶变换 $V(f)$ 。
- (b) 通过计算图 P2.20 中低通滤波器输出信号的傅里叶变换, 确定为完全恢复原始消息信号 $m(t)$, $H_Q(f)$ 与 $H(f)$ 必须满足的条件(不包括比例因子)。

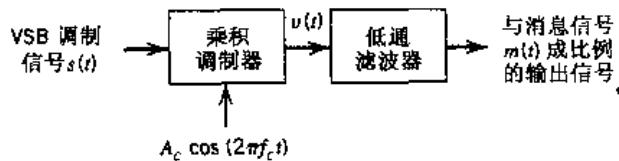


图 P2.20

2.21 单音频调制信号 $m(t) = A_m \cos(2\pi f_m t)$ 用于产生 VSB 信号

$$s(t) = \frac{1}{2} a A_m A_c \cos[2\pi(f_c + f_m)t] + \frac{1}{2} A_m A_c (1 - a) \cos[2\pi(f_c - f_m)t]$$

式中, a 为小于 1 的常量, 表示上边频的衰减。

- (a) 找出 VSB 信号 $s(t)$ 的正交分量。
 (b) 将 VSB 信号和载波 $A_c \cos(2\pi f_c t)$ 通过包络检波器, 确定由正交分量引起的失真。
 (c) 常数 a 取何值时, 失真最严重?
- 2.22 本题研究超外差接收机中混频器的原理。图 P2.22 为由乘积调制器、具有可变频率 f_l 的本地振荡器以及带通滤波器级联而成的混频器的框图。输入信号为 AM 信号, 其带宽为 10 kHz, 载频为 0.535 MHz 到 1.605 MHz 之间的任意值。这些参数是 AM 无线广播中用到的典型值。混频器需将此 AM 信号搬移至中心频率固定在中频 0.455 MHz 的频带上。求出满足此要求的本地振荡器的调频范围。

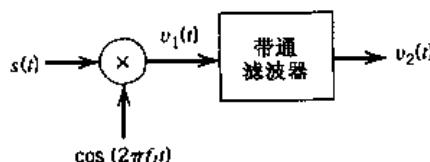


图 P2.22

2.23 图 P2.23 为外差式频谱分析器的框图。它由变频振荡器、乘法器、带通滤波器和均方根仪(RMS meter)组成。振荡器幅度为 A , 工作在 f_0 到 $f_0 + W$ 范围内, f_0 为滤波器的中心频率, W 为信号带宽。假设 $f_0 = 2 W$, 滤波器带宽 Δf 小于 f_0 , 通带范围内滤波器的幅度响应为 1。若输入信号为低通信号 $g(t)$, 确定均方根仪的输出值。

角度调制

- 2.24 画出由图 P2.24 所示锯齿波产生的 PM 和 FM 信号的波形草图。
- 2.25 在调频雷达中, 发射载波的瞬时频率变化如图 P2.25 所示, 这是由三角波调制信号产生的。图 P2.25 中还用虚线画出了接收到的反射信号的瞬时频率。 τ 为往返时间。将发射信号和接收到的反射信号加在混频器上, 保留差频分量。假定 $f_0 \tau \ll 1$,

计算 1 s 内混频器输出的平均拍数, 用载频频偏峰值 Δf , 时延 τ 和发射信号的重复频率 f_0 表示。

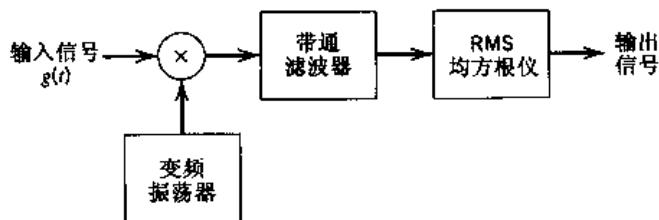


图 P2.23

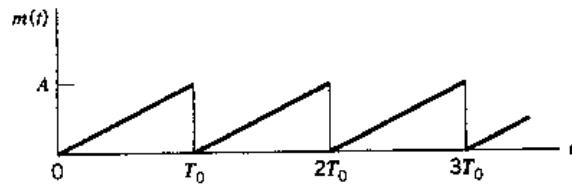


图 P2.24

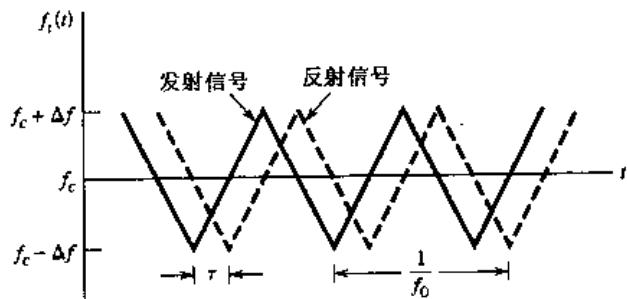


图 P2.25

2.26 当 $|t| \leq T/2$, 正弦信号的瞬时频率等于 $f_c - \Delta f$; 当 $|t| > T/2$, 正弦信号的瞬时频率等于 f_c , 计算此调频信号的频谱。提示: 将时间分为三段间隔: $-\infty < t < -T/2$, $-T/2 \leq t \leq T/2$, $T/2 < t < \infty$ 。

2.27 SSB 调制可看成幅度调制和频率调制的一种复合调制形式。试计算在下面两种情况下, SSB 信号的包络和瞬时频率。

- (a) 只发送上边带;
- (b) 只发送下边带。

2.28 窄带 FM 信号的近似表达式如下:

$$s(t) \approx A_c \cos(2\pi f_c t) - \beta A_c \sin(2\pi f_c t) \sin(2\pi f_m t)$$

- (a) 计算该已调信号的包络, 包络的最大值与最小值之比是多少? 设 $0 \leq \beta \leq 0.3$, 画出此比值对 β 的曲线。
- (b) 计算此窄带 FM 信号的平均功率, 用未调制载波平均功率的百分比表示。设 $0 \leq \beta \leq 0.3$, 画出平均功率对 β 的曲线。
- (c) 将窄带 FM 信号 $s(t)$ 的角度 $\theta_c(t)$ 展成幂级数形式, 限定调制指数 β 的最大值为 0.3 弧度, 证明:

$$\theta_i(t) \approx 2\pi f_c t + \beta \sin(2\pi f_m t) - \frac{\beta^3}{3} \sin^3(2\pi f_m t)$$

在 $\beta = 0.3$ 时, 三次谐波功率与基波功率之比是多少?

2.29 将正弦调制信号

$$m(t) = A_m \cos(2\pi f_m t)$$

加在相位灵敏度为 k_p 的相位调制器上。未调制载波的幅度为 A_c , 频率为 f_c 。

- (a) 计算 PM 信号的频谱, 假定最大相偏 $\beta_p = k_p A_m$ 不大于 0.3 弧度。
 - (b) 画出已调信号的矢量图, 并与相应的窄带 FM 信号矢量图进行比较。
- 2.30 假定习题 2.29 中 PM 信号的最大相偏 β_p 取任意值。将此已调信号加在理想的带通滤波器上, 滤波器的中心频率为 f_c , 通带从 $f_c - 1.5f_m$ 到 $f_c + 1.5f_m$ 。计算滤波器输出信号的包络值、相位以及瞬时频率关于时间的函数。
- 2.31 载波经频率为 f_m , 幅度为 A_m 的正弦信号进行频率调制。
- (a) 当 FM 信号的载波分量降为零时, 计算对应的调制指数 β 值。计算中可采用表 A6.5 中的 $J_0(\beta)$ 值。
 - (b) 令 $f_m = 1$ kHz, 增大 A_m (从 0 V 开始), 会发现当 $A_m = 2$ V 时, FM 信号的载波分量第一次降为零, 调制器的频率灵敏度是多少? 当载波分量第二次降为零时, A_m 值为多大?
- 2.32 将调制指数 $\beta = 1$ 的 FM 信号通过理想的带通滤波器发送, 滤波器的中心频率为 f_c , 带宽为 $5f_m$, f_c 是载频, f_m 是正弦调制信号的频率。计算滤波器的输出幅度谱。
- 2.33 将频率为 100 MHz 的载波用幅度为 20 V, 频率 100 kHz 的正弦信号进行频率调制。调制器的频率灵敏度为 25 kHz/V。
- (a) 用卡逊准则计算 FM 信号的近似带宽。
 - (b) 当只发送幅值大于未调载波幅值 1% 的边频时, 计算传输带宽。计算中使用图 2.26 中的标准曲线。
 - (c) 当调制信号的幅度加倍时, 重复上述计算。
 - (d) 当调制频率加倍时, 重复上述计算。
- 2.34 由正弦调制信号 $A_m \cos(2\pi f_m t)$ 产生宽带 PM 信号, 调制器的相位灵敏度等于 k_p 弧度每伏。
- (a) 若 PM 信号的最大相偏值大于 1 弧度, 证明 PM 信号的带宽随调制频率 f_m 呈线性变化。
 - (b) 比较宽带 PM 信号与宽带 FM 信号的性质。
- 2.35 图 P2.35 为按频率调制原理工作的实时频谱分析仪框图。将已知信号 $g(t)$ 和调频信号 $s(t)$ 加在乘法器上, 再将输出 $g(t)s(t)$ 送入冲激响应为 $b(t)$ 的滤波器里。 $s(t)$ 和 $b(t)$ 为线性 FM 信号, 它们的瞬时频率均随时间线性变化, 但二者变化方向相反, 表达式为

$$s(t) = \cos(2\pi f_c t - \pi kt^2)$$

$$b(t) = \cos(2\pi f_c t + \pi kt^2)$$

其中, k 为常数。证明滤波器输出信号的包络值与输入信号 $g(t)$ 的幅度谱成正比, 这里的 ks 起到频率 f 的作用。提示: 用附录 2 中分析带通信号和带通滤波器的复数表示法。

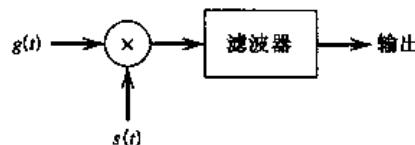


图 P2.35

- 2.36 将频偏为 10 kHz, 调制频率为 5 kHz 的 FM 信号加在级联的两个频率乘法器上。第一个乘法器使将频率值乘以 2, 第二个乘法器将频率值乘以 3。计算第二个乘法器输出 FM 信号的频偏值和调制指数。此 FM 信号的相邻边频之间的间隔是多少?

- 2.37 图 P2.37 为间接调频的宽带频率调制器的框图。调制器用于发送频率范围从 100 Hz 到 15 kHz 的音频信号。由一个晶体振荡器产生的频率 $f_1 = 0.1 \text{ MHz}$ 的载频被送入窄带相位调制器。另一个晶体振荡器则为混频器提供频率为 9.5 MHz 的正弦信号。对系统的具体说明如下:

发射机输出端的载频 $f_c = 100 \text{ MHz}$

最小频偏 $\Delta f = 75 \text{ kHz}$

相位调制器的最大调制指数为 0.3 弧度

(a) 计算上述条件下, 倍频数 n_1 和 n_2 (混频器之前和之后) 的取值。

(b) 标出图 P2.37 调制器中各点的载频值和频偏值。

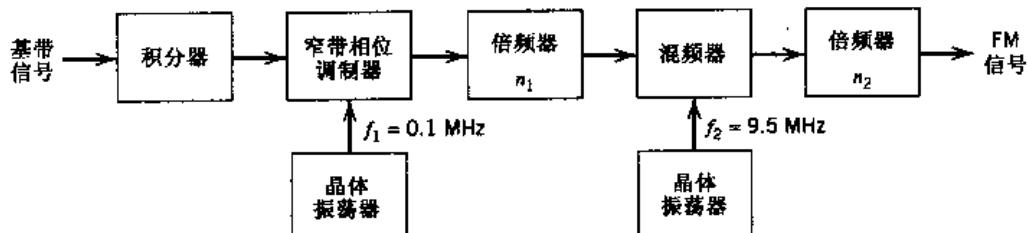


图 P2.37

- 2.38 图 P2.38 为压控振荡器的选频网络。将调制信号 $A_m \sin(2\pi f_m t)$ 与偏压 V_b 加在一对变容二极管上来进行频率调制, 这对变容二极管与一个 $200 \mu\text{H}$ 的电感和 100 pF 的电容相并联。每个二极管的容量与加在其电极两端的电压 V (单位为 V) 的关系式为

$$C = 100V^{-1/2} \text{ pF}$$

未调振荡频率是 1 MHz。将 VCO 的输出加在倍频器上, 产生载频为 64 MHz、调制指数为 5 的 FM 信号。试确定:

(a) 偏压 V_b 的幅度;

(b) $f_m = 100 \text{ kHz}$ 时, 调制信号的幅度 A_m 。

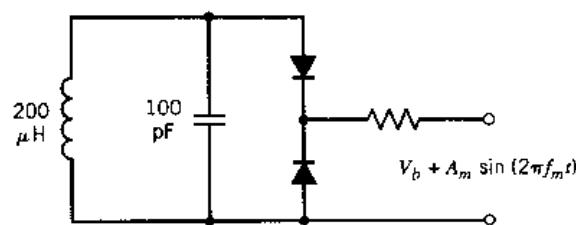


图 P2.38

2.39 将 FM 信号

$$s(t) = A_c \cos \left[2\pi f_c t + 2\pi k_f \int_0^t m(\tau) d\tau \right]$$

加在图 P2.39 由高通 RC 滤波器和包络检波器组成的系统上。假设：

- (a) 对于 $s(t)$ 所有主要的频率分量, 电阻 R 均小于电容 C 的电感值;
- (b) 包络检波器无滤波功能。假定对于所有 t , 有 $|k_f m(t)| < f_c$, 计算包络检波器的输出信号。

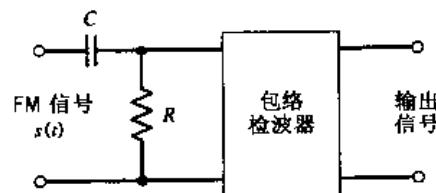


图 P2.39

2.40 考虑图 P2.40 中的频率解调方法, 使输入 FM 信号 $s(t)$ 通过延迟器, 从而在载频 f_c 处发生了 $\pi/2$ 弧度的相移。用输入 FM 信号减去延迟器的输出信号, 对产生的复合信号进行包络检波。该解调器用于微波 FM 信号的解调中。假设：

$$s(t) = A_c \cos[2\pi f_c t + \beta \sin(2\pi f_m t)]$$

当调制指数 β 小于 1, 且延迟器产生的时延 T 足够小, 使得下面的近似式成立时, 分析该解调器的工作原理:

$$\cos(2\pi f_m T) \approx 1$$

和

$$\sin(2\pi f_m T) \approx 2\pi f_m T$$

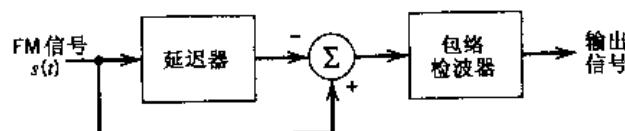


图 P2.40

2.41 图 2.41 为用于解调 FM 信号的过零检测器框图。它由限幅器、在输入的每个过零点产生窄脉冲的脉冲发生器以及提取调制信号的低通滤波器组成。

- (a) 证明输入 FM 信号的瞬时频率和时间间隔 $t - (T_1/2)$ 到 $t + (T_1/2)$ 内的过零点数

与 T_1 的商成正比。假设在该时间间隔内, 调制信号不变。

- (b) 将图 P2.24 的锯齿波作为调制信号, 分析该解调器的工作原理。

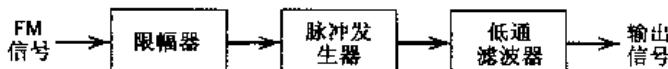


图 P2.41

2.42 假设 FM 系统接收到的信号中含有--部分具有正幅度 $a(t)$ 的残留 AM 信号, 即

$$s(t) = a(t) \cos[2\pi f_c t + \phi(t)]$$

其中, f_c 为载频。相位 $\phi(t)$ 与调制信号 $m(t)$ 的关系式如下:

$$\phi(t) = 2\pi k_f \int_0^t m(\tau) d\tau$$

其中, k_f 为常数。假设信号 $s(t)$ 被限制在带宽为 B_T , 中心频率为 f_c 的频带内, 其中, B_T 是无幅度调制时 FM 信号的传输带宽, 幅度调制的变化慢于 $\phi(t)$ 的变化。证明对于输入为 $s(t)$ 的理想鉴频器, 其输出与 $a(t)m(t)$ 成正比。提示: 用附录 2 中已调信号 $s(t)$ 的复数表示法。

2.43 (a) 将习题 2.42 中的已调信号 $s(t)$ 加在硬限幅器上, 硬限幅器输出 $z(t)$ 的定义如下:

$$\begin{aligned} z(t) &= \operatorname{sgn}[s(t)] \\ &= \begin{cases} +1, & s(t) > 0 \\ -1, & s(t) < 0 \end{cases} \end{aligned}$$

证明限幅器输出信号可由下面的傅里叶级数表示:

$$z(t) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \cos[2\pi f_c t(2n+1) + (2n+1)\phi(t)]$$

(b) 将限幅器的输出送入带通滤波器, 滤波器通带内幅度响应为 1, 带宽为 B_T , 中心频率为 f_c , B_T 是无幅度调制时 FM 信号的传输带宽。假设 f_c 远远大于 B_T , 证明滤波器输出等于

$$y(t) = \frac{4}{\pi} \cos[2\pi f_c t + \phi(t)]$$

将该输出与习题 2.42 中定义的原始已调信号 $s(t)$ 做比较, 说明该结果的实际作用。

2.44 (a) 考虑由调制信号 $m(t)$ 产生的载频为 f_c 的 FM 信号。假设 f_c 足够大, 可将 FM 信号看成是窄带信号。求该 FM 信号希尔伯特变换的近似表达式。

(b) 对于由特殊的正弦调制信号 $m(t) = A_m \cos(2\pi f_m t)$ 得到的 FM 信号, 求其希尔伯特变换的准确表达式。在这种情况下, (a) 中所得近似表达式的误差是多大?

2.45 单边带角度调制的定义如下:

$$s(t) = \exp[-\hat{\phi}(t)] \cos[2\pi f_c t + \phi(t)]$$

其中 $\hat{\phi}(t)$ 为相位函数 $\phi(t)$ 的希尔伯特变换, f_c 为载频。

(a) 证明已调信号 $s(t)$ 的频谱在 $-f_c < f < f_c$ 内没有频率分量, 且其频谱无穷延伸。

(b) 已知相位函数

$$\phi(t) = \beta \sin(2\pi f_m t)$$

其中, β 为调制指数, f_m 为调制频率, 推导已调信号 $s(t)$ 的相应表达式。

注意: 对于习题 2.44 和 2.45, 需要参考附录 2 中的希尔伯特变换。

CW 调制系统中的噪声

2.46 DSB-SC 已调信号通过噪声信道传输, 噪声的功率谱密度如图 P2.46 所示。消息带宽为 4 kHz, 载频为 200 kHz。假设已调信号的平均功率是 10 W, 计算接收机的输出信噪比。

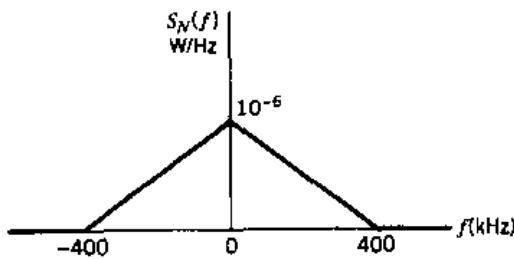


图 P2.46

2.47 计算下列情况中, 相干检测器输入窄带噪声的同相分量和正交分量的自相关函数和互相关函数。

- (a) DSB-SC 系统;
- (b) 发送下边带的 SSB 系统;
- (c) 发送上边带的 SSB 系统。

2.48 在使用相干检测的接收机里, 本地振荡器产生的正弦信号与载波 $\cos(2\pi f_c t)$ 之间存在相位误差 $\theta(t)$ 。假设 $\theta(t)$ 是均值为零, 方差为 σ_θ^2 的高斯过程的一个样本函数, 且在绝大部分时间里, $\theta(t)$ 的最大值均小于 1。求 DSB-SC 调制下, 接收机输出的均方误差。均方误差定义为接收机输出信号与接收机输出的消息信号分量之差的平方的期望值。

2.49 按照 2.11 节中对 DSB-SC 接收机的分析过程, 分析 SSB 接收机中的噪声, 并计算:

- (a) 输出信噪比;
- (b) 信道信噪比。

并证明 SSB 接收机的解调增益恰好与 DSB-SC 接收机的解调增益相同。注意, 与 DSB-SC 接收机不同, SSB 接收机前端的窄带滤波噪声的谱密度函数的中心频率相对于载频 f_c 偏移了 $W/2$, W 为消息带宽。

2.50 消息信号 $m(t)$ 采用 SSB 调制进行传输。 $m(t)$ 的功率谱密度为

$$S_M(f) = \begin{cases} \alpha \frac{|f|}{W}, & |f| \leq W \\ 0, & \text{其他} \end{cases}$$

其中, a 和 W 为常数。在接收机输入端, SSB 已调信号叠加了均值为零, 功率谱密度为 $N_0/2$ 的高斯白噪声。求出接收机输出信噪比的表达式。

2.51 考虑式(2.92)定义的包络检波器的输出信号, 其表达式如下:

$$y(t) = \{[A_c + A_c k_a m(t) + n_I(t)]^2 + n_Q^2(t)\}^{1/2}$$

(a) 若下式成立的概率不大于 δ_1 :

$$|n_Q(t)| > \epsilon A_c |1 + k_a m(t)|$$

其中 $\epsilon \ll 1$ 。求正交分量 $n_Q(t)$ 的影响可忽略的概率是多大?

(b) 调整关于消息信号 $m(t)$ 的 k_a 值, 使得下式成立的概率等于 δ_2 :

$$A_c [1 + k_a m(t)] + n_I(t) < 0$$

那么近似式

$$y(t) \approx A_c [1 + k_a m(t)] + n_I(t)$$

成立的概率是多少?

(c) 当 δ_1 和 δ_2 均小于 1 时, 说明(b)中所得结果的意义。

2.52 将幅度为 A_c , 频率为 f_c 的未调制载波与带限白噪声之和通过理想的包络检波器。若噪声功率谱密度为 $N_0/2$, 带宽为 $2W$, 中心在载频 f_c 处。计算当载噪比很高时, 输出信噪比是多大?

2.53 用 R 表示在固定时间内观察包络检波器输出所得的随机变量。如果随机变量 R 大于载波幅度 A_c 的概率为 0.5, 则包络检波器刚好工作在门限区。另一方面, 如果上述概率只有 0.01, 那么包络检波将不存在门限效应, 消息不会发生丢失。

(a) 假设检测器输入的窄带噪声是零均值高斯白噪声, 其功率谱密度为 $N_0/2$, 消息带宽为 W , 证明事件 $R \geq A_c$ 的概率为

$$P(R \geq A_c) = \exp(-\rho)$$

式中, ρ 为载噪比:

$$\rho = \frac{A_c^2}{4WN_0}$$

(b) 用此概率公式, 计算下面两种情况下的载噪比:

- (1) 包络检波器工作在门限区时;
- (2) 包络检波器正常工作时。

2.54 考虑一个 PM 系统的已调信号表达式为

$$s(t) = A_c \cos[2\pi f_c t + k_p m(t)]$$

其中, k_p 为常数, $m(t)$ 为消息信号。相位检测器输入端的加性噪声 $n(t)$ 为

$$n(t) = n_I(t) \cos(2\pi f_c t) - n_Q(t) \sin(2\pi f_c t)$$

假设检测器输入端的载噪比大于 1, 计算:

- (a) 输出信噪比;
- (b) 系统的解调增益。

在正弦调制下,将计算结果与 FM 系统相比较。

- 2.55 FDM 系统采用 SSB 调制将 12 个独立的语音信号结合在一起,然后通过频率调制发送此复合基带信号。每一个语音信号的平均功率为 P ,占用的频带为 $0.3 \sim 3.4$ kHz,系统为语音信号分配的带宽是 4 kHz。对于每一个语音信号,只发送下边带。用于第一级调制的副载波定义如下:

$$c_k(t) = A_k \cos(2\pi k f_0 t), \quad 0 \leq k \leq 11$$

接收信号由传输的 FM 信号和均值为零,功率谱密度为 $N_0/2$ 的高斯白噪声组成。

- (a) 画出鉴频器输出信号的功率谱密度草图,并标出信号和噪声分量。
 (b) 为使已调语音信号有相同的信噪比,找出副载波幅度 A_k 之间应满足的关系。

- 2.56 在 2.13 节关于 FM 门限效应的讨论中,给出了用窄带噪声 $n(t)$ 的包络 $r(t)$ 和相位 $\psi(t)$ 表示的正向咔嗒声和负向咔嗒声出现的条件。试给出用 $n(t)$ 的同相分量 $n_I(t)$ 和正交分量 $n_Q(t)$ 表示的两种咔嗒声出现的条件。

- 2.57 采用图 2.50(a) 所示的预加重滤波器,并将语音信号作为调制信号,则 FM 发射机产生的信号在较低音频时主要为频率调制,在较高音频时主要为相位调制。解释这种现象发生的原因。

- 2.58 假设 FM 系统中预加重和去加重滤波器的转移函数为:

$$H_{pe}(f) = k \left(1 + \frac{if}{f_0} \right)$$

和

$$H_{de}(f) = \frac{1}{k} \left(\frac{1}{1 + if/f_0} \right)$$

比例因子 k 的选择应使加重后消息信号的平均功率与原始消息信号 $m(t)$ 的平均功率相等。

(a) 当消息信号 $m(t)$ 的功率密度谱为

$$S_M(f) = \begin{cases} \frac{S_0}{1 + (f/f_0)^2}, & -W \leq f \leq W \\ 0, & \text{其他} \end{cases}$$

时,计算加重后消息信号与原始消息信号的平均功率相等时的 k 值。

(b) 由这对预加重和去加重滤波器得到的相应的改善因子 I 是多少? 将该值与例 2.6 中的值相比较。改善因子 I 的定义见式(2.160)。

- 2.59 PM 系统使用一对预加重和去加重滤波器,它们的转移函数为

$$H_{pe}(f) = 1 + \frac{if}{f_0}$$

和

$$H_{de}(f) = \frac{1}{1 + (if/f_0)}$$

证明:使用了这对滤波器带来的输出信噪比的改善因子为

$$I = \frac{W/f_0}{\tan^{-1}(W/f_0)}$$

其中, W 为消息带宽。计算 $W = 15 \text{ kHz}$, $f_0 = 2.1 \text{ kHz}$ 时的改善因子, 并将结果与相应 FM 系统的改善因子进行比较。

计算机实验

2.60 在本实验中研究图 P2.5 所示包络检波器的作用, 系统说明如下:

电源内阻 $R_s = 75 \Omega$;

负载电阻 $R_L = 10 \text{ k}\Omega$;

电容 $C = 0.01 \mu\text{F}$;

二极管正偏时内阻为 25Ω , 反偏时电阻为无穷大。

若输入正弦 AM 信号的调制百分比为 50%, 调制频率为 1 kHz , 载频为 20 kHz , 计算包络检波器的输出波形。

2.61 本实验继续研究 2.14 节中的锁相环:

(a) 计算锁相环中 VCO 瞬时频率的变化量, 锁相环参数如下:

$$\text{环路增益参数 } K_0 = \frac{50}{2\pi} \text{ Hz}$$

$$\text{自由振荡频率 } f_n = \frac{1}{2\pi} \text{ Hz}$$

$$\text{阻尼因子 } \zeta = 0.707$$

分别针对以下频率步进值进行计算: $\Delta f = 0.125, 0.5, 7/12, 2/3 \text{ Hz}$

(b) 根据 2.14 节实验 2 中锁相环的参数值, 说明相对频偏 $\Delta f f_n / f_n^2$ 的变化是怎样影响锁相环相位误差的峰值的。

第3章 脉冲调制

本章是模拟通信到数字通信的过渡,主要讨论下列问题:

- 抽样,是所有形式的脉冲调制的基础。
- 脉冲幅度调制,是脉冲调制中最简单的形式。
- 量化,与抽样结合,可将模拟信号以幅度和时间均离散的形式表示。
- 脉冲编码调制,这是用数字方式传输模拟消息信号的标准方法。
- 时分复用,通过脉冲调制使多用户分时共享同一信道。
- 数字信号复接器,将很多低速比特流合成一个高速比特流。
- 其他数字脉冲调制形式,如增量调制和差分脉冲编码调制。
- 线性预测,是以低比特率对模拟消息信号进行编码的基础(相对于差分脉冲编码调制方法)。
- 自适应差分脉冲编码调制和自适应增量调制。
- MPEG-1/音频编码标准,是透明的、感觉无损的压缩系统。

3.1 引言

在第2章对连续波调制的研究中,正弦载波的某些参数随消息信号连续变化。相应地,在本章研究的脉冲调制中,脉冲串的某些参数随消息信号发生变化。脉冲调制可分为两种:模拟脉冲调制和数字脉冲调制。在模拟脉冲调制中,将周期性脉冲串作为载波,每一脉冲的某些特征量(如幅度、宽度或位置)随消息信号的相应抽样值而连续变化。因此,在模拟脉冲调制中,信息本质上是在离散时间上以模拟形式进行传送的。而在数字脉冲调制中,消息信号以时间和幅度均离散的形式表示,因此可用一系列编码脉冲以数字形式进行信号传输,这种信号传输形式没有相对应的连续波。

以编码脉冲传输模拟信号的方法是数字通信的一个基本组成部分。因此,可将本章视为从模拟通信到数字通信的过渡。对于所有脉冲调制系统,无论是模拟的还是数字的,抽样都是一个基本过程,因此本章从抽样开始进行介绍。

3.2 抽样过程

抽样过程经常在时域内进行描述。它是数字信号处理和数字通信中的基本操作。通过抽样,模拟信号变成了相应的时间间隔相等的序列。显然,要使抽样过程具有实际意义,选择正确的抽样速率是非常必要的。只有这样,才能保证能从样值序列中恢复出原始模拟信号,这是抽样定理的实质。接下来讨论抽样定理。

考虑能量有限的任意信号 $g(t)$, $g(t)$ 在所有时刻上均有值。信号 $g(t)$ 的一部分如图 3.1(a) 所示。假设对信号 $g(t)$ 以恒定速率, 每隔 T_s 秒进行一次瞬时抽样, 则将得到间隔为 T_s 秒的无穷抽样序列。抽样序列用 $\{g(nT_s)\}$ 表示, n 为所有可能的整数值。 T_s 称为抽样周期, 其倒数 $f_s = 1/T_s$ 称为抽样速率。这种理想抽样形式称为瞬时抽样。

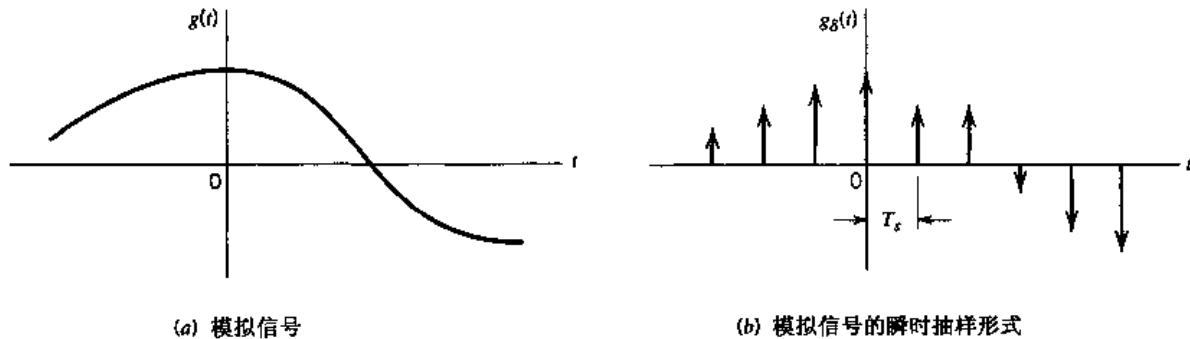


图 3.1 抽样过程

令 $g_s(t)$ 表示用序列 $\{g(nT_s)\}$ 对间隔为 T_s 秒的周期冲激函数序列加权所得信号, $g_s(t)$ 如下式所示 [见图 3.1(b)]:

$$g_s(t) = \sum_{n=-\infty}^{\infty} g(nT_s) \delta(t - nT_s) \quad (3.1)$$

$g_s(t)$ 称为理想抽样信号。 $\delta(t - nT_s)$ 表示在时刻 $t = nT_s$ 处的冲激函数。由冲激函数的定义可知, 理想冲激函数的面积为 1(见附录 2)。因此式(3.1)中的乘积因子 $g(nT_s)$ 可视为分配给冲激函数 $\delta(t - nT_s)$ 的“重量”。加权后的冲激函数可用宽度为 Δt , 幅度为 $g(nT_s)/\Delta t$ 的矩形脉冲近似地表示。 Δt 越小, 近似效果越好。

由傅里叶变换表, 可得(见表 A6.3 最后一项):

$$g_s(t) \rightleftharpoons f_s \sum_{m=-\infty}^{\infty} G(f - mf_s) \quad (3.2)$$

其中, $G(f)$ 为原始信号 $g(t)$ 的傅里叶变换, f_s 为抽样速率。式(3.2)表明, 对有限能量的连续时间信号进行均匀抽样, 将得到周期等于抽样速率的周期频谱。

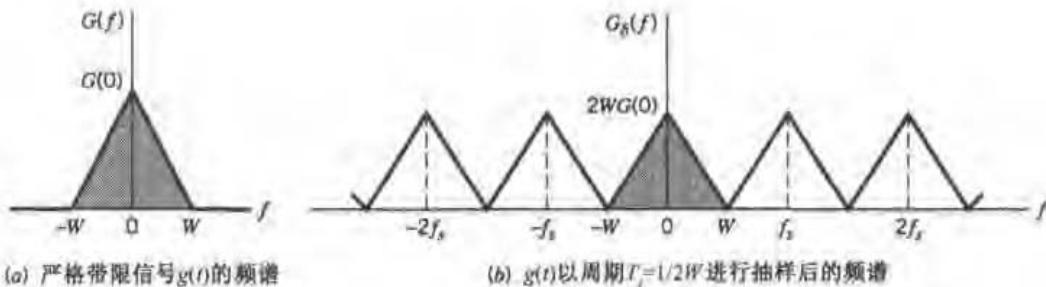
对式(3.1)两边同时取傅里叶变换, 可得理想抽样信号 $g_s(t)$ 的另一个有用表达式。需注意的是, 冲激函数 $\delta(t - nT_s)$ 的傅里叶变换等于 $\exp(-j2\pi nf_s T_s)$ 。令 $G_s(f)$ 表示 $g_s(t)$ 的傅里叶变换, 可得

$$G_s(f) = \sum_{n=-\infty}^{\infty} g(nT_s) \exp(-j2\pi nf_s T_s) \quad (3.3)$$

该关系式称为离散时间傅里叶变换。可将它作为周期频率函数 $G_s(f)$ 的复数傅里叶级数表达式, 由抽样序列 $\{g(nT_s)\}$ 来决定该级数的系数。

这里推导出的关系式, 适用于任意能量有限、持续时间无限的连续时间信号 $g(t)$ 。如果 $g(t)$ 是严格带限的, 则不存在高于 W 赫的频率分量。即信号 $g(t)$ 的傅里叶变换 $G(f)$ 在 $|f| \geq W$ 处为零, 如图 3.2 所示。该图为示意图。同时假定抽样周期为 $T_s = 1/2W$, 则相应抽样信号 $g_s(t)$ 的频谱 $G_s(f)$ 如图 3.2(b) 所示。将 $T_s = 1/2W$ 代入式(3.3), 有

$$G_s(f) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \exp\left(-j\frac{\pi n f}{W}\right) \quad (3.4)$$

图 3.2 $g(t)$ 抽样前后的频谱

由式(3.2)容易看出, $g_s(t)$ 的傅里叶变换也可表示为

$$G_s(f) = f_s G(f) + f_s \sum_{m=-\infty}^{\infty} G(f - mf_s) \quad (3.5)$$

从而,在下列两个条件下:

1. 当 $|f| \geq W$ 时, $G(f) = 0$
2. $f_s = 2W$

由式(3.5),有

$$G(f) = \frac{1}{2W} G_s(f), \quad -W < f < W \quad (3.6)$$

将式(3.4)代入式(3.6),可得

$$G(f) = \frac{1}{2W} \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \exp\left(-j\frac{\pi n f}{W}\right), \quad -W < f < W \quad (3.7)$$

因此,如果对于所有 n 值,信号 $g(t)$ 的抽样值 $g(n/2W)$ 均为已知,则信号的傅里叶变换 $G(f)$ 由式(3.7)所示的离散时间傅里叶变换惟一确定。由于 $g(t)$ 与 $G(f)$ 为一对傅里叶变换,因此当 $-\infty < n < \infty$ 时,信号 $g(t)$ 本身仅由抽样值 $g(n/2W)$ 决定。也就是说,序列 $\{g(n/2W)\}$ 包含了 $g(t)$ 中的全部信息。

接下来研究如何由抽样值序列 $\{g(n/2W)\}$ 重构信号 $g(t)$ 。将式(3.7)代入用 $G(f)$ 定义的傅里叶逆变换式 $g(t)$,有

$$\begin{aligned} g(t) &= \int_{-\infty}^{\infty} G(f) \exp(j2\pi f t) df \\ &= \int_{-W}^{W} \frac{1}{2W} \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \exp\left(-j\frac{\pi n f}{W}\right) \exp(j2\pi f t) df \end{aligned}$$

交换求和与积分的顺序

$$g(t) = \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \frac{1}{2W} \int_{-W}^{W} \exp\left[j2\pi f\left(t - \frac{n}{2W}\right)\right] df \quad (3.8)$$

容易计算式(3.8)中的积分式,从而得到最后的结果为

$$\begin{aligned} g(t) &= \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \frac{\sin(2\pi Wt - n\pi)}{(2\pi Wt - n\pi)} \\ &= \sum_{n=-\infty}^{\infty} g\left(\frac{n}{2W}\right) \text{sinc}(2Wt - n), \quad -\infty < t < \infty \end{aligned} \quad (3.9)$$

式(3.9)提供了由抽样序列 $\{g(n/2W)\}$ 重构原始信号 $g(t)$ 的插值公式,函数 $\text{sinc}(2Wt)$ 起到了插值函数的作用。每个抽样值均与一个延迟插值函数相乘,所有波形之和即为 $g(t)$ 。

对于能量有限的严格带限信号而言,抽样定理可表述为等价的两部分,分别应用于脉冲调制系统的发射机和接收机:

1. 能量有限的带限信号,不存在高于 W 赫的频率分量,完全可由时间相隔为 $1/(2W)$ 秒的瞬时信号值表示。
2. 能量有限的带限信号,不存在高于 W 赫的频率分量,完全可从抽样速率为每秒 $2W$ 的样值序列中恢复出来。

对于信号带宽为 W 赫的信号,每秒 $2W$ 抽样点的抽样速率称为奈奎斯特速率,其倒数 $1/(2W)$ (以秒计)称为奈奎斯特间隔。

此处对抽样定理的推导,是基于信号 $g(t)$ 严格带限这一假设的。但在实际情况中,承载有信息的信号并不是严格带限的,从而产生了某种程度的欠抽样。因此,在抽样过程中会发生混叠现象。混叠是指信号频谱中的高频分量在其抽样频谱的低频处发生叠加的现象,如图3.3所示。图3.3(b)中的实线是混叠的频谱,它是图3.3(a)所示频谱对应的“欠抽样”形式。

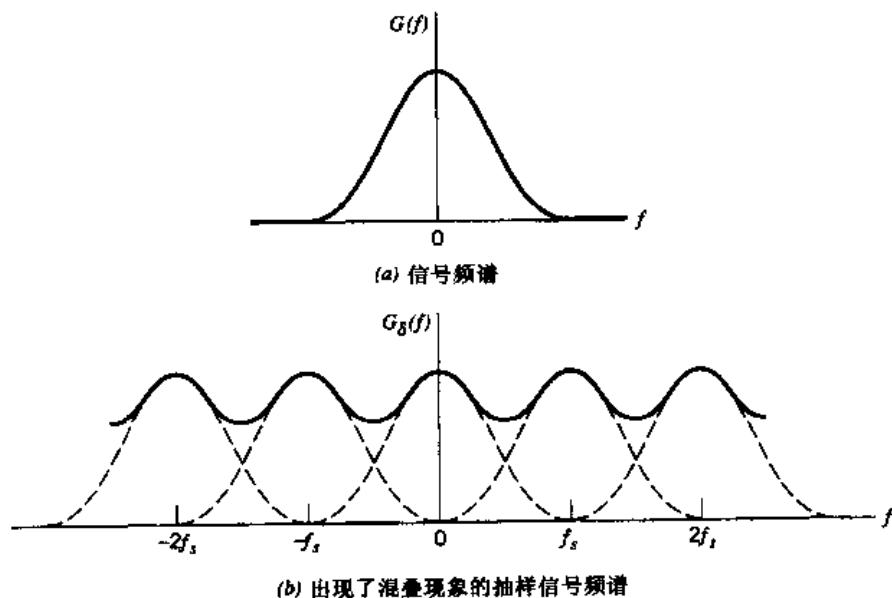


图 3.3 混叠现象

在实际情况中,为了抑制混叠的影响,可采用下述两个校正措施:

1. 在抽样之前,让信号经过低通抗混叠滤波器,以抑制那些对所传送信息来讲并不重要的高频分量。
2. 滤波后的信号抽样速率稍高于奈奎斯特速率。

抽样速率高于奈奎斯特速率,也有利于简化重构滤波器的设计,重构滤波器的作用是从抽样值中恢复出原始信号。经过抗混叠(低通)滤波器的信号频谱如图 3.4(a)所示。抽样速率高于奈奎斯特速率时相应的瞬时抽样频谱如图 3.4(b)所示。由图 3.4(b),易得如图 3.4(c)的重构滤波器设计方案如下:

- 重构滤波器是低通的,通带从 $-W$ 到 W , W 的值由抗混叠滤波器决定。
- 该滤波器存在从 W 到 $f_s - W$ 的过渡带(对于正频率),其中 f_s 为抽样速率。

重构滤波器有确定过渡带的事实,说明它是物理可实现的。

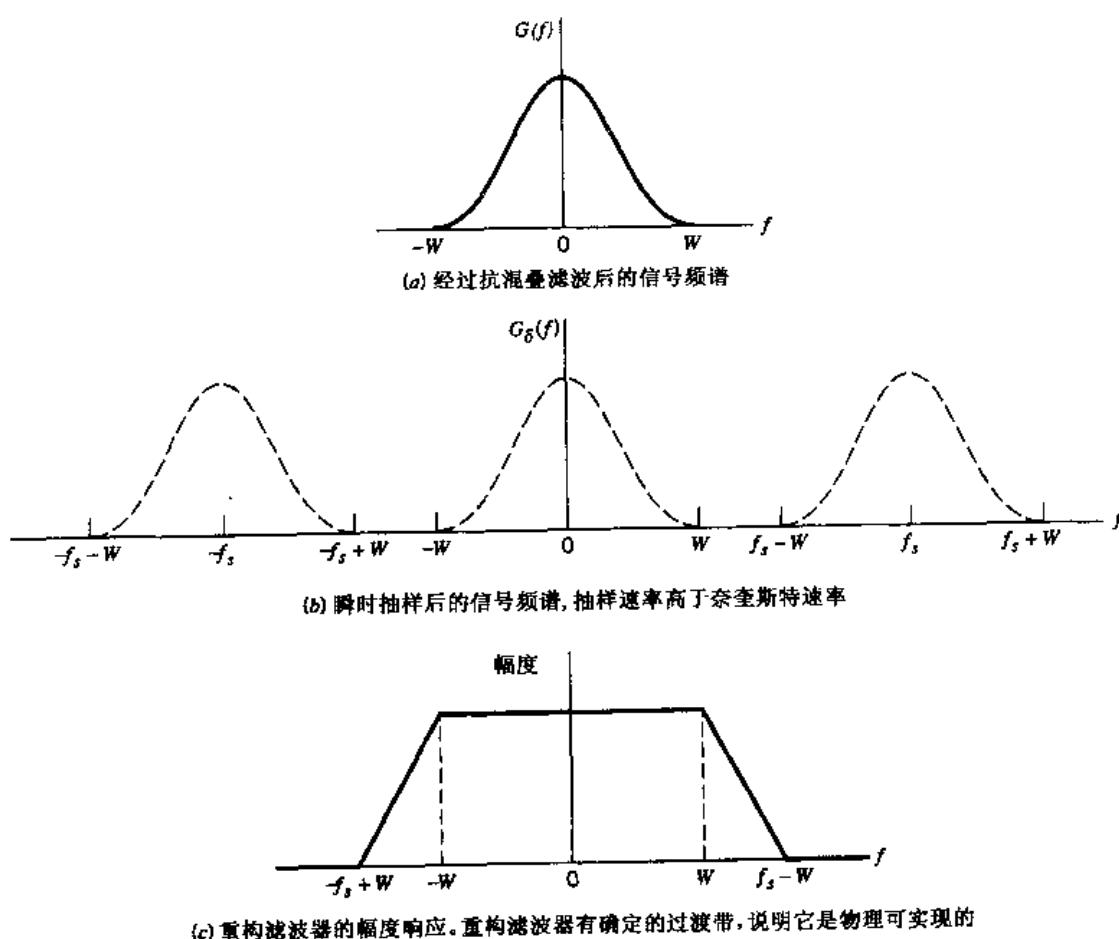


图 3.4 信号频谱和重构滤波器幅度响应

3.3 脉冲幅度调制

脉冲幅度调制(PAM)是最简单、最基本的模拟脉冲调制形式。在 PAM 中,等间隔脉冲的幅度随连续消息信号的相应抽样值呈正比例变化。脉冲可以是矩形脉冲或其他合适的脉冲形式。这里定义的 PAM 与自然抽样有些类似。在自然抽样中,消息信号与周期性的矩形脉冲串相乘。二者的区别在于,在自然抽样中,每个已调矩形脉冲的顶部形状随消息信号而变化,但在 PAM 中,矩形脉冲的顶部是平坦的。习题 3.2 进一步探讨了自然抽样。

图 3.5 为 PAM 信号的波形。该图中的虚线为消息信号 $m(t)$ 的波形，由实线表示的幅度调制矩形脉冲序列即为相应的 PAM 信号 $s(t)$ 。产生 PAM 信号的过程如下：

1. 每隔 T_s 秒对消息信号 $m(t)$ 进行瞬时抽样。其中，抽样速率 $f_s = 1/T_s$ 是根据抽样定理选取的。
2. 延长每个抽样值的持续时间，使之等于常量 T 。

在数字电路技术中，将这两步合称为“抽样和保持”。特意延长每个样值持续时间的一个重要原因是为了避免占用过多的信道带宽，因为带宽与脉冲宽度成反比。我们通过下面的分析来说明，应当认真选取抽样持续时间 T 。

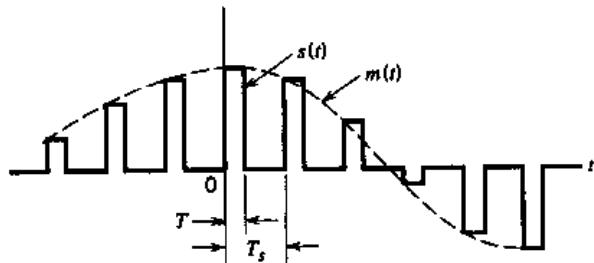


图 3.5 表示模拟信号的平顶抽样

令 $s(t)$ 表示由图 3.5 描述的方法产生的平顶脉冲序列。可将 PAM 信号表示为

$$s(t) = \sum_{n=-\infty}^{\infty} m(nT_s) h(t - nT_s) \quad (3.10)$$

其中， T_s 为抽样周期， $m(nT_s)$ 为在时刻 $t = nT_s$ 得到的 $m(t)$ 的抽样值。 $h(t)$ 是幅值为 1、宽度为 T 的标准矩形脉冲，其定义如下（如图 3.6(a)）：

$$h(t) = \begin{cases} 1, & 0 < t < T \\ \frac{1}{2}, & t = 0, t = T \\ 0, & \text{其他} \end{cases} \quad (3.11)$$

由定义， $m(t)$ 的瞬时抽样值为

$$m_\delta(t) = \sum_{n=-\infty}^{\infty} m(nT_s) \delta(t - nT_s) \quad (3.12)$$

其中， $\delta(t - nT_s)$ 为时移冲激函数。对 $m_\delta(t)$ 和脉冲 $h(t)$ 做卷积，可得

$$\begin{aligned} m_\delta(t) \star h(t) &= \int_{-\infty}^{\infty} m_\delta(\tau) h(t - \tau) d\tau \\ &= \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} m(nT_s) \delta(\tau - nT_s) h(t - \tau) d\tau \\ &= \sum_{n=-\infty}^{\infty} m(nT_s) \int_{-\infty}^{\infty} \delta(\tau - nT_s) h(t - \tau) d\tau \end{aligned} \quad (3.13)$$

由冲激函数的移位特性（见附录 2）可得

$$m_\delta(t) \star h(t) = \sum_{n=-\infty}^{\infty} m(nT_s) h(t - nT_s) \quad (3.14)$$

由式(3.10)和式(3.14)可知，PAM 信号 $s(t)$ 是 $m(t)$ 的瞬时抽样值 $m_\delta(t)$ 与脉冲 $h(t)$ 的卷积，

也就是说, $s(t)$ 与该卷积在数学形式上是等价的, 即

$$s(t) = m_s(t) \star h(t) \quad (3.15)$$

对式(3.15)两边取傅里叶变换, 并注意到两个函数在时域中的卷积对应于频域中它们的傅里叶变换的乘积, 即有

$$S(f) = M_s(f)H(f) \quad (3.16)$$

其中, $S(f) = F[s(t)]$, $M_s(f) = F[m_s(t)]$, $H(f) = F[h(t)]$ 。利用式(3.2), 可得傅里叶变换 $M_s(f)$ 与原始消息信号 $m(t)$ 的傅里叶变换 $M(f)$ 有如下关系:

$$M_s(f) = f_s \sum_{k=-\infty}^{\infty} M(f - kf_s) \quad (3.17)$$

其中, f_s 为抽样速率。将式(3.17)代入式(3.16)可得

$$S(f) = f_s \sum_{k=-\infty}^{\infty} M(f - kf_s)H(f) \quad (3.18)$$

给定 PAM 信号 $s(t)$, 其傅里叶变换 $S(f)$ 由式(3.18)定义, 怎样恢复出原始消息信号 $m(t)$ 呢?

重构的第一步, 先将 $s(t)$ 通过频率响应如图 3.4(c) 所示的低通滤波器。这里假定消息被限制在带宽 W 以内, 并且抽样速率 f_s 大于奈奎斯特速率 $2W$ 。由式(3.18)可得滤波器输出的频谱等于 $M(f)H(f)$ 。此滤波器的输出与原始消息信号 $m(t)$ 通过另一个频率响应为 $H(f)$ 的低通滤波器的输出是相等的。

由式(3.11)得出矩形脉冲 $h(t)$ 的傅里叶变换为

$$H(f) = T \operatorname{sinc}(fT) \exp(-j\pi fT) \quad (3.19)$$

相应的曲线如图 3.6(b) 所示。由此可见, 通过平顶抽样产生 PAM 信号, 会产生幅度失真和 $T/2$ 的时延。这与由电视中有限大的扫描孔径造成的传输频率变化十分类似, 因此, 采用 PAM 模拟载频信号而造成的失真就称为孔径效应。

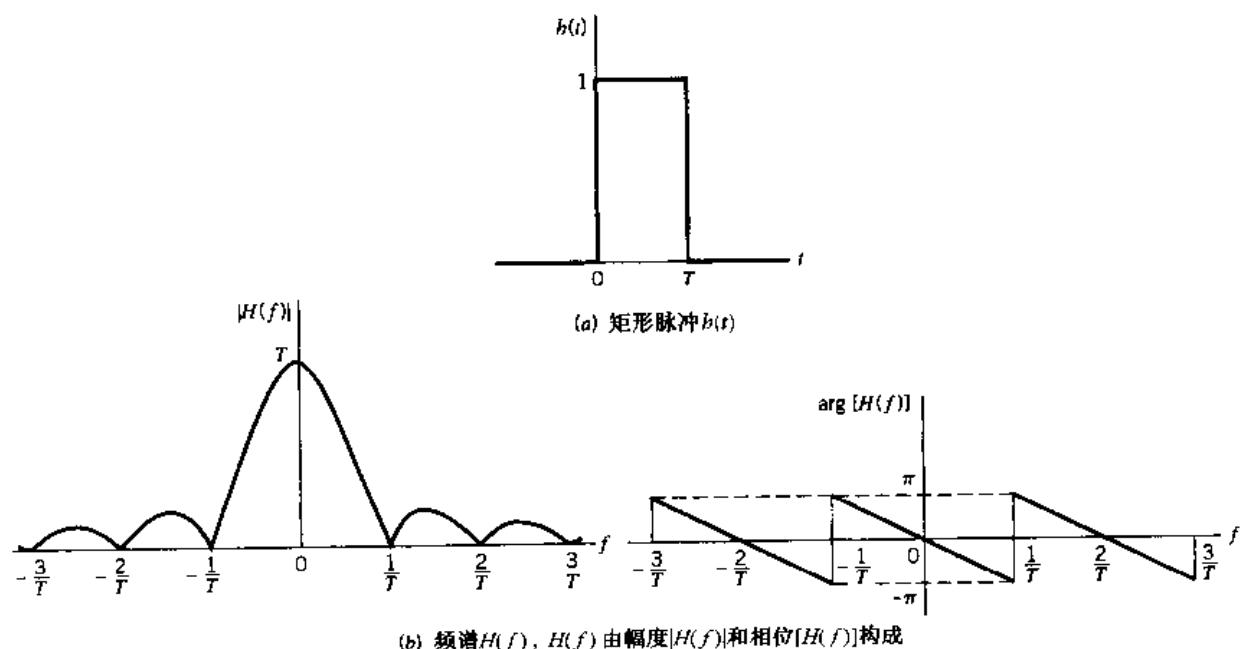


图 3.6 矩形脉冲及其频谱

这种失真可通过均衡器和低通重构滤波器的级联得到校正,如图 3.7 所示。均衡器通过频率增加从而弥补了孔径效应,因此起到了减少重构滤波器带内损失的作用。理想情况下,均衡器的幅度响应如下:

$$\frac{1}{|H(f)|} = \frac{1}{T \operatorname{sinc}(fT)} = \frac{\pi f}{\sin(\pi fT)} \quad (3.20)$$

在实际情况中,需要采用均衡的情况很少。当占空因数 $T/T_s \leq 0.1$ 时,幅度失真低于 0.5%。这种情况下,就可以省去均衡的步骤。

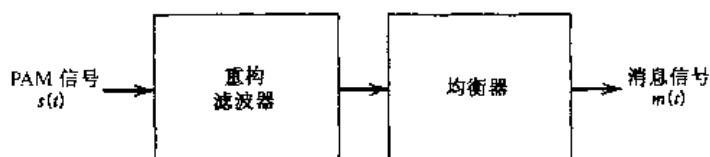


图 3.7 从 PAM 信号 $s(t)$ 中恢复消息信号 $m(t)$ 的系统框图

由于传输脉冲很窄,所以 PAM 信号的传输对信道幅度响应和相位响应有很严格的要求。同时,PAM 系统的噪声性能永远好于基带信号传输。因此,对于长距离的信号传输,PAM 只是作为时分复用中消息处理的一种手段而存在。接下来,我们将讨论其他形式的脉冲调制。时分复用将在 3.9 节中进行探讨。

3.4 其他脉冲调制形式

在脉冲调制系统中,可通过增加脉冲所消耗的带宽来提高系统抗噪声性能。将消息信号的抽样值用脉冲的其他特征量(不是幅度)表示,就可以达到上述目的。

- 脉冲持续时间调制(PDM),也称为脉宽调制,载波脉冲的宽度随消息信号的抽样值而变化。
- 脉冲相位调制(PPM),即脉冲对未调制时的相对位置,随消息信号的变化而变化。

在正弦调制信号下这两种形式的脉冲调制如图 3.8 所示。

在 PDM 中,持续时间较长的脉冲消耗的功率很大,但并没有携带更多的信息。如果将这部分无用功率从 PDM 中提取出来,只保留时间瞬间,那么,就得到了 PPM。与 PDM 相比,PPM 是更加有效的脉冲调制方式。

在 PPM 系统中,已调脉冲的相对位置包含着所传递的信息。因此,附加噪声的存在会使已调脉冲的出现时间偏离正确值,从而影响系统的性能。如果能够迅速地产生脉冲,从而使受噪声干扰影响的脉冲持续时间非常短,就可以消除噪声的影响。实际上,如果接收到的脉冲是理想矩形脉冲,那么加性噪声对脉冲的位置就不会产生任何影响,因为噪声只会在垂直方向带来干扰。然而,接收理想矩形脉冲需要无穷大的信道带宽,这当然是不切实际的。实际上,对于有限的信道带宽,接收到的脉冲存在有限的上升时间。因此,与预料的一样,PPM 接收机的性能会受到噪声的影响。

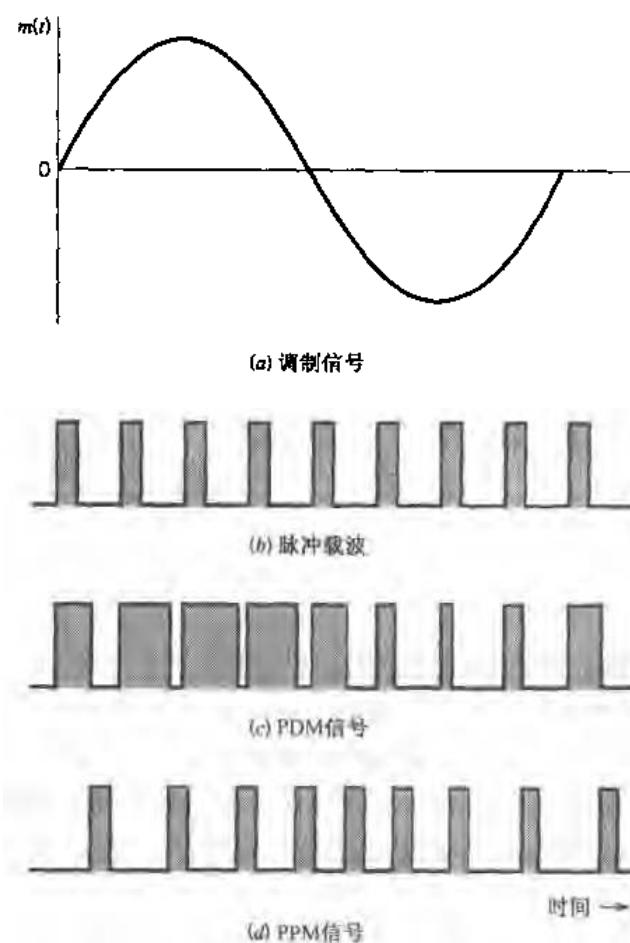


图 3.8 正弦调制信号下,两种不同形式的脉冲~时间调制

与连续波调制系统一样,PPM 系统的噪声性能可用输出信噪比进行描述。基带传输中由 PPM 带来的抗噪声性能的改善程度,可用 PPM 系统的输出信噪比除以信道信噪比即解调增益来表示。解调增益见 2.10 节。假设信道噪声的平均功率小于峰值脉冲功率,则 PPM 系统的解调增益与传输带宽 B_T 关于消息带宽 W 归一化结果的平方成正比。但当输入信噪比低于临界值时,在接收机输出端会损失有用的消息信号。即 PPM 系统会受到自身门限效应的影响。

3.5 带宽 – 噪声权衡

就噪声性能而言,PPM 系统是最佳的模拟脉冲调制形式。对 PPM 系统的噪声分析,揭示了 PPM 系统和 FM 系统具有相似的噪声性能,总结如下:^[1]

1. 两个系统的解调增益均与传输带宽关于消息带宽的归一化值的平方成正比。
2. 随着信噪比的减小,两个系统均出现了门限效应。

第一点的实际含义是,可通过增加传输带宽来提高噪声性能。在连续波调制和模拟脉冲调制系统中,达到此目的最有效的方法就是使用平方律。在讨论第一点时会产生这样一个问题:能产生比平方律更好的方法吗?答案明显是肯定的,数字脉冲调制就能够做到这一点。数字脉

冲调制是与连续波调制截然不同的一种调制方法。

特别地,在数字脉冲调制的一种基本形式,即脉冲编码调制(PCM)^[2]中,用时间和幅度均离散的形式来表示消息信号。这种信号表示方法将消息信号作为二进制脉冲编码序列进行传输。给定这样的序列,只要使所传输的二进制 PCM 信号的平均功率与噪声平均功率相比足够大,就可以将信道噪声对接收机输出造成的影响降至可忽略的程度。

产生二进制 PCM 信号波形要经过两个基本的过程:抽样和量化。抽样过程将消息信号表示成在时间上离散的形式。需注意的是,抽样必须遵照 3.2 节所述的抽样定理来进行。量化过程将消息信号表示成在幅度上离散的形式。下一节我们将具体讨论量化的问题。将抽样与量化相结合,才能将消息信号以编码的形式进行传输。这种编码形式可使带宽与噪声之间遵循一种指数律的交换,这个问题也将在下一节得到证明。

3.6 量化过程^[3]

连续信号(如语音)的幅度是连续的。因此,其抽样值也是连续的。换言之,在有限的信号幅值范围内,可找出无穷多个幅值电平。事实上,并不需要精确地传送抽样幅值。作为最终接收者的人类的任何感官(耳或眼),都只能察觉出有限的强度差异。这意味着可将原始的连续信号近似地表示为离散幅度组成的信号,该信号在可能的幅值集合中具有最小的误差。存在有限个离散幅值电平是脉冲编码调制的基本条件。显然,如果选择的离散幅值间隔足够小,那么,近似信号与原始连续信号几乎没有差别的。

幅值量化是将消息信号 $m(t)$ 在时刻 $t = nT_s$ 的抽样幅值 $m(nT_s)$ 转化为离散幅值 $v(nT_s)$ 的过程, $v(nT_s)$ 从可能幅值的有限集合中进行选取。假设量化过程是无记忆和瞬时的,即时刻 $t = nT_s$ 的量化值与该时刻之前和之后的消息信号抽样值均无关。这种最简单的等级量化形式虽然不是最佳的,但却应用广泛。

在讨论无记忆量化器时,我们去掉时间下标以简化符号。用符号 m 代替 $m(nT_s)$,如图 3.9(a)中的量化器框图所示。在图 3.9(b)中,如果信号幅值位于分隔区间内:

$$\mathcal{J}_k: \{m_k < m \leq m_{k+1}\}, \quad k = 1, 2, \dots, L \quad (3.21)$$

那么,信号幅值 m 由下标 k 确定。其中 L 是量化器的总级数。量化器输入端的离散幅值 m_k ($k = 1, 2, \dots, L$) 称为判决电平或判决门限。在量化器的输出端,下标 k 被转化为代表了幅度区间 \mathcal{J}_k 的幅值 v_k , 离散幅值 v_k ($k = 1, 2, \dots, L$) 称为量化级或重构电平。相邻量化级之间的间隔称为量阶或步长。因此,如果输入信号样值 m 在幅度区间 \mathcal{J}_k , 那么, 量化器的输出 v 就等于 v_k 。如图 3.9(a), 映射

$$v = g(m) \quad (3.22)$$

是量化器的特性,由定义可知这是一个阶梯函数。

量化器可以是均匀的,也可以是非均匀的。在均匀量化器中,量化级是均匀分布的。反之,量化级就是非均匀分布的。本节中只考虑均匀量化器。非均匀量化器将在 3.7 节进行讨论。量化器按特性也可分为中平型和中升型。图 3.10(a)为中平型均匀量化器的输入 - 输出特性,这种量化器被称为中平型,是因为原点位于阶梯图阶梯平台的中点。图 3.10(b)为中升

型均匀量化器相应的输入 - 输出特性, 图中原点位于阶梯图上升部分的中点。图 3.10 中, 中平型和中升型均匀量化器均关于原点对称。

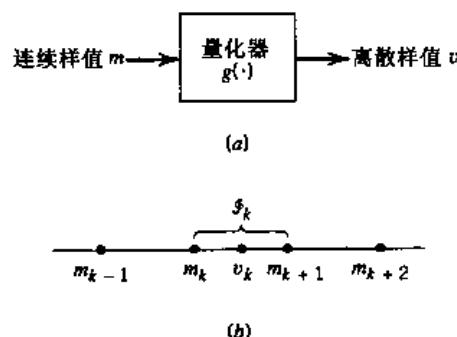


图 3.9 无记忆量化器的描述

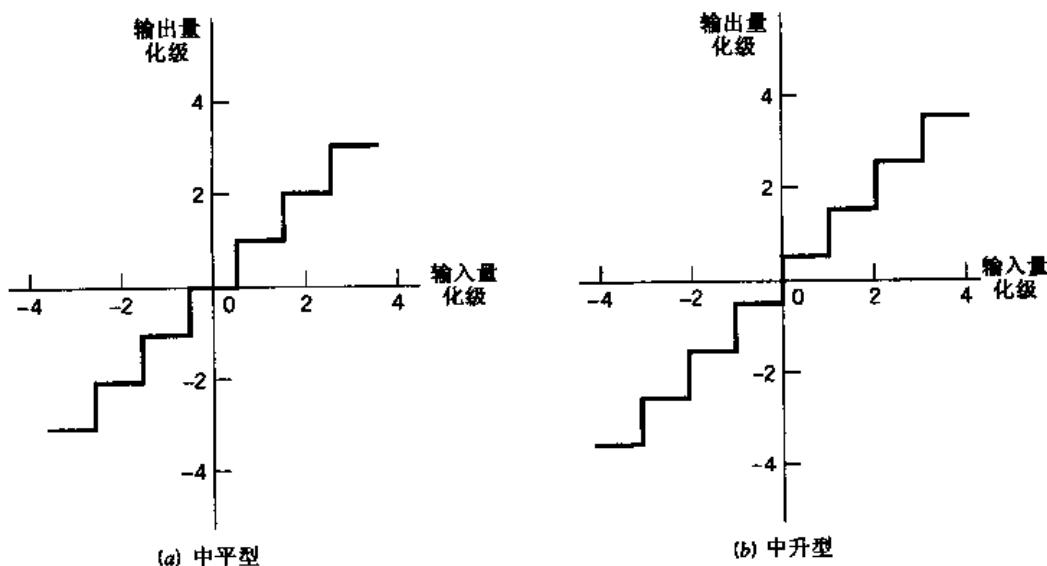


图 3.10 两种类型的量化器

3.6.1 量化噪声

量化的应用使输入信号 m 与输出信号 v 之间产生了差值, 即引入了误差。这个误差称为量化噪声。图 3.11 为量化噪声随时间变化的典型曲线, 假设采用中平型均匀量化器。

以零均值随机变量 M 的抽样值作为量化器输入 m 。(如果输入均值不为零, 通常在输入端减去均值, 量化后再把均值加上。)量化器 $g(\cdot)$ 将幅值连续的输入随机变量 M 映射为离散随机变量 V , 它们各自的抽样值 m 和 v 之间的关系式由式(3.22)给出。用抽样值为 q 的随机变量 Q 代表量化误差, 则有

$$q = m - v \quad (3.23)$$

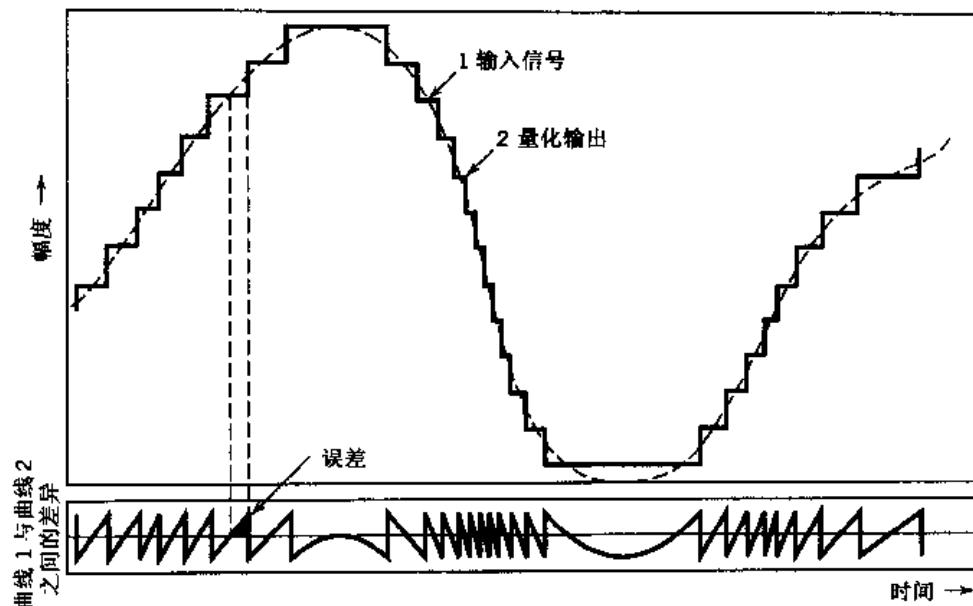


图 3.11 量化过程(经 AT&T 允许改编自 Bennett, 1948)

或相应地有

$$Q = M - V \quad (3.24)$$

若输入 M 是零均值的;且量化器如图 3.10 所示是对称的,那么量化器的输出 V 与量化误差 Q 也将是零均值的。因此,在用输出信噪比或信号 - 量化噪声比描述量化器的统计特性时,只需求出量化误差 Q 的均方值即可。

下面考虑连续幅值范围在 $(-m_{\max}, m_{\max})$ 内的输入 m 。假设采用图 3.10(b) 所示的中升型均匀量化器,则量化步长为

$$\Delta = \frac{2m_{\max}}{L} \quad (3.25)$$

式中, L 为量化级总数。对于均匀量化器,量化误差 Q 抽样值的取值范围为 $-\Delta/2 \leq q \leq \Delta/2$ 。如果步长 Δ 足够小(即量化级数 L 足够大),则可假设置量误差 Q 是服从均匀分布的随机变量,而量化噪声对量化器输入的干扰类似于热噪声的影响。因此可将量化误差 Q 的概率密度函数表示为如下形式:

$$f_Q(q) = \begin{cases} \frac{1}{\Delta}, & -\frac{\Delta}{2} < q \leq \frac{\Delta}{2} \\ 0, & \text{其他} \end{cases} \quad (3.26)$$

若要此式成立,必须保证输入信号没有超出量化器的范围。由于量化误差的均值为零,所以其方差 σ_Q^2 与均方值相等:

$$\sigma_Q^2 = E[Q^2] = \int_{-\Delta/2}^{\Delta/2} q^2 f_Q(q) dq \quad (3.27)$$

将式(3.26)代入式(3.27),有

$$\begin{aligned}\sigma_Q^2 &= \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} q^2 dq \\ &= \frac{\Delta^2}{12}\end{aligned}\quad (3.28)$$

通常情况下,用 L 进制数 k 来表示量化器的第 k 个量化级, k 以二进制形式被传送到接收端。令 R 表示二进制编码中每个样值的比特数。则

$$L = 2^R \quad (3.29)$$

或写成等价形式:

$$R = \log_2 L \quad (3.30)$$

将式(3.29)代入式(3.25),得到步长为

$$\Delta = \frac{2m_{\max}}{2^R} \quad (3.31)$$

将式(3.31)代入式(3.28)可得

$$\sigma_Q^2 = \frac{1}{3} m_{\max}^2 2^{-2R} \quad (3.32)$$

令 P 表示消息信号 $m(t)$ 的平均功率。则均匀量化器的输出信噪比可表示为

$$\begin{aligned}(\text{SNR})_O &= \frac{P}{\sigma_Q^2} \\ &= \left(\frac{3P}{m_{\max}^2} \right) 2^{2R}\end{aligned}\quad (3.33)$$

式(3.33)说明量化器的输出信噪比随着每样值的比特数 R 的增大呈指数级增长。而 R 的增大需要信道(传输)带宽 B_T 成比例地增大,由此可见,用二进制编码表示消息信号(例如在 PCM 中)可以提供一种用增大信道带宽来换取噪声性能的方法,这种方法的效率比 FM 或 PPM 的效率都要高。上述结论的前提是,假设 FM 系统和 PPM 系统的性能均受到接收机噪声的限制,而二进制编码调制系统的性能受到量化噪声的限制。对于后一个问题在 3.8 节中还将继续进行探讨。

例 3.1 正弦调制信号

考虑幅值为 A_m 的满幅正弦调制信号,它利用了所提供的全部量化级。信号的平均功率为(假定是 1Ω 的负载):

$$P = \frac{A_m^2}{2}$$

因为调制信号在 $-A_m$ 和 A_m 之间波动,所以量化器输入的总范围是 $2A_m$ 。令 $m_{\max} = A_m$,在此条件下,利用式(3.32)算出量化噪声的平均功率(方差)为

$$\sigma_Q^2 = \frac{1}{3} A_m^2 2^{-2R}$$

对于满幅测试音信号,均匀量化器的输出信噪比为

$$(\text{SNR})_O = \frac{A_m^2/2}{A_m^2 2^{-2R}/3} = \frac{3}{2} (2^{2R}) \quad (3.34)$$

将信噪比用 dB 表示,则有

$$10 \log_{10}(\text{SNR})_O = 1.8 + 6R \quad (3.35)$$

对于不同的 L 和 R 值, 相应的信噪比由表 3.1 给出。在正弦调制下, 由表 3.1 可迅速地估算出在指定的输出信噪比下, 每个样值所需要的比特数。

表 3.1 正弦调制下不同量化级数对应的信噪比(或信号 - 量化噪声比)

量化级数, L	每样值比特数, R	信噪比(dB)
32	5	31.8
64	6	37.8
128	7	43.8
256	8	49.8

本节进一步讨论了如何描述无记忆均匀量化器以及如何评估其性能。但是, 在讨论的过程中并未涉及量化器最优化设计的问题, 即在固定的量化级数下, 为了使平均量化功率最小, 应如何选择量化级和分隔区间。遗憾的是, 由于量化过程的高非线性, 对于最优化问题并没有完整的解决方法。但通过一些有效的算法, 可以迭代地寻找最优设计方案。下面将讨论一种值得一提的著名算法, 即 Lloyd-Max 量化器。

3.6.2 标量量化器最优化的条件

在设计标量量化器时, 问题在于当量化级数固定时, 如何选取量化级和分隔区间, 从而使得平均量化功率最小化。

为了从数学上阐明这个问题, 首先考虑由平稳过程 $M(t)$ 得到的消息信号 $m(t)$ 。设 $m(t)$ 的动态范围 $-A \leq m \leq A$ 被分为 L 个区间, 如图 3.12 所示。分割区间的边界用一组实数 m_1, m_2, \dots, m_{L+1} 表示, 它们满足下面的条件:

$$\begin{aligned} m_1 &= -A \\ m_{L+1} &= A \\ m_k &\leq m_{k+1} \quad k = 1, 2, \dots, L \end{aligned}$$

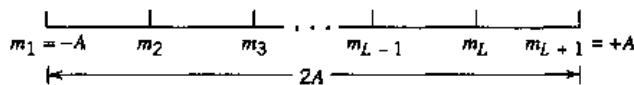
定义第 k 个分隔区间为

$$\mathcal{S}_k: m_k < m \leq m_{k+1} \quad k = 1, 2, \dots, L \quad (3.36)$$

用 v_k ($k = 1, 2, \dots, L$) 代表量化级(即量化值)。当用 v_k 表示输入 m 落在幅度区间 \mathcal{S}_k 的所有值时, 失真量用 $d(m, v_k)$ 表示。接下来的目标是找到使平均失真达到最小的两个集合 $\{v_k\}_{k=1}^L$ 和 $\{\mathcal{S}_k\}_{k=1}^L$ 。平均失真的表达式为

$$D = \sum_{k=1}^L \int_{m \in \mathcal{S}_k} d(m, v_k) f_M(m) dm \quad (3.37)$$

其中, $f_M(m)$ 表示抽样值为 m 的随机变量 M 的概率密度函数。

图 3.12 动态范围为 $-A \leq m \leq A$ 的消息信号 $m(t)$ 被分割为 L 个区间的示意图

衡量失真的普遍公式为

$$d(m, v_k) = (m - v_k)^2 \quad (3.38)$$

这里只考虑了均方失真。这里讲述的最优化问题都是非线性的，尚不存在确定的完整解决方法。为了克服这个困难，我们需要借助于一种迭代的算法。

从结构上看，量化器由具有相关设计参数的两部分组成：

- 位于发射机中用分隔区间集合 $\{\vartheta_k\}_{k=1}^L$ 表示的编码器。
- 位于接收机中用量化级集合 $\{v_k\}_{k=1}^L$ 表示的译码器。

相应地，可以定义两种极端情况，这两种情况能够为所有最优量化算法提供数学依据。其中一种情况是假设给定译码器，寻求发射端的最优编码器。另一种情况是假设给定编码器，寻求接收端的最优译码器。今后，分别称这两种情况为情况 I 和情况 II。

情况 I 给定译码器时编码器的最优化

给定译码器意味着有确定的码本。将码本定义如下：

$$\mathcal{C} : \{v_k\}_{k=1}^L \quad (3.39)$$

给定码本 \mathcal{C} ，此时的问题就是找到一组幅度区间 $\{\vartheta_k\}_{k=1}^L$ ，使得平均失真 D 为最小。也就是说，希望找到由下述非线性映射定义的编码器：

$$g(m) = v_k, \quad k = 1, 2, \dots, L \quad (3.40)$$

于是有

$$D = \int_{-A}^A d(m, g(m)) f_M(m) dm \geq \sum_{k=1}^L \int_{m \in \vartheta_k} [\min_{v_j \in \mathcal{C}} d(m, v_j)] f_M(m) dm \quad (3.41)$$

为了达到式(3.41)中的下界，要求式(3.40)的非线性映射仅在下述条件下成立：

$$d(m, v_k) \leq d(m, v_j) \quad j \neq k \quad (3.42)$$

在给定码本 \mathcal{C} 时，式(3.42)所示最优化编码器的必要条件，称为最近邻条件。具体地说，最近邻条件要求幅度区间 ϑ_k 包含输入 m 的所有比码本 C 中任意元素都更接近于 v_k 的值。

情况 II 给定编码器时译码器的最优化

下面考虑与情况 I 相反的情形，即假设给定编码器的幅度区间集合 $\mathcal{G} = \{\vartheta_k\}_{k=1}^L$ ，要寻找最优译码器的所有码本 $\{v_k\}_{k=1}^L$ 。用平均(均方)失真作为最优化的标准：

$$D = \sum_{k=1}^L \int_{m \in \vartheta_k} (m - v_k)^2 f_M(m) dm \quad (3.43)$$

概率密度函数 $f_M(m)$ 明显独立于码本 \mathcal{C} 。将 D 对量化级 v_k 求导，易得

$$\frac{\partial D}{\partial v_k} = -2 \sum_{k=1}^L \int_{m \in \vartheta_k} (m - v_k) f_M(m) dm \quad (3.44)$$

令 $\partial D / \partial v_k$ 等于零并求解 v_k , 得到其最优值

$$v_{k,\text{opt}} = \frac{\int_{m \in \mathcal{S}_k} m f_M(m) dm}{\int_{m \in \mathcal{S}_k} f_M(m) dm} \quad (3.45)$$

式(3.45)的分母是抽样值为 m 的随机变量 M 位于幅度区间 \mathcal{S}_k 内的概率 p_k , 即

$$\begin{aligned} p_k &= P(m_k < M \leq m_k + 1) \\ &= \int_{m \in \mathcal{S}_k} f_M(m) dm \end{aligned} \quad (3.46)$$

相应地, 可将式(3.45)的最优化条件解释为, 在 M 位于幅度区间 \mathcal{S}_k 时, 量化级 v_k 应等于随机变量 M 的条件均值。因此, 当给定编码器时, 译码器最优化的条件为:

$$v_{k,\text{opt}} = E[M | m_k < M \leq m_{k+1}] \quad (3.47)$$

式中, E 表示取期望值, 式(3.47)在直观上也是成立的。

注意, 给定译码器时, 编码器最优化的最近邻条件(情况 I)是在普通的平均失真情况下得到的。而给定编码器时, 译码器最优化的条件平均要求(情况 II)是在均方失真的特殊情况下得到的。在任何时候, 这两个条件都是对标量量化器进行最优化所必须的。基本上, 设计量化器时, 总是先根据情况 I 最优化编码器, 然后根据情况 II 最优化译码器, 反复交替地进行这两个过程, 直至平均失真 D 达到最小值。按照这种方法设计的最优化量化器, 称为 Lloyd-Max 量化器^[4]。

3.7 脉冲编码调制

了解了抽样和量化过程后, 下面我们将认识脉冲编码调制(PCM)。前面提到, PCM 是最基本的数字脉冲调制方式。在 PCM 中, 消息信号用编码脉冲序列来表示, 即信号的时间和幅值都被表示为离散的形式。PCM 系统发射端的基本步骤是抽样、量化和编码, 如图 3.13(a)所示。抽样之前的低通滤波器作用是避免消息信号的混叠。量化和编码操作通常在同一电路中完成, 此电路称为模 - 数转换器。接收端的基本步骤是畸变信号的再生、译码和量化样值的重构, 如图 3.13(c)所示。必要时, 在传输路径的中间点也会加入再生过程, 如图 3.13(b)所示。当采用时分复用时, 为使整个系统正常工作, 收发两端必须建立同步, 这一点将在 3.9 节进行讨论。下面描述 PCM 系统的基本组成部分。

3.7.1 抽样

为了接近瞬时抽样过程, 用一连串矩形窄脉冲对输入消息信号进行抽样。根据抽样定理, 接收端若想重构消息信号, 抽样速率必须大于消息信号最高频率分量 W 的两倍。在实际情况中, 抽样器前端会放置一个低通抗混叠滤波器, 用于在抽样前滤掉大于 W 的频率。抽样过程将连续变化的消息信号(持续时间有限)简化为每秒存在有限个离散值的信号。

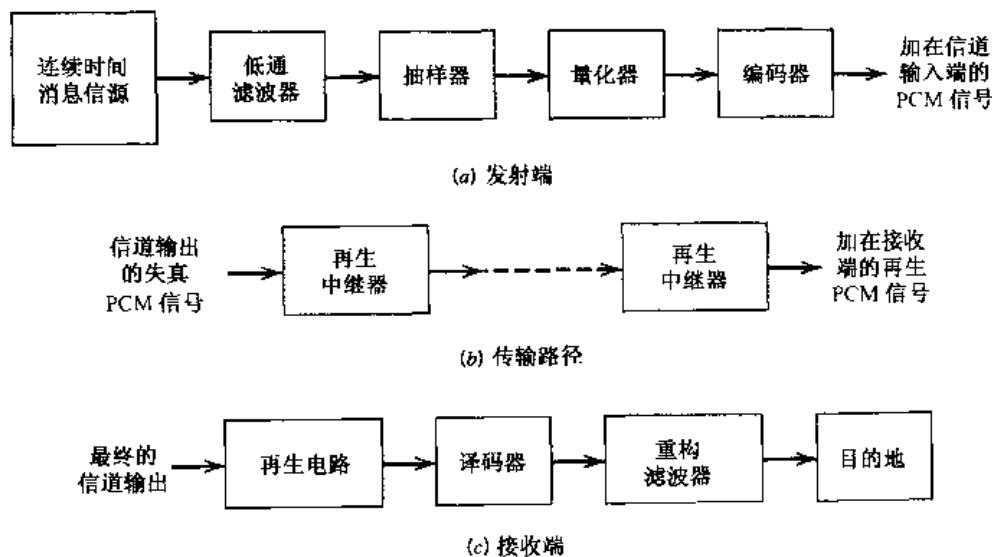


图 3.13 PCM 系统的基本组成

3.7.2 量化

将抽样后的消息信号进行量化，就可用时间和幅值均离散的形式来表示信号了。量化过程可采用 3.6 节讲到的均匀量化。但在电话通信中，一般更倾向于采用可变间隔的量化级。例如，从高声语音峰值到低声语音信号，其电压覆盖的范围为从 1000 到 1。非均匀量化器的步长随着输入 - 输出幅值特性曲线与原点之间距离的增大而增加，在量化器末端采用大的步长值，能够处理较少出现并位于大幅值范围内的语音信号。也就是说，以牺牲大信号为代价，使需要更多保护的小信号得到了改善。这种方式使得量化过程在输入信号的大部分幅度范围内都能够达到均匀的精度，其结果是在同等条件下，非均匀量化所需的量化级数少于均匀量化。

使用非均匀量化器，等价于将基带信号通过压缩器后再进行均匀量化。实际应用中，有一种特殊形式的压缩律，称为 μ 律^[5]，其定义为

$$|v| = \frac{\log(1 + \mu|m|)}{\log(1 + \mu)} \quad (3.48)$$

其中， m 和 v 为归一化的输入和输出电压， μ 为正的常数。图 3.14(a) 绘出了三个不同 μ 值下的 μ 律曲线。其中 $\mu = 0$ 时对应于均匀量化。给定 μ 值，定义压缩曲线斜率的倒数（即 $|m|$ 对 $|v|$ 的导数）为量化间隔，因此有

$$\frac{d|m|}{d|v|} = \frac{\log(1 + \mu)}{\mu} (1 + \mu|m|) \quad (3.49)$$

由此可见， μ 律既不是严格线性的也不是严格对数形式的。但是，在低输入电平，对应的 $\mu|m| \ll 1$ 时，曲线是近似线性的；在高输入电平，对应的 $\mu|m| \gg 1$ 时，曲线是近似对数形式的。

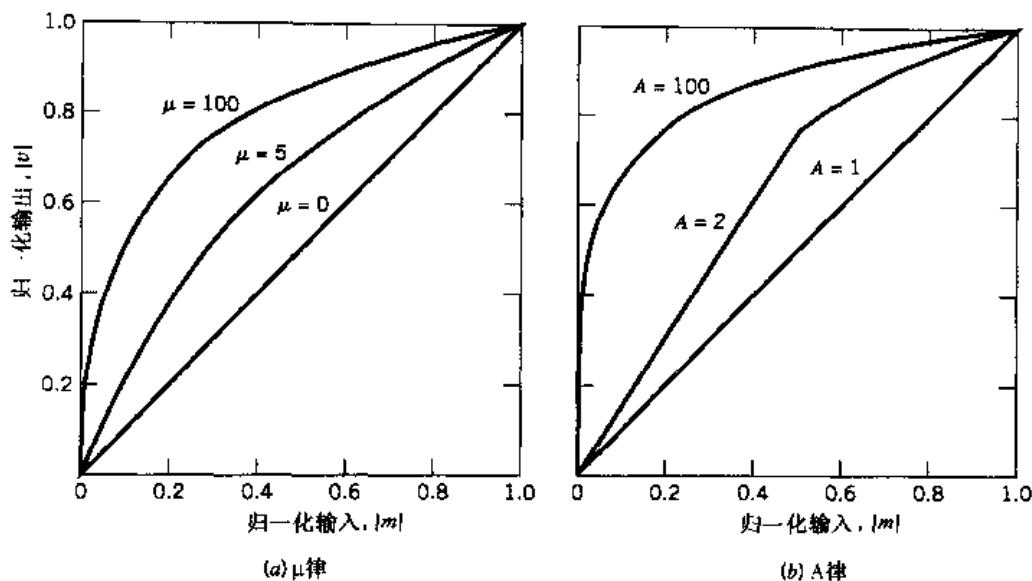


图 3.14 压缩律

实际应用中的另外一个压缩律称为 A 律, 定义为

$$|\nu| = \begin{cases} \frac{A|m|}{1 + \log A}, & 0 \leq |m| \leq \frac{1}{A} \\ \frac{1 + \log(A|m|)}{1 + \log A}, & \frac{1}{A} \leq |m| \leq 1 \end{cases} \quad (3.50)$$

图 3.14(b)画出了不同 A 值对应的曲线。其中 $A = 1$ 对应于均匀量化。 $|m|$ 对 $|\nu|$ 的导数等于第二条压缩曲线斜率的倒数, 用公式表示为(由输入 $|m|$ 归一化的值决定):

$$\frac{d|m|}{d|\nu|} = \begin{cases} \frac{1 + \log A}{A}, & 0 \leq |m| \leq \frac{1}{A} \\ (1 + A)|m|, & \frac{1}{A} \leq |m| \leq 1 \end{cases} \quad (3.51)$$

为了将信号抽样值恢复成正确的相应电平值, 必须在接收端使用与压缩器特性相逆的器件。这种器件称为扩张器, 理想情况下, 扩张律和压缩律恰好互逆, 因此, 若不考虑量化的影响, 扩张器的输出应等于压缩器的输入。压缩器和扩张器合称为压扩器。

对 μ 律和 A 律而言, 压缩器的动态范围分别随着 μ 和 A 的增大而扩大。低电平信号信噪比的增大是以降低高电平信号信噪比为代价的。为了同时满足这两个相互矛盾的要求(即使低电平和高电平信号均达到适当的信噪比), 在选择参数 μ 和 A 时, 常取折中值。实际使用的典型值为 $\mu = 255$ 和 $A = 87.6$ 。

值得注意的是, 在实际 PCM 系统的压缩电路中, 得到的并不是与图 3.14 完全吻合的非线性压缩曲线, 而是理想曲线的分段线性近似曲线。通过选取足够的线段, 可使近似曲线与真正的压缩曲线非常接近。例 3.2 阐述了这种近似情况。

3.7.3 编码

将抽样和量化结合,连续消息(基带)信号就变有限个离散值的集合。但这并不是最适于在电话线或无线路径中传输的信号形式。为了更好地利用抽样和量化的优势,使传输信号具有更强的抗噪、抗干扰和抗信道畸变的能力,还需采用编码过程,将抽样值的离散集合变为更适于传输的信号形式。任何将离散集合中每个值都表示为离散事件的一个特定排列的方案均可称为码。码中的一个离散事件称为码元或符号。例如,脉冲的有或无就是一个符号。码中用于表示离散集合中某个值的符号特定排列,称为码字或字符。

在二进制码中,每个符号可以是两个不同值(或不同种类)中的某一个,例如脉冲的有或无。习惯上,二进制码中的两个符号用 0 和 1 表示。在三进制码中,每个符号可以是三个不同值(或不同种类)中的某一个,其他代码以此类推。由于二进制符号抵抗噪声电平的能力相对较强,因此其再生也相对容易,所以一般采用二进制码。在二进制码中,每个码字由 R 比特(bit)组成(bit 是 binary 的缩写)。 R 代表每个样值的比特数。采用这种码,可表示 2^R 个不同的数值。例如,一个量化级为 256 的样值可用 8 比特的码字表示。

在量化级和码字之间建立一一对应关系的方法有很多。一种简单的方法是将表示量化级的十进制数转化为二进制数。在二进制数字系统中,每个数字的权重为 2。表 3.2 所示为每个样值采用 4 比特编码(即 $R = 4$)时的情况。

表 3.2 $R = 4$ 比特/样值的二进制系统

量化级的十进制数	用 2 的加权和表示的量化级数	二进制数
0		0000
1	2^0	0001
2	2^1	0010
3	$2^1 + 2^0$	0011
4	2^2	0100
5	$2^2 + 2^0$	0101
6	$2^2 + 2^1$	0110
7	$2^2 + 2^1 + 2^0$	0111
8	2^3	1000
9	$2^3 + 2^0$	1001
10	$2^3 + 2^1$	1010
11	$2^3 + 2^1 + 2^0$	1011
12	$2^3 + 2^2$	1100
13	$2^3 + 2^2 + 2^0$	1101
14	$2^3 + 2^2 + 2^1$	1110
15	$2^3 + 2^2 + 2^1 + 2^0$	1111

3.7.4 线路码

任何一组线路码都可用于表示二进制数据流电平。图 3.15 为数据流 01101001 对应的 5 种重要线路码的波形。图 3.16 为这 5 种线路码各自对应的随机二进制数据的功率谱(正频率

部分),其中,假设:

- 符号 0 和 1 是等概率出现的。
- 归一化平均功率等于 1。
- 将频率 f 关于比特率 $1/T_b$ 进行归一化(归一化得到的公式用于绘制图 3.16 所示的功率谱图,见习题 3.11)。

对图 3.15 所示五种线路码的描述分别如下:

1. 单极性不归零(NRZ)信号

在符号持续时间内,用幅值为 A 的脉冲表示符号 1,无脉冲表示符号 0,如图 3.15(a)所示。这种线路码也称为通-断信号。由于这种信号需传送直流电平,且发射信号的功率谱在零频处不为零,因此通-断信号的缺点是浪费功率。

2. 双极性不归零(NRZ)信号

用幅值为 $+A$ 和 $-A$ 的脉冲分别表示符号 1 和 0,如图 3.15(b)所示。这种线路码容易产生,其缺点是在零频附近,信号的功率谱值较大。

3. 单极性归零(RZ)信号

用占空比为 $1/2$,幅值为 A 的矩形脉冲表示符号 1,无脉冲表示符号 0,如图 3.15(c)所示。这种线路码的显著特点是,所传输信号的功率谱在 $f = 0, \pm 1/T_b$ 处出现了冲激函数,可用于接收端比特定时(bit-timing)的提取。在同样的符号差错概率下,它需要的功率比双极性归零信号多 3 dB 。第 4 章习题 4.10 将探讨这一问题。

4. 双极性归零(BRZ)信号

如图 3.15(d)所示,这种线路码采用了三个幅值电平。交替使用幅值相同(如 $+A$ 和 $-A$)的正、负脉冲来表示符号 1,其中每一脉冲的占空比为 $1/2$;用无脉冲来表示符号 0。BRZ 信号的一个有用特性是传输信号的功率谱没有直流(DC)分量,且当符号 1 和 0 等概率出现时,功率谱仅含有很少的低频分量。这种线路码也称为交替极性反转(AMI)信号。

5. 多相(Split-phase)码(曼彻斯特码)

如图 3.15(e)所示,在这种线路码中,用一个幅值为 A 的正脉冲后接一个幅值为 $-A$ 的负脉冲来表示符号 1,两个脉冲的占空比均为 $1/2$ 。将这两个脉冲的极性反转来表示符号 0。曼彻斯特码抑制了直流分量,且仅含有较少的低频分量,这与信号的统计特性无关。在某些应用中,这种性质相当重要。

3.7.5 差分编码

这种编码方法通过信号的跃变来编码信息。特别地,可以用跃变来表示输入二进制数据流中的符号 0;而用无跃变来表示符号 1,如图 3.17 所示。图 3.17(b)是对图 3.17(a)中的数据进行差分编码后的数据流。这里使用的原始二进制数据流与图 3.15 中相同。图 3.17(c)为差分编码数据的波形。从图 3.17 中可明显地看出,将差分编码信号反转并不会影响译码。仅通过比较相邻二进制符号的极性,确定是否发生了跃变,就能够恢复出原始的二进制信号。注意,在编码过程开始之前,应设定一个参考比特。在图 3.17 中,参考比特为符号 1。

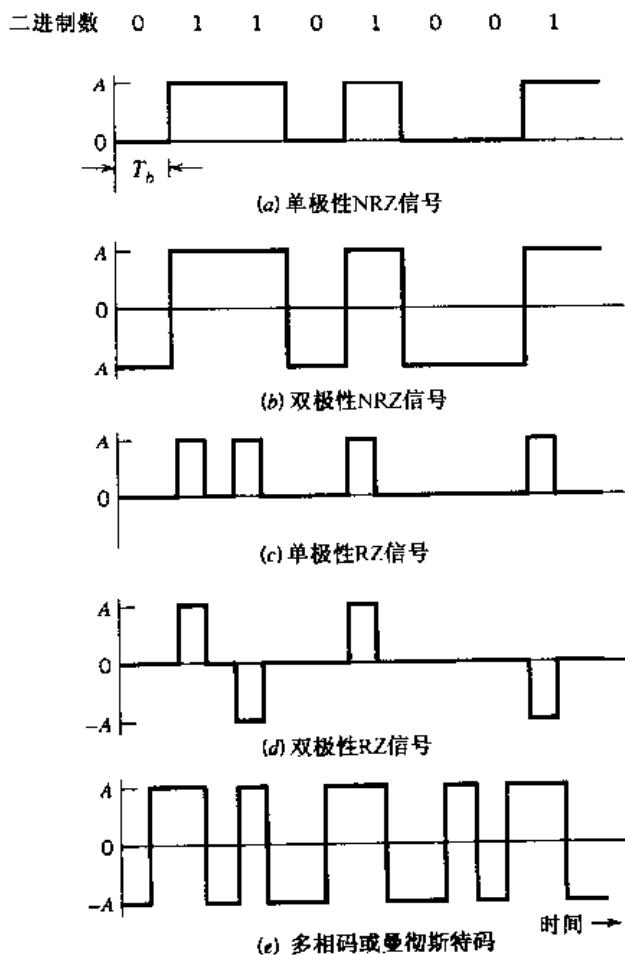


图 3.15 表示二进制数据电平的线路码

3.7.6 再生

PCM 系统最显著的特点是具有控制失真和噪声的能力,这种失真和噪声是 PCM 信号通过信道传输时带来的。在传输线路上放置一组距离足够近的再生中继器,用以重构 PCM 信号,就能够实现 PCM 系统对失真和噪声的控制能力。如图 3.18 所示,再生中继器具有三个基本功能:均衡、定时和判决。均衡器对接收到的脉冲进行整形,以补偿由于信道传输特性的不理想所引起的脉冲串幅度和相位的失真。在定时电路中,由接收脉冲产生周期性的脉冲串,用于在信噪比最大的瞬时对均衡后的脉冲进行抽样。然后将提取到的每个样值与判决器里的预定门限进行比较。在每比特持续时间内,根据样值是否超过门限值来确定接收符号是 1 还是 0。如果样值大于门限值,就将一个表示符号 1 的新脉冲送入下一个中继器。否则,就发送一个表示符号 0 的新脉冲。假定判决过程中的干扰不足以造成误码,则中继器之间失真和噪声的累积就能够通过上述过程完全去除。在理想情况下,除了存在时延以外,再生信号应与原始信号完全相同。但在实际情况中,由于以下两个主要原因,再生信号与原始信号之间存在差异:

1. 不可避免的信道噪声和干扰,使得中继器偶尔会发生错误判决,从而在再生信号中引入了比特差错。

2. 如果接收脉冲之间的距离偏离了固定分配值,那么在再生脉冲位置上就会产生抖动,从而造成失真。

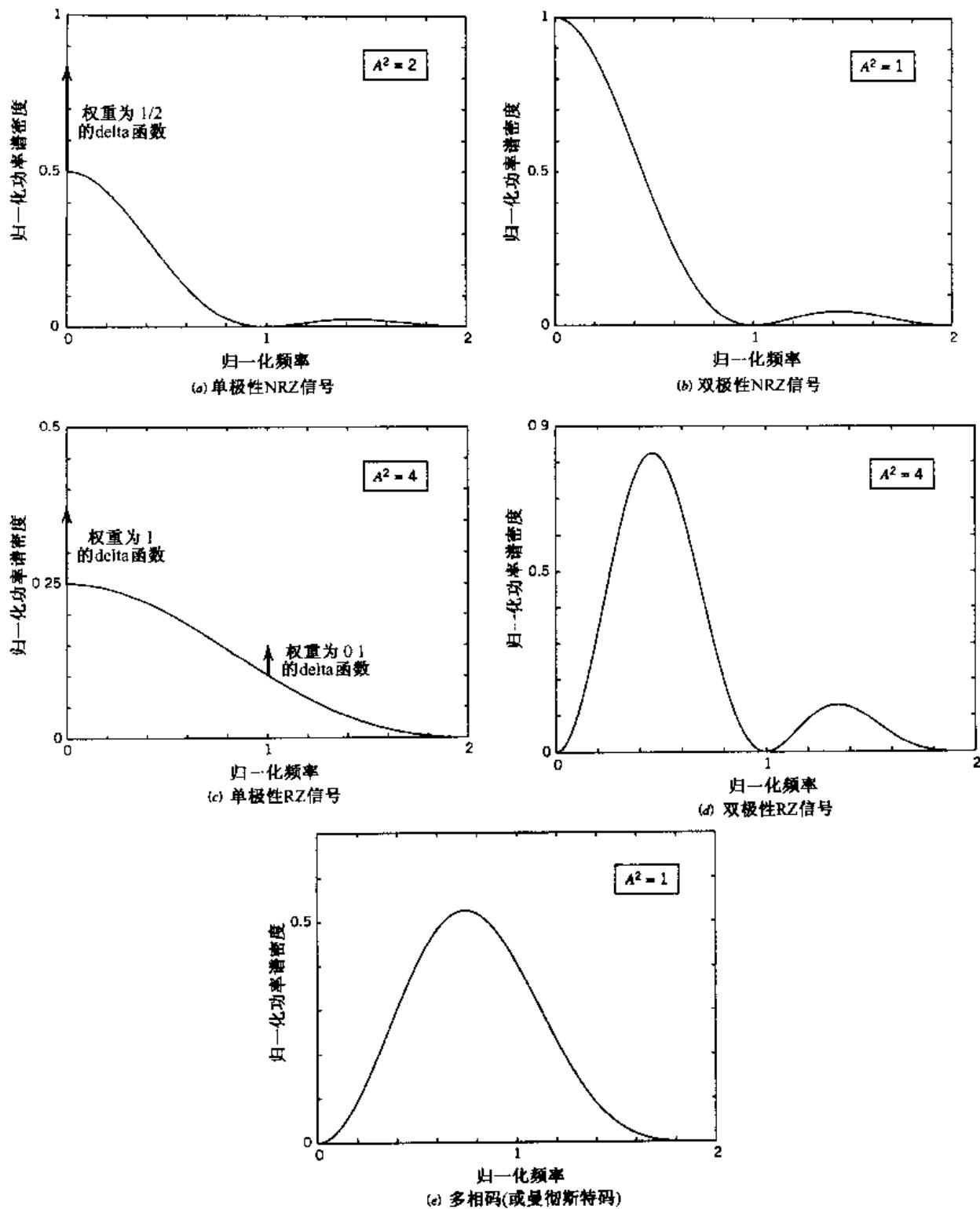


图 3.16 线路码的功率谱。其中频率关于比特率 $1/T_b$ 进行了归一化,且平均功率归一化为 1

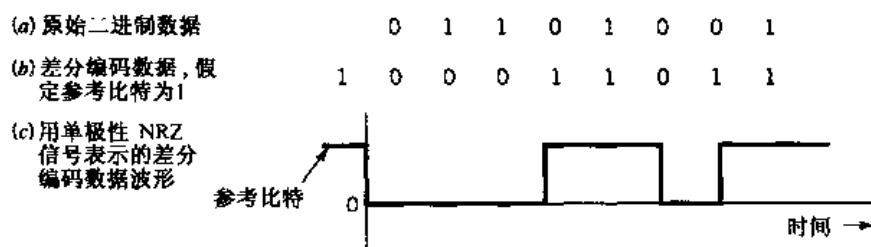


图 3.17 差分编码

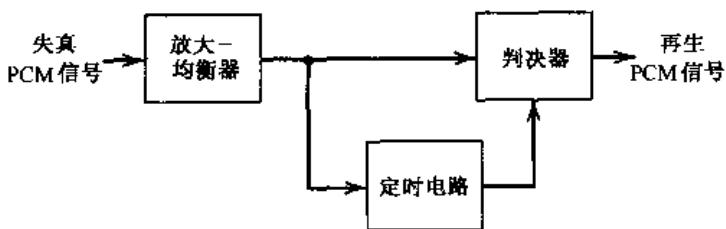


图 3.18 再生中继器框图

3.7.7 译码

接收机的第一步操作是最后一次再生(即再成形和净化)接收脉冲。接着将这些净化后的脉冲重新组成码字,然后译码(即反映射)成已量化的PAM信号。译码过程即产生一个脉冲,其幅值等于将码字中所有脉冲按各自在代码中的位置值($2^0, 2^1, 2^2, \dots, 2^{R-1}$)进行加权操作后的线性和,其中 R 为每样值的比特数。

3.7.8 滤波

接收机的最后一步操作是将译码器的输出送入截止频率等于消息带宽 W 的低通重构滤波器,从而恢复出消息信号。假设传输路径是无误差的,那么,恢复出来的信号中除了包含由量化过程带来的原始失真外,并不包含噪声。

3.8 PCM系统中的噪声

PCM系统的性能主要受两个噪声源影响:

1. **信道噪声**。在发射机输出端和接收机输入端之间的任何位置均可引入。只要设备处于工作状态,就会存在信道噪声。
2. **量化噪声**。在发射机中产生,并一直被带到接收机的输出中。与信道噪声对系统性能不同的是,量化噪声依赖于信号,即切断消息信号时,量化噪声将会消失。

只要PCM系统处于工作状态,这两种噪声源就会同时出现。但传统的分析方法是将它们分别进行考虑,以便考察它们各自对系统性能造成的影响。

信道噪声对系统性能的主要影响是在接收信号中引入了比特差错。在二进制PCM系统中,比特差错表现为将符号1误认为符号0,或将符号0误认为符号1。很明显,比特差错发生

的频率越高,接收机输出信号与原始消息信号之间的差异就越大。在存在信道噪声的情况下,可以用平均符号差错概率的来衡量 PCM 信息传输的逼真度。平均符号差错概率的定义为,接收机输出的重构符号与所传输的二进制符号不相同的平均概率。在原始二进制波形中的所有比特均具有相同重要性的条件下,平均符号差错概率又称为误比特率(BER)。但是,在重构原始消息信号的模拟波形时,不同的符号差错可能需要区别对待。例如,码字(表示消息信号的量化抽样值)中最最重要的比特发生的差错,要比最不重要的比特发生的差错有害得多。

在存在信道噪声的情况下,为了将系统性能最优化,需使平均符号差错概率降至最低。在平均符号差错概率的计算中,习惯上将信道噪声看成是加性的高斯白噪声。实际上,在 PCM 系统再生中继器之间的距离足够短,且信号能量与噪声密度的比值足够大的情况下,信道噪声的影响是可以忽略的。在这种情况下,PCM 系统的性能仅受量化噪声的限制。

从 3.6 节对量化噪声的讨论可知,量化噪声实质上是受设计者控制的。当量化器量化级数足够大,且压缩方法适合于所传输的消息信号特性时,量化噪声带来的影响是可以忽略的。由此可见,采用 PCM 调制可以构造出性能不受信道噪声限制的通信系统,这是任何连续波调制或模拟脉冲调制系统都不具备的优点。

3.8.1 差错门限

计算 PCM 系统误比特率的基础理论将在第 4 章介绍。在二进制编码 PCM 接收机中,由加性高斯白噪声引起的平均符号差错概率仅由 E_b/N_0 决定。 E_b/N_0 的定义为每比特的发射信号能量 E_b 与噪声功率谱密度 N_0 的比值。注意,虽然 E_b 和 N_0 的物理含义不同,但比值 E_b/N_0 是无量纲的。表 3.3 总结了平均符号差错概率与 E_b/N_0 之间的关系,此时二进制 PCM 系统采用的是双极性 NRZ 信号。表中最后一列的结果是在比特率为 10^5 bps 的条件下得到的。

表 3.3 E_b/N_0 对差错概率的影响

E_b/N_0	差错概率 P_e	比特率为 10^5 bps 时, 发生一个错误的大约时间间隔
4.3 dB	10^{-2}	10^{-3} 秒
8.4	10^{-4}	10^{-1} 秒
10.6	10^{-6}	10 秒
12.0	10^{-8}	20 分
13.0	10^{-10}	1 天
14.0	10^{-12}	3 个月

由表 3.3 可看出差错门限的存在(大约在 11 dB)。当 E_b/N_0 低于差错门限时,接收机将大量的差错;当 E_b/N_0 高于差错门限时,信道噪声的影响实际上是可以忽略的。也就是说,如果比值 E_b/N_0 超过了差错门限,信道噪声对接收机的性能实质上没有任何影响,这正是 PCM 所希望达到的目的。但是,当 E_b/N_0 低于差错门限时,接收机的出错率会急剧增大。由于判决错误会导致不正确码字的产生,所以当错误频繁发生时,接收机输出的重构信号与原始信号的差异会相当大。

采用双极性 NRZ 信号的 PCM 系统差错门限是 11 dB。而在幅度调制中,高质量传输语音信号需要的 E_b/N_0 是 60~70 dB。比较这两者可见,虽然在 PCM 系统中平均噪声功率随着带宽呈 R 次幂增长, R 为一个码字中的比特数(即每样值的比特数),系统所需功率还是相当小的。

在大多数传输系统中,各段链路的噪声和失真的影响会累积起来。在整体传输质量一定的情况下,发射机和接收机之间的物理距离越长,对系统内各段链路的性能要求就越高。但是,在PCM系统中,由于信号在必要的时候可以进行再生,所以,一条链路发生的幅度、相位或非线性的失真(如果不是很严重)对下一条链路的再生输入实际上是没有影响的。若比值 E_b/N_0 高于门限值,则信道噪声的影响实质上也是可以忽略的。因此,实际上,对PCM链路的传输要求几乎与通信信道的物理长度无关。

PCM系统的另一个重要特点是抗干扰能力强,这种干扰是由脉冲漂移或串话造成的。信道噪声和干扰同时存在势必会导致保证PCM系统正常工作的差错门限值增大。如果能预先对差错门限多留些余量,那么系统就能够抵抗更多的干扰。也就是说,PCM系统具有很强的抗信道噪声和抗干扰能力。

3.9 时分复用

抽样定理为传输带限消息信号 $m(t)$ 携带的信息提供了理论基础, $m(t)$ 携带的信息是指用稍高于奈奎斯特速率的速率对 $m(t)$ 进行均匀抽样得到的抽样值序列。抽样过程的一个主要特点是时间守恒。也就是指消息样值的传输周期性,即消息样值只在抽样间隔的部分时间内占用通信信道。这样,相邻样值的时间间隔可被不同的独立消息源按分时的方式使用。由此,产生了时分复用(TDM)系统,此系统允许很多相互独立的消息源在互不干扰的情况下共享通信信道。

TDM的概念由图3.19的框图说明。首先,用低通抗混叠滤波器滤掉不必要的高频分量,从而将每个输入消息信号限制在带宽范围内。然后,将低通滤波器的输出送入转接器中,转接器通常是电子控制的切换电路。转接器的功能有两个:(1)以稍高于 $2W$ 的速率 f_s 对 N 个输入信号分别进行窄带抽样,其中 W 为抗混叠滤波器的截止频率;(2)将这 N 个抽样值顺序排列在抽样间隔 T_s 内。事实上,转接器的后一个功能正是时分复用的实质。在转接过程之后,将复用信号送入脉冲调制器中,其目的是为了将复用信号转换成适合在共用信道上传输的信号形式。很明显,时分复用引入了带宽扩张因子 N ,因为这种复用方案必须将从 N 个相互独立的消息源得到的 N 个抽样值放入等于1个抽样间隔的时隙中。在系统的接收端,将收到的信号送入脉冲解调器中,脉冲解调器的作用与脉冲调制器正好相反。通过反转接器将脉冲解调器输出的窄带样值分配给合适的低通重构滤波器,反转接器与发射端的转接器是同步工作的。同步对于系统的正常工作是非常必要的。这种同步的完成由传输复用样值序列时所采用的脉冲调制方式决定。

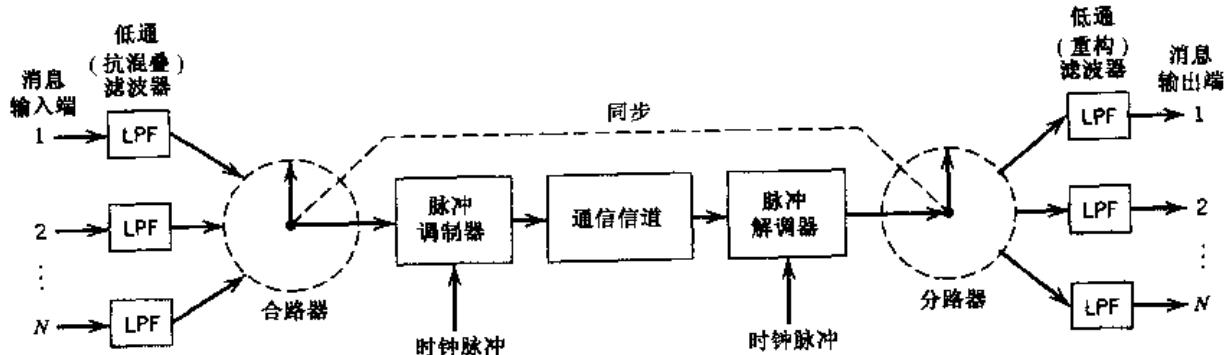


图3.19 TDM系统框图

TDM 系统对于共用信道的弥散现象非常敏感。弥散现象是指幅度随频率发生了变化,或者相位与频率不再是正比例关系了。因此,对信道的幅度和相位响应进行正确的均衡是保证系统正常工作的必要条件,这个问题将在第 4 章中进行讨论。不同于频分复用的是,时分复用的一阶近似作为串话源的时候不受信道非线性的影响。其原因是不同的消息信号并没有同时加在信道上。

3.9.1 同步

采用 PCM 时,很自然会想到将不同的消息源通过时分方式进行复用,每个消息源从发端到收端保持各自的独立性不变。这种独立性使得在 TDM 系统中取出或插入消息源变得比较容易。由于所有消息源必须排列在等于它们抽样速率倒数的同一时间间隔内,因此随着独立消息源数量的增加,分配给每个消息源的时间间隔肯定会减小。反过来说,这意味着代表单个样值的码字的持续时间减小了。当持续时间降低时,脉冲的产生和传输会变得更加困难。进一步说,如果脉冲变得太窄,传输媒质中的畸变可能就会干扰系统的正常工作。因此,在实际应用中,有必要限制同一时分群内的独立消息源数量。

在任何情况下,除了传输以及再生中继需要损耗掉的时间外,接收端的时序操作都应当紧随发射端的相应操作进行工作,这是保证时分复用 PCM 系统正常工作的必要条件。通常,除了接收机中的本地时钟比发射机中用于发射信号的标准时钟稍慢以外,两者在时间步调上要保持一致。使发射机和接收机时钟保持同步的一种可行方法是在每帧(由各个独立消息源的码字连在一起组成)的末尾留出一个码元或脉冲,然后每隔一帧只传输这个脉冲。在这种情况下,接收机中的电路将交替搜索 0 和 1 模式,搜索速率为帧速率的一半,从而在发射机和接收机之间建立起同步。

当传输路径被中断时,发射机和接收机的时钟将不会长时间地继续保持一致。因此,在进行同步过程时,必须建立一个有序的过程来检测同步脉冲。该过程需要逐一地观察码元,直到检测到同步脉冲为止。即通过足够长的时间观察某个码元并确保它不是同步脉冲后,接收机的时钟将放过这个码元,接着观察下一个。重复此搜索过程,直至检测到同步脉冲。由此看出,同步需要的时间由重新建立正确传输的时间点决定。

例 3.2 T1 系统

此例描述 PCM 系统中的 T1 系统^[6]的重要特性,T1 系统用互相分离的电缆对来承载 24 个语音信道,其中,再生中继器之间的距离约为 2 km。T1 载波系统是 3.10 节中介绍的北美数字交换系列的基础。

语音信号(男声或女声)主要限制在 300~3100 Hz 的频带内,这个范围之外的频率对语音清晰度的影响不大。事实上,规定电话线路只对此频率范围产生响应,会带来很好的服务质量。所以,通常在抽样之前先将语音信号通过截止频率约为 3.1 kHz 的低通滤波器。由 $W = 3.1 \text{ kHz}$ 可知奈奎斯特速率为 6.2 kHz。对滤波后的语音信号,通常以稍高的速率(如 8 kHz)进行抽样。8 kHz 是电话系统中的标准抽样速率。

在压缩信号时,T1 系统利用分段线性特性(由 15 折线组成)来近似表示式(3.48)中的对数 μ 律曲线,取 $\mu = 255$ 。通过使各线段的端点位于式(3.48)求得的压缩曲线上,并使各线段在纵轴上的投影等间距,即可构造出上述近似表示。表 3.4 给出了各线段的端点在横轴上的投影值以及各段的步长。此表中的值对 8199 进行了归一化,因此,所有值均为整数。折线的

第0段是直线段,它经过原点,且含有30个均匀判决电平。线段 $1a, 2a, \dots, 7a$ 在横轴的上方,而线段 $1b, 2b, \dots, 7b$ 在横轴的下方,这14条线段均含有16个均匀判决电平。对于第0段,量化器输入端的判决电平为 $\pm 1, \pm 3, \dots, \pm 31$,量化器输出端的相应量化级为0, $\pm 1, \dots, \pm 15$ 。对于线段 $1a$ 和 $1b$,量化器输入端的判决电平为 $\pm 31, \pm 35, \dots, \pm 95$,量化器输出端的相应量化级为 $\pm 16, \pm 17, \dots, \pm 31$,其他各段以此类推。

表 3.4 15 段压缩曲线($\mu = 255$)的特性

线段数	量化间隔	线段端点在水平轴上的投影
0	2	± 31
$1a, 1b$	4	± 95
$2a, 2b$	8	± 223
$3a, 3b$	16	± 479
$4a, 4b$	32	± 991
$5a, 5b$	64	± 2015
$6a, 6b$	128	± 4063
$7a, 7b$	256	± 8159

上述15段压缩特性曲线共有 $31 + (14 \times 16) = 255$ 个量化级。为了提供这些量化级,24个语音信道中的每一个均采用8比特码字的二进制编码。码字中第一个比特用来表示输入语音抽样值的极性。如果该比特为1,说明样值为正;若该比特为0,则说明样值为负。接下来的3个比特用于表示输入语音样值的幅度位于哪一个线段内,最后的4比特表示样值位于该线段内的哪个量化级。

在8 kHz的抽样速率下,复用信号的每一帧占125 μs 的周期。每一帧由24个8比特的码字和位于帧尾用于同步的1个比特组成。因此每一帧的比特总数为 $(24 \times 8) + 1 = 193$ 比特。相应地,每个比特的持续时间等于0.647 μs ,传输速率等于1.544 Mbps。

除了语音信号,电话系统还需要专门的监控信号,即信令信息。信令信息用于传送拨号脉冲和电话摘/挂机信号。在T1系统中,信令信息按照下述方法进行传送。每隔6帧,将每个语音信道中最不重要的比特(即第8比特)删除,并在此位置上插入信令比特,因此每个语音输入信号的平均信息比特数为 $7\frac{5}{6}$ 比特。从而,信令比特序列的传输速率等于抽样速率8 kHz除以6,即1.333 Kbps。该信令速率适用于24个输入信道中的每一个信道。

3.10 数字复接器

3.9节中介绍了时分复用的思想,即以同一抽样速率按时间顺序对一组模拟信号(如语音信号)进行抽样,经过复用后通过共用线路进行传输。本节将探讨不同比特率的数字信号的复用。这种复用使得多种数字信号如计算机输出、数字化的语音信号、数字化传真以及电视信号能够合成一个数据流(其比特率比其中一个输入信号的比特率高出很多)进行传输。

图 3.20 为数字复接 - 分接原理框图。

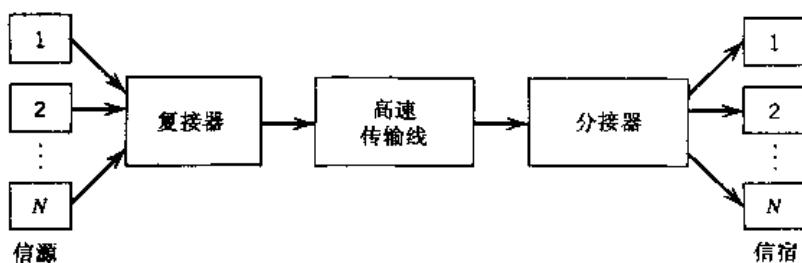


图 3.20 复接 - 分接原理框图

使用选择开关按顺序从每个输入线路中取出 1 比特, 然后将这些比特放入高速共用线路中。通过这种逐比特的插入过程, 即可完成数字信号的复用。在系统的接收端, 共用线路的输出被分为多个相应的低速码流, 然后被送至各自的信宿。

数字复接器主要分为两组。一组用于复用来自数字计算机的低比特率数据流, 并将复用后的低速数据通过公众交换电话网(PSTN)进行时分复用的传输。这类复接器的实现需要使用调制解调器, 这将在第 6 章中进行讨论。

第二组数字复接器组成了电信载体, 如 AT&T 的部分数据传输服务系统。特别地, 这些复接器构成了一种数字系列, 可将低速率比特流时分复用成更高速率的比特流。各国采用的复用系列比特率虽然各不相同, 但均存在一个普遍特点: 最低速率为 64 Kbps(即语音信号的标准 PCM 传输速率)。任何 64 Kbps 的输入比特流, 无论来自哪一类终端, 均称为零次群数字信号(DS0)。美国、加拿大和日本^[7]采用北美数字时分复用系列, 描述如下:

- 第一级复接器将 24 个 DS0 比特流合成为 1.544 Mbps 的一次群数字信号(DS1), 此过程在例 3.2 描述的 T1 系统中进行。1.544 Mbps 称为数字系列的基本速率, 因为它是数字交换机能够处理的最低比特率。数字交换机是一种有记忆的逻辑器件, 其功能只是进行数字信号的交换, 由此得名。
- 第二级复接器将 4 个 DS1 比特流合成为 6.312 Mbps 的二次群数字信号(DS2)。
- 第三级复接器将 7 个 DS2 比特流合成为 44.736 Mbps 的三次群数字信号(DS3)。
- 第四级复接器将 6 个 DS3 比特流合成为 274.176 Mbps 的四次群数字信号(DS4)。
- 最后一级即第五级复接器将两个 DS4 比特流合成为 560.160 Mbps 的五次群数字信号(DS5)。

需注意, 上述复接器产生的数字信号的比特率, 均稍高于输入比特率与相应复用次数的乘积, 因为在每一级复用中均进行了比特填充。比特填充将在随后的内容中进行讨论。

此外, 应认识到数字传输设备的功能只是传递比特流, 而对于比特本身的含义不予解释。而由传输设备两端的数字交换机来对数据流中的比特含义进行解释, 如比特是代表语音还是数据, 是帧格式还是信令格式, 等等。

无论哪组数字复接器, 设计中均需遵循以下原则:

1. 数字信号不能直接进行插入操作, 除非它们的比特率锁定在同一时钟频率下。此外, 为使输入数字信号同步, 从而保证其正确插入, 需要对其进行存储。
2. 复用信号必须包含某种形式的标志, 这样在接收端才能识别出各自的分量。

3. 复接器必须有能力处理输入数字信号比特率的微小变化。例如，一根载有每秒 3×10^8 个脉冲的长为 1000 km 的同轴电缆，将同时传递约 100 万个脉冲，每个脉冲约占用 1 m 长的电缆。如果温度降低 1°F，传输时延就会增加 0.01%，这将使电缆少传 100 个脉冲。显然，复接器必须有能力进行迅速的调整，以确保全部脉冲都能正确接收。

采用比特填充技术既能够达到同步的要求，又能够使复接器调整速率以适应输入数据速率的微小变化。比特填充的思想是额外地填充不含信息量的脉冲，使复接器的输出比特率稍高于各输入信道比特率的最大值之和。所有的输入数字信号中都填充了足够多的比特数，使得填充后的比特率等于本地时钟产生的比特率。为了完成比特填充过程，应将每个输入数字信号或比特流送入复接器的缓冲寄存器(Elastic Store)中。缓冲寄存器是这样一种器件，经过该器件存储的比特流的读出速率可以不等于它的写入速率。在分接器中，必须将填充比特从复用信号中取出。这里需采用一种识别填充比特的方法。为了阐述这种方法，也为了对帧同步进行说明，下面我们介绍 AT&T M12 复接器的信号格式。该复接器将四个 DS1 比特流复用为一个 DS2 比特流。这是前面讨论的第二级数字复用等级。

例 3.3 AT&T M12 复接器的信号格式

图 3.21 说明了 M12 复接器的信号格式。每帧被分成四个子帧。先传第一个子帧(图 3.21 的第一行)，然后是第二个、第三个、第四个，按此顺序进行传送。

对输入的四个 DS1 比特流各取 12 比特，进行逐位插入，合成 48 比特的码流。然后由复接器插入一个控制比特。每帧共含 24 个控制比特，分别被含 48 个比特的数据序列隔开。M12 复接器共有三种类型的控制比特，分别用于提供同步、帧标识以及确认四个输入信号中的哪一个经过了填充。在图 3.21 中用 F 、 M 和 C 来表示这些控制比特。它们的功能如下：

1. 每个子帧含两个 F 控制比特，构成主帧脉冲。 F 控制比特的下标表示实际传输的比特(0 或 1)。因此主帧序列为 $F_0 F_1 F_0 F_1 F_0 F_1 F_0 F_1$ 或 01010101。
2. 每个子帧含一个 M 控制比特，构成识别四个子帧的次帧脉冲。同理， M 控制比特的下标也表示实际传输的比特(0 或 1)。因此，次帧序列为 $M_0 M_1 M_1 M_1$ 或 0111。
3. 每个子帧含三个 C 控制比特，为填充标识符。特别地， C_I 表示输入信道 I， C_{II} 表示输入信道 II，以此类推。例如，第一个子帧中 M_0 后跟随的三个 C 控制比特是第一个 DS1 比特流的填充标识符。将三个 C 控制比特都设置为 1，表明在该 DS1 比特流中插入了填充比特。若未进行填充，则将它们全部设置为 0。如果这三个 C 控制比特均设置为 1，则填充比特应位于同一子帧内 F_1 控制比特之后第一个 DS1 比特流中第一个信息比特所处的位置上。同理，如果需要的话，第二、第三、第四个 DS1 比特流也照此法进行填充。在接收端通过多数逻辑译码，就能够检测出三个 C 控制比特中任意一个的错误。这种译码方法只是简单地通过 C 控制比特中的大多数比特来决定传送的是全 1 还是全 0 序列。因此三个 1 或者两个 1 与一个 0 均表示信息序列中相关子帧里控制比特 F_1 之后的是填充比特。另外，三个 0 或者两个 0 与一个 1 则表示没有进行填充。

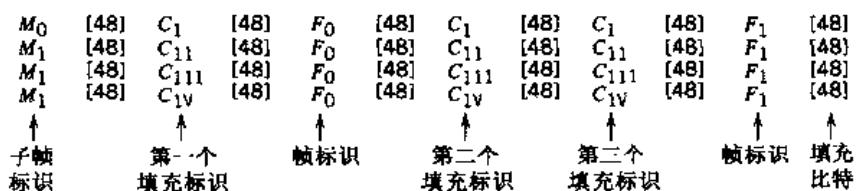


图 3.21 AT&T M12 复接器的信号格式

M12 接收单元的分接器首先搜索主帧序列 $F_0 F_1 F_0 F_1 F_0 F_1 F_0 F_1$ 。由此确认收到了四个输入 DS1 比特流，也收到了 M 和 C 控制比特。由 $M_0 M_1 M_1 M_1$ 序列，证实收到了 C 控制比特的正确帧。最后，将这四个 DS1 比特流进行适当分路并去掉填充比特。

上述信号格式有两个保护措施：

1. 某个输入 DS1 比特流可能会包含与 $F_0 F_1 F_0 F_1 F_0 F_1 F_0 F_1$ 相同的序列。虽然这种情况发生的可能极小,但是一旦发生,就会使接收机锁定在错误序列上。而 $M_0 M_1 M_1 M_1$ 的出现将验证 $F_0 F_1 F_0 F_1 F_0 F_1 F_0 F_1$ 序列,由此保证了这四个 DS1 比特流的正确分路。
 2. C 控制比特纠正单个错误的能力确保了在四个 DS1 比特流中能够正确地去掉填充比特。

M12 复接器适应输入数据速率微小变化的能力, 可由图 3.21 的格式计算得到。将包含一个 $M_0 M_1 M_2 M_3$ 周期的比特间隔定义为 M 帧, 在每个 M 帧里, 向四个输入 DS1 比特流的每一个中填充 1 比特。则在每个 M 帧中, 这样的信号共有 $12 \times 6 \times 4 = 288$ 个位置可用于比特填充。T1 比特流的比特率等于 1.544 Mbps。因此输入比特率可增加

$$1.544 \times 10^3 \times \frac{1}{288} = 5.4 \text{ Kbps}$$

该结果远远超出了输入 DS1 比特流速率的预计变化量。因此，对每个输入信道，在每个帧里填充 1 比特，完全能够适应输入信号速率可能发生的变化量。

本地时钟决定了输出比特率,也决定了填充速率 S 。填充速率即在每个信道中任一帧里填充的平均比特数。M12 复接器是按 $S = 1/3$ 设计的。因而 DS2 比特流的标称比特速率等于

$$1.544 \times 4 \times \frac{49}{48} \times \frac{288}{288-S} = 6.312 \text{ Mbps}$$

这也保证了TMS2时钟频率是8kHz(语音信号的抽样速率)的整数倍,这正是我们所希望得到的。

3.11 PCM 的优点、局限性和改进

总地看来,PCM是传输承载语音和视频信号等模拟信号的最好调制方法。PCM的优点可追溯到使用编码脉冲作为模拟信号的数字表示,正是这个特点使PCM区别于其他模拟调制方法。可将PCM的重要优点归纳如下:

1. 抗信道噪声和抗干扰的能力很强。
 2. 在传输路径上能有效地再生编码信号。
 3. 可有效地用信道带宽来换取信噪比的提高,这种交换遵循指数律。
 4. 将各种基带信号按统一的格式进行传输,因此 PCM 可与其他数字数据形式共存于同一网络。
 5. 在 TDM 系统中取出或再插入消息源比较容易。

6. 通过采用特殊的调制方法或加密以达到安全通信的目的。附录5中讨论了数据的加密和解密。

然而,这些优点是以增加系统的复杂度和牺牲信道带宽为代价得到的。下面就讨论复杂度和带宽问题。

尽管采用PCM会引入许多复杂的操作,但这些操作都可以通过商用或民用VLSI(超大规模集成)芯片低成本地实现。也就是说,实现PCM系统所需的器件技术已经具备了。此外,随着VLSI技术的不断进步,PCM在模拟信号数字传输中也必将得到更加广泛的应用。

如果必须采用简单的调制方法,则可用增量调制(DM)代替PCM。在DM中,有意对基带信号进行“过抽样”,是为了采用简单的量化方法来构造编码信号。DM将在3.12节中进行讨论。

下面转向带宽问题的讨论。PCM带宽需求的增加是过去人们关心的一个问题,但如今它已经不再是人们真正担心的问题了,原因有两个。首先,越来越多的宽带通信信道的出现,意味着带宽不再是系统的约束条件了。用于广播的通信卫星和用于网络的光纤大大地解放了带宽的限制。与通信信道相关的概念在“背景与预览”中曾经提到过。

其次,通过采用复杂的数据压缩技术,实际上能够去掉PCM信号中的冗余量,从而在系统性能不发生严重劣化的前提下,降低了传输数据的比特率。事实上,处理过程复杂度的增加(因而增加了造价)是为了降低比特率,从而降低对带宽的要求。降低比特率的目的主要是为了保证较低容量无线信道的安全通信。

3.12 增量调制

在增量调制(DM)^[8](DM)中,对输入消息信号进行了过抽样(即速率远高于奈奎斯特速率),其目的是为了增加信号相邻样值之间的相关性。

在DM的基本形式里,对消息信号过抽样后的形式给出了阶梯近似,如图3.22(a)所示。将输入信号与这种近似信号之间的差值量化为两个量化级,用 $\pm\Delta$ 表示正差值和负差值。如果在抽样周期内,近似曲线位于信号以下,则将近似曲线增加一个台阶 Δ 。如果近似曲线位于信号以上,则将近似曲线下降一个台阶 Δ 。假如信号从一个样值到另一个样值的变化不是太快的话,阶梯近似曲线将保持在输入信号偏差 $\pm\Delta$ 的范围内。

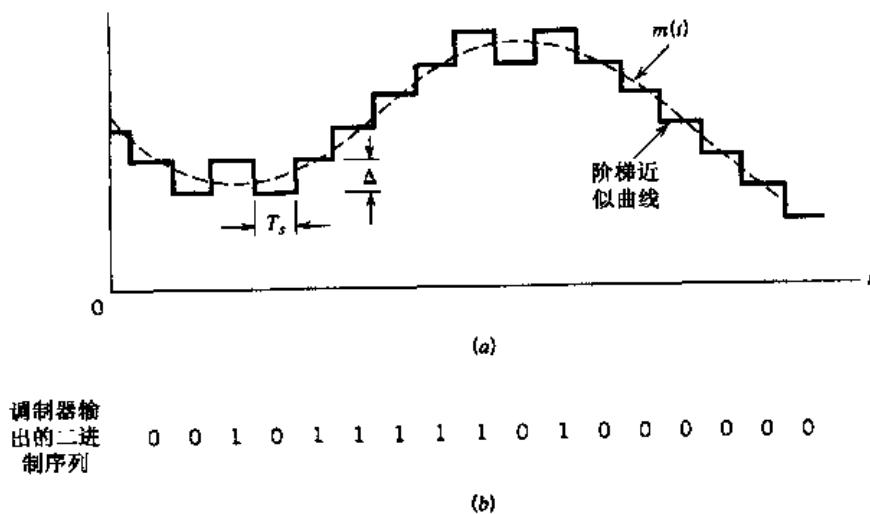


图3.22 DM的示意图

用 $m(t)$ 表示输入(消息)信号, $m_q(t)$ 表示其阶梯近似。为了便于表达, 采用数字信号处理文献中的习惯表示法:

$$m[n] = m(nT_s), \quad n = 0, \pm 1, \pm 2, \dots$$

其中, T_s 为抽样周期, $m(nT_s)$ 为信号 $m(t)$ 在时刻 $t = nT_s$ 的抽样值(对于其他连续时间信号的抽样也是如此)。

接着, 可用下列离散时间关系式描述 DM 的基本规律:

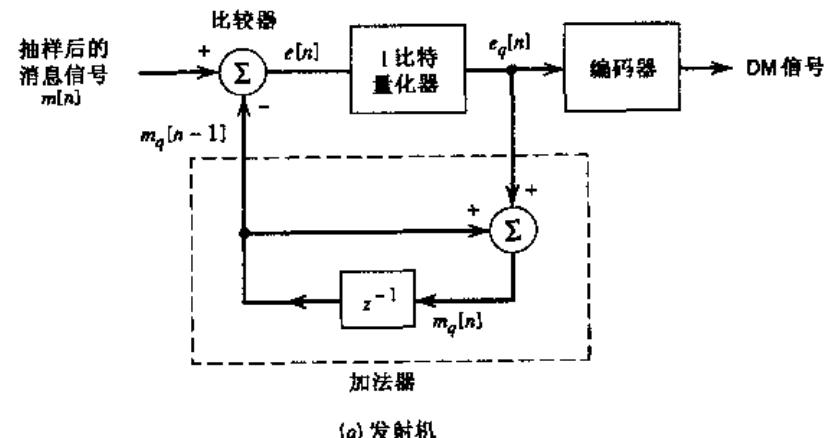
$$e[n] = m[n] - m_q[n - 1] \quad (3.52)$$

$$e_q[n] = \Delta \operatorname{sgn}(e[n]) \quad (3.53)$$

$$m_q[n] = m_q[n - 1] + e_q[n] \quad (3.54)$$

其中, $e[n]$ 为误差信号, 它表示此刻输入信号的抽样值 $m[n]$ 与前一时刻的近似值 $m_q[n - 1]$ 之差, $e_q[n]$ 为 $e[n]$ 量化后的形式, $\operatorname{sgn}(\cdot)$ 为符号函数。最后, 对量化器的输出 $m_q[n]$ 进行编码, 产生 DM 信号。

图 3.22(a) 说明了阶梯近似曲线 $m_q(t)$ 是怎样根据式(3.52)~式(3.54), 随输入信号 $m(t)$ 的变化而变化的, 而图 3.22(b) 则给出相应二进制序列的解调和输出。容易看出, 在 DM 系统中, 信息传输速率等于抽样速率 $f_s = 1/T_s$ 。



(a) 发射机

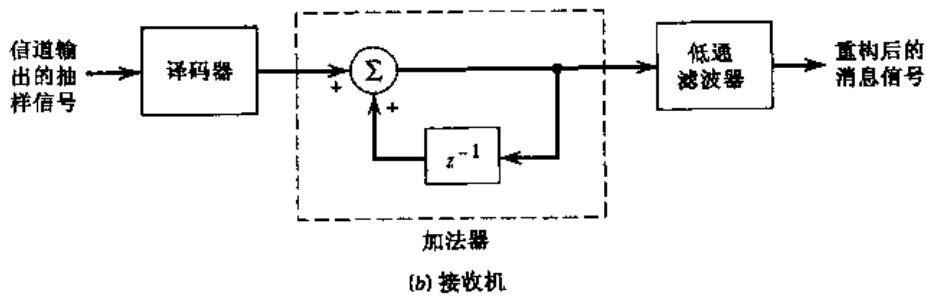


图 3.23 DM 系统

DM 的主要优点就是实现简单。将抽样后的消息信号送入由比较器、量化器和加法器级联而成的调制器中, 就可以产生 DM 信号, 如图 3.23(a)。加法器中标有 z^{-1} 字样的方框表示单位时延, 即时延等于一个抽样周期(变量 z 普遍用于 z 变换中。 z 变换是离散时间和离散系

统的基本分析方法)。调制器的原理是直接由式(3.52)~式(3.54)得到的。比较器用于计算两个输入信号的差值。量化器由硬限幅器组成,其输入-输出关系为,信号取符号函数,再与一个比例因子相乘,即得到输出。然后将量化器输出加在加法器上,得到结果为

$$m_q[n] = \Delta \sum_{i=1}^n \operatorname{sgn}(e[i]) = \sum_{i=1}^n e_q[i] \quad (3.55)$$

通过对式(3.53)和式(3.54)求解 $m_q[n]$ 得到了式(3.55)的结果。因此,在抽样时刻 nT_s , 加法器为近似曲线提供一个增量 Δ , Δ 的方向可能为正,也可能为负,这要由误差样值 $e[n]$ 的代数符号决定。如果输入样值 $m[n]$ 大于前一时刻的近似值 $m_q[n]$, 近似值就增加一个正的台阶 $+\Delta$ 。反之,则近似值增加一个负的台阶 $-\Delta$ 。这种方式使得加法器每次都能通过一个台阶(幅度为 $+\Delta$ 或 $-\Delta$)较好地跟踪输入样值。在图 3.23(b)所示的接收机中,译码器输出的正、负脉冲序列通过(与发射机中加法器原理相似的)加法器后,可重构出阶梯近似值 $m_q(t)$ 。然后将 $m_q(t)$ 送入图 3.23(b)所示的低通滤波器,则叠加在阶梯信号 $m_q(t)$ 高频部分的带外量化噪声被低通滤波器滤除,该滤波器的带宽等于原始消息带宽。

DM 中存在两种量化误差:斜率过载失真和颗粒噪声。首先讨论斜率过载失真。

式(3.54)表示的是正、负幅度增量 Δ 的累积,因此可视为积分的一种数字等效形式。也可以用 $q[n]$ 表示量化误差,即

$$m_q[n] = m[n] + q[n] \quad (3.56)$$

由式(3.52)可得到量化器的输入为

$$e[n] = m[n] - m[n - 1] - q[n - 1] \quad (3.57)$$

由此可见,如果不考虑量化误差 $q[n - 1]$,量化器的输入就等于输入信号的一阶后向差分,这在数值上近似等于输入信号的导数,也就是积分的逆过程。考虑原始输入信号 $m(t)$ 斜率的最大值。显然,在 $m(t)$ 斜率最大的区间内,为使样值序列 $\{m_q[n]\}$ 与输入样值序列 $\{m[n]\}$ 增加得一样快,应满足以下条件:

$$\frac{\Delta}{T_s} \geq \max \left| \frac{dm(t)}{dt} \right| \quad (3.58)$$

否则,若台阶 Δ 太小,当阶梯近似曲线 $m_q(t)$ 跟随输入波形 $m(t)$ 的陡峭部分变化时, $m_q(t)$ 就无法跟上 $m(t)$ 的变化了,如图 3.24 所示。这种情况称为斜率过载,由此带来的量化误差称为斜率过载失真(噪声)。由于阶梯近似曲线 $m_q(t)$ 的最大斜率是由台阶 Δ 的大小决定的,因此, $m_q(t)$ 会沿直线增加或减小。正是由于这个原因,采用固定台阶大小的增量调制器常称为线性增量调制器。

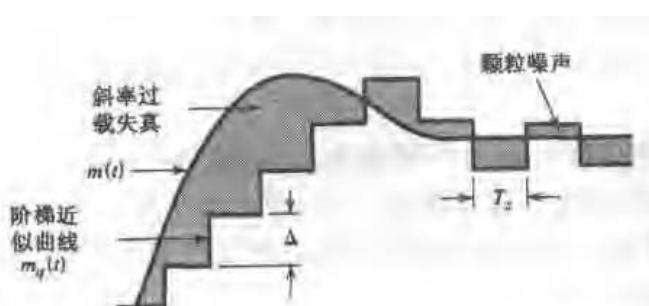


图 3.24 DM 中两种不同形式量化误差的示意图

与斜率过载失真相反,颗粒噪声发生在台阶 Δ 相对于输入波形 $m(t)$ 的局部斜率特性来说过大时。颗粒噪声导致阶梯近似曲线 $m_q(t)$ 围绕输入波形相对平坦的部分上下波动,如图3.24所示。颗粒噪声与PCM系统中的量化噪声类似。

由此可见,需要用大的台阶去适应较宽的动态范围,而用小的台阶来精确表示相对低电平的信号。因此,应为选择最优尺寸的台阶,使线性增量调制器中量化误差的均方值达到最小,从而得到斜率过载失真和颗粒噪声的折中结果。为达到此目的,应将增量调制器改为“自适应的”,也就是说,应当使台阶尺寸可随输入信号的变化而变化。这个问题将在3.16节计算机实验中进行进一步讨论。

3.12.1 增量 – 总和调制

前面提到,在传统形式的DM中,量化器的输入可近似看成输入消息信号的导数。这个特性给DM带来了一个缺点,即传输中的干扰(如噪声)会在解调信号中形成误差积累。在增量调制前,先对消息信号进行积分,就可以克服这个缺点。这种方法的优点是:

- 预加重了输入信号的低频分量。
- 增大了DM中输入相邻样值的相关性,从而降低了量化器输入端误差信号的方差,提高了系统性能。
- 简化了接收机的设计。

输入端包含了积分过程的增量调制称为增量 – 总和调制(D – ΣM)^[9]。准确地说,应称为总和 – 增量调制,因为积分过程实际是在增量调制之前进行的。但增量 – 总和调制是在文献中普遍使用的一个术语,因此本书仍沿用这个名称。

图3.25(a)是增量 – 总和调制系统的框图。在这个框图中,消息信号 $m(t)$ 为连续时间形式,说明此时的脉冲调制器是由一个硬限幅器后跟一个乘法器组成。乘法器的另一个输入分量是由外部脉冲发生器(时钟)提供的1比特编码信号。由于发射机输入端采用了积分,在接收机中就必须存在相反的信号处理过程,即微分。但是,由于传统的DM接收机中存在与之相抵消的积分过程,因此在接收端不再需要微分过程。因此,增量 – 总和调制系统的接收机仅由低通滤波器构成,如图3.25(a)所示。

积分本质上是一个线性运算。因此可将图3.25(a)中发射端的两个积分器1和2合并为一个积分器,并置于比较器之后,如图3.25(b)所示,这样就简化了发射机的设计。这种形式的增量 – 总和调制系统不但比图3.25(a)所示的系统更为简单,同时还对1比特PCM的“平滑”形式——增量 – 总和调制做出了解释:术语“平滑”是指比较器的输出在量化前先经过了积分;术语“1比特”只是再次强调量化器是由硬限幅器组成的,只有两个量化级。

在DM中,通过采用远高于PCM所需速率的抽样速率,简化了发射机和接收机的实现。简化的代价是传输带宽与信道带宽的增大。但在很多情况下,信道带宽是既定的,这就与DM的要求相矛盾。这种情况下,可以考虑通过增加系统复杂度来减小信道带宽。这可以通过预测来实现。下面讨论线性预测。

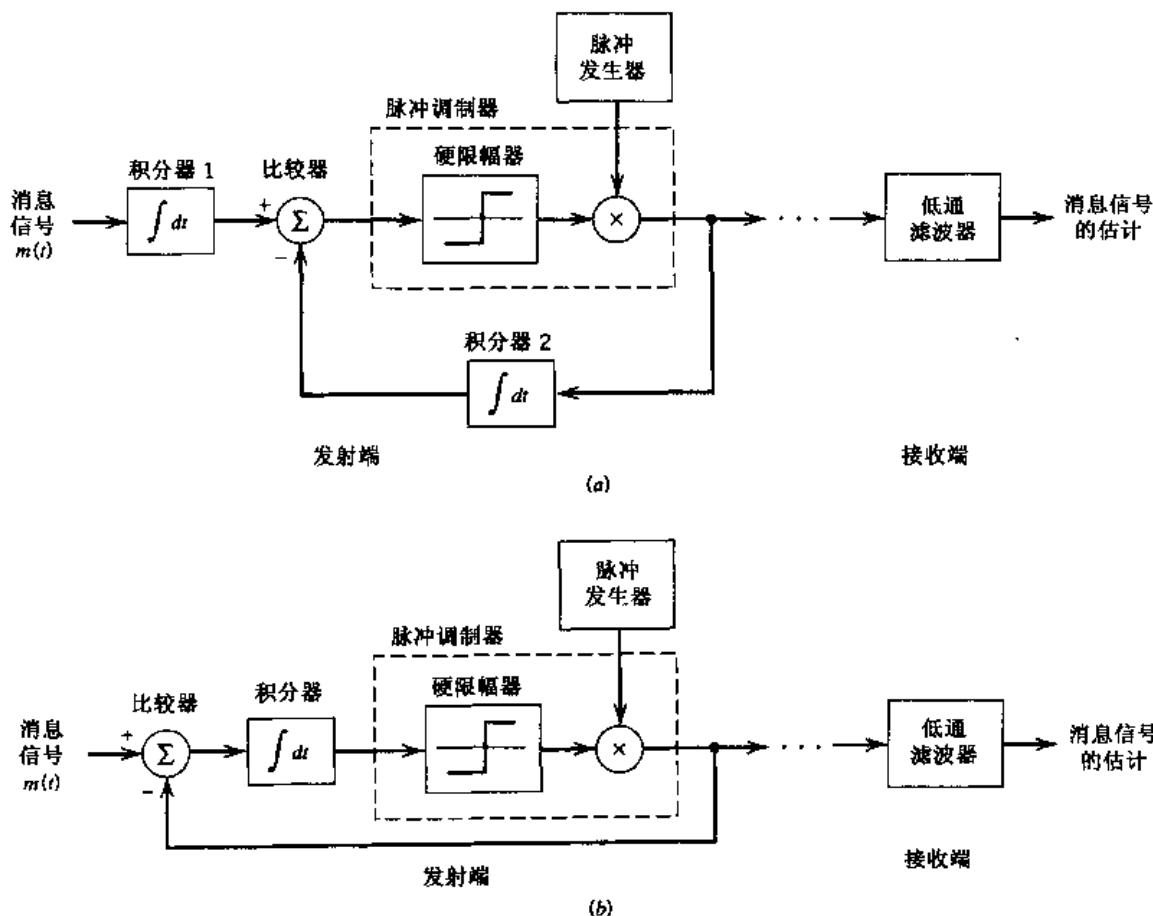


图 3.25 两种等价形式的增量 - 总和调制系统

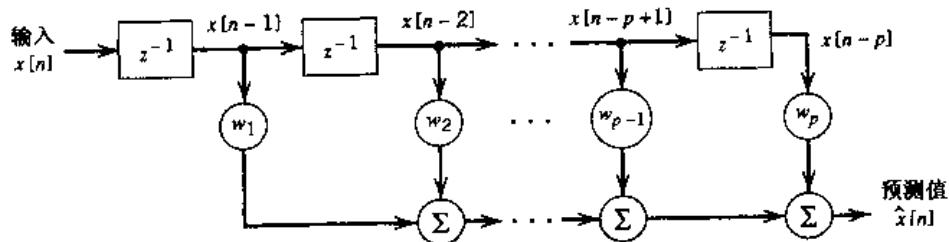
3.13 线性预测

考虑图 3.26 中的有限长冲激响应(FIR)离散时间滤波器,此滤波器包含三个功能块:

1. p 个单位延迟单元的集合,其中,每个单元用 z^{-1} 表示。
2. 滤波系数为 w_1, w_2, \dots, w_p 的乘法器集合。
3. “加法器”的集合,用于对添加了比例因子并经过延迟的输入 $x[n-1], x[n-2], \dots, x[n-p]$ 进行求和,以产生输出 $\hat{x}[n]$ 。因此,滤波器的输出 $\hat{x}[n]$ (更准确地说是输入的线性预测)可用以下卷积和进行定义:

$$\hat{x}[n] = \sum_{k=1}^p w_k x[n-k] \quad (3.59)$$

其中,单位延迟单元总数 p 称为预测阶数。

图 3.26 p 阶线性预测滤波器框图

时刻 nT_s 的实际抽样值为 $x[n]$ 。用 $e[n]$ 表示预测误差, 即 $x[n]$ 和预测值 $\hat{x}[n]$ 之间的差值, 则

$$e[n] = x[n] - \hat{x}[n] \quad (3.60)$$

设计目标是选择适当的滤波器系数 w_1, w_2, \dots, w_p , 使性能指数的 J 达到最小值, 性能指数的定义为均方误差

$$J = E[e^2[n]] \quad (3.61)$$

将式(3.59)和式(3.60)代入式(3.61), 然后展开, 则性能指数的公式变为

$$J = E[x^2[n]] - 2 \sum_{k=1}^p w_k E[x[n]x[n-k]] + \sum_{j=1}^p \sum_{k=1}^p w_j w_k E[x[n-j]x[n-k]] \quad (3.62)$$

假定输入信号 $x(t)$ 是零均值平稳过程 $X(t)$ 的抽样函数, 即对于所有 n 值, $E[x(t)]$ 均为 0。定义

$$\begin{aligned} \sigma_X^2 &= X(t) \text{ 在时刻 } nT_s \text{ 的采样函数方差} \\ &= E[x^2[n]] - (E[x[n]])^2 \\ &= E[x^2[n]] \end{aligned}$$

$$\begin{aligned} R_X(kT_s) &= X(t) \text{ 的时延为 } kT_s \text{ 的自相关函数} \\ &= R_X[k] \\ &= E[x[n]x[n-k]] \end{aligned}$$

从而可将式(3.62)简化为

$$J = \sigma_X^2 - 2 \sum_{k=1}^p w_k R_X[k] + \sum_{j=1}^p \sum_{k=1}^p w_j w_k R_X[k-j] \quad (3.63)$$

将性能指数 J 对滤波系数 w_k 求导, 并令结果等于 0, 然后对各项进行重新组合, 可得

$$\sum_{j=1}^p w_j R_X[k-j] = R_X[k] = R_X[-k], \quad k = 1, 2, \dots, p \quad (3.64)$$

最优式(3.64)称为线性预测的 Wiener-Hopf 等式。

将 Wiener-Hopf 等式(3.64)改写成矩阵形式将更为方便。令

$\mathbf{w}_o = p \times 1$ 的最优系数矢量

$$= [w_1, w_2, \dots, w_p]^T$$

$\mathbf{r}_X = p \times 1$ 的自相关系数矢量

$$= [R_X[1], R_X[2], \dots, R_X[p]]^T$$

$\mathbf{R}_X = p \times p$ 的自相关系数矩阵

$$= \begin{bmatrix} R_X[0] & R_X[1] & \cdots & R_X[p-1] \\ R_X[1] & R_X[0] & \cdots & R_X[p-2] \\ \vdots & \vdots & & \vdots \\ R_X[p-1] & R_X[p-2] & \cdots & R_X[0] \end{bmatrix}$$

则可将式(3.64)的集合简记为

$$\mathbf{R}_X \mathbf{w}_o = \mathbf{r}_X \quad (3.65)$$

假定自相关矩阵 \mathbf{R}_X 是非奇异的, 即存在逆矩阵。将式(3.65)两边同时乘以逆矩阵 \mathbf{R}_X^{-1} , 求出

系数矢量 \mathbf{w}_o 的最优解

$$\mathbf{w}_o = \mathbf{R}_x^{-1} \mathbf{r}_x \quad (3.66)$$

注意,自相关矩阵 \mathbf{R}_x 主对角线上的所有元素均为 $R_x[0] = \sigma_x^2$,而与主对角线平行的其他对角线上的元素也各自相等。具有这种性质的方阵称为特普利茨(Toeplitz)方阵,这是假定输入信号 $x(t)$ 来自平稳过程的直接结果。特普利茨性质的实际含义是,自相关矩阵 \mathbf{R}_x 由自相关值 $R_x[0], R_x[1], \dots, R_x[p-1]$ 唯一确定。自相关矢量 \mathbf{r}_x 由自相关值 $R_x[1], R_x[2], \dots, R_x[p]$ 定义。因而线性最优预测器的 p 个滤波器系数由方差 $\sigma_x^2 = R_x[0]$ 和 $X(t)$ 的自相关函数在 $T_s, 2T_s, \dots, pT_s$ 时刻的 p 个值唯一确定。

将式(3.64)代入式(3.63)可得预测误差的最小均方值(已经过简化):

$$J_{\min} = \sigma_x^2 - \mathbf{r}_x^T \mathbf{R}_x^{-1} \mathbf{r}_x \quad (3.67)$$

二次项 $\mathbf{r}_x^T \mathbf{R}_x^{-1} \mathbf{r}_x$ 通常为正值。因此,由式(3.67)定义的最优线性预测器的均方误差 J_{\min} 通常小于被预测输入样值的方差 σ_x^2 。

3.13.1 线性自适应预测

用式(3.66)计算线性预测器的加权矢量时,需要知道输入序列 $|x[n]|$ 在 $k=0, 1, \dots, p$ 时的自相关函数值 $R_x[k]$,其中 p 为预测阶数。但是,很多情况下,我们无法确定全部 k 值对应的 $R_x[k]$ 。这种情况下可采用自适应预测器。

在下面的前提下,预测器是自适应的:

- 给定任意的时延加权初值,由“递归”方式计算时延加权值 w_k ,其中 $k=1, 2, \dots, p$ 。
- 用于调节时延加权的算法(反复进行)是“自设计的”,只对可用数据进行操作。

算法的目的是找出误差抛物曲面的最小点,该曲面描述了代价函数 J 与时延加权值之间的依赖关系。因此,可沿着误差曲面下降最快的方向,相继调节预测器的时延加权值,也就是说,应按与梯度矢量相反的方向进行调节。梯度矢量的定义如下:

$$g_k = \frac{\partial J}{\partial w_k}, \quad k = 1, 2, \dots, p \quad (3.68)$$

这实际上是最佳梯度法的思想。令 $w_k[n]$ 表示第 n 次迭代时第 k 个时延的加权值。那么在第 $n+1$ 次迭代时这个加权值变为

$$w_k[n+1] = w_k[n] - \frac{1}{2} \mu g_k, \quad k = 1, 2, \dots, p \quad (3.69)$$

其中, μ 为步长参数,用来控制调节的速度,为了后面表达的方便,引入了因子 $1/2$ 。将式(3.63)中的代价函数 J 对 w_k 求导,易得

$$\begin{aligned} g_k &= -2R_x[k] + 2 \sum_{j=1}^p w_j R_x[k-j] \\ &= -2E[x[n]x[n-k]] + 2 \sum_{j=1}^p w_j E[x[n-j]x[n-k]], \quad k = 1, 2, \dots, p \end{aligned} \quad (3.70)$$

用瞬时值作为自相关函数 $R_x[k]$ 和 $R_x[k-j]$ 的估计值,可将求解 g_k 的公式进一步简化。即

忽略式(3.70)中的期望运算以简化自适应过程。因此,可将第 n 次迭代时 g_k 的相应估计值记为

$$\hat{g}_k[n] = -2x[n]x[n-k] + 2 \sum_{j=1}^p w_j[n]x[n-j]x[n-k], \quad k = 1, 2, \dots, p \quad (3.71)$$

注意,由平稳过程得到的输入 $x[n]$,其梯度 g_k 是个确定量,而估计值 $\hat{g}_k[n]$ 是随机变量的一个样值。

将式(3.71)代入式(3.69)并提出公因子 $x[n-k]$,则有

$$\begin{aligned} \hat{w}_k[n+1] &= \hat{w}_k[n] + \mu x[n-k] \left(x[n] - \sum_{j=1}^p \hat{w}_j[n]x[n-j] \right) \\ &= \hat{w}_k[n] + \mu x[n-k]e[n], \quad k = 1, 2, \dots, p \end{aligned} \quad (3.72)$$

其中 $e[n]$ 为预测误差,其定义如下:

$$e[n] = x[n] - \sum_{j=1}^p \hat{w}_j[n]x[n-j] \quad (3.73)$$

在式(3.72)和式(3.73)里,将第 k 个时延加权值的估计值用 \hat{w}_k 表示,以便与实际值 w_k 区分开来。需注意在计算预测器时延加权值的递归调整量时, $x[n]$ 相当于“目标响应”。

式(3.72)和式(3.73)构成了线性自适应预测的最小均方(LMS)算法,其运算过程如图 3.27 所示。该自适应滤波算法之所以具有普遍性,是因为它易于实现。算法的计算复杂度由加法和乘法的数量决定,并与预测阶数 p 呈线性关系。

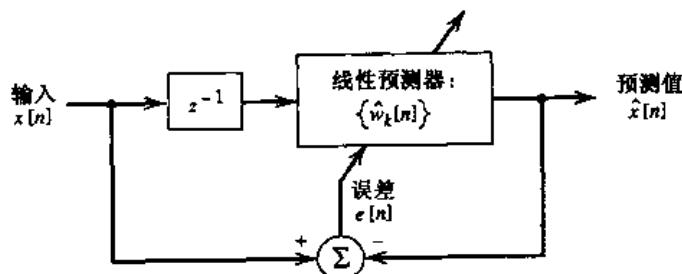


图 3.27 线性自适应预测过程框图

LMS 算法是一种随机自适应滤波算法。随机的含义是指由 $\{w_k[0]\}_{k=1}^p$ 定义的初始条件开始计算,沿着一种“之”字形路径去寻找误差平面的最低点,但却永远不会准确地找到这个最低点。一旦建立了稳态条件,预测点就以随机方式落在误差平面最低点的附近。

了解了线性预测,我们就可以进一步讨论怎样在实际中提高 PCM 的性能了。

3.14 差分脉冲编码调制

当在 PCM 中以稍高于奈奎斯特速率的速率对语音或视频信号进行抽样时,抽样信号的相邻抽样值之间存在着较高的相关性。高相关性的含义是指,平均来说,信号不会从一个样值迅速变到下一个样值,其结果是相邻样值差值的方差小于信号本身的方差。当对这些高相关性的样值进行编码时(例如在标准 PCM 系统中),编码信号将包含冗余信息。这说明编码过程产

生了对于信息的传输来说并不必要的符号。在编码之前,先去掉冗余信息,从而得到更高效的编码信号,就是差分脉冲编码调制(DPCM)的基本思想。

如果已知一个信号在过去某个时间点之前的变化规律,则可用3.13节讲述的预测方法来得到信号未来值的一个估计。假定以速率 $f_s = 1/T_s$ 对基带信号 $m(t)$ 进行抽样,得到样值间隔为 T_s 秒的序列 $\{m[n]\}$ 。信号 $m(t)$ 未来值的预测可通过图3.28(a)所示的差分量化方法来实现。量化器的输入信号定义如下:

$$e[n] = m[n] - \hat{m}[n] \quad (3.74)$$

它是未经量化的输入样值 $m[n]$ 与其预测值 $\hat{m}[n]$ 之差。预测值由线性预测滤波器产生。该滤波器的输入即为输入样值 $m[n]$ 的量化值。差值信号 $e[n]$ 就是预测滤波器未能正确预测输入值所造成的预测误差。如图3.28(a)所示,对量化器输出进行编码,即得到PCM的变体,称为DPCM^[10]。

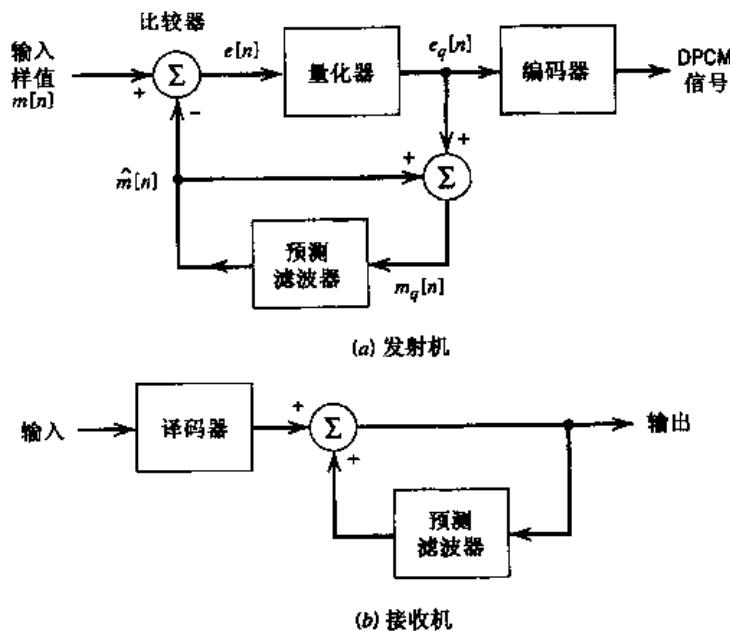


图3.28 DPCM系统

可将量化器输出表示为

$$e_q[n] = e[n] + q[n] \quad (3.75)$$

其中, $q[n]$ 为量化误差。如图3.28(a)所示,预测值 $\hat{m}[n]$ 与量化器输出 $e_q[n]$ 之和构成了预测滤波器的输入信号

$$m_q[n] = \hat{m}[n] + e_q[n] \quad (3.76)$$

将式(3.75)代入式(3.76),得

$$m_q[n] = \hat{m}[n] + e[n] + q[n] \quad (3.77)$$

由式(3.74)可看出,和项 $\hat{m}[n] + e[n]$ 正好等于输入样值 $m[n]$ 。因此,可将式(3.77)简化为

$$m_q[n] = m[n] + q[n] \quad (3.78)$$

此式是输入样值 $m[n]$ 的量化形式。即无论预测滤波器的性质如何,预测滤波器的输入量化

样值 $m_q[n]$ 与原始输入样值 $m[n]$ 之间总是存在量化误差 $q[n]$ 。如果预测滤波器性能很好，则预测误差 $e[n]$ 的方差将小于输入样值 $m[n]$ 的方差。所以，对于给定量化级数的量化器，可对其进行调节，使之产生的量化误差的方差小于输入样值 $m[n]$ 在标准 PCM 系统中直接进行量化所产生的误差的方差。

重构量化输入信号的接收机如图 3.28(b) 所示。图中的译码器用于重构量化误差信号。通过在译码器输出端采用与图 3.28(a) 所示发射机中相同的预测滤波器，可重构出原始输入信号的量化形式。当不存在信道噪声时，接收机的输入与发射机的输出应完全一致。此时，相应的接收机输出信号应等于 $m_q[n]$ ，它与原始输入 $m[n]$ 只相差一个量化误差 $q[n]$ ，这是对预测误差 $e[n]$ 进行量化的结果。

由前面的分析可知，在无噪声环境下，发射机和接收机中的预测滤波器处理的是相同的样值序列 $m_q[n]$ 。之所以在发射机中的量化器上增加一个如图 3.28(a) 所示的反馈路径，就是为了达到这个目的。

DM 是 DPCM(差分脉冲编码调制)的特例。比较图 3.28 所示的 DPCM 系统和图 3.23 所示的 DM 系统，可发现两者基本相似，但存在两个重要的差别：在 DM 中采用的是 1 比特(两级)量化器，并用一个延迟单元(即零阶预测)代替了 DPCM 中的预测滤波器。简单地说，DM 就是 DPCM 的 1 比特形式。与标准 PCM 系统不同，DPCM 和 DM 的发射机中均采用了反馈。

与 DM 一样，当输入信号变化太快以至于预测滤波器很难对其进行跟踪时，DPCM 也将发生斜率过载失真。同 PCM 一样，DPCM 也存在量化噪声。

3.14.1 处理增益

图 3.28 所示 DPCM 系统的输出信噪比为

$$(SNR)_O = \frac{\sigma_M^2}{\sigma_Q^2} \quad (3.79)$$

其中， σ_M^2 为原始输入样值 $m[n]$ 的方差，假定 $m[n]$ 均值为零， σ_Q^2 为量化误差 $q[n]$ 的方差。可将式(3.79)写成两个因式的乘积，即

$$\begin{aligned} (SNR)_O &= \left(\frac{\sigma_M^2}{\sigma_E^2} \right) \left(\frac{\sigma_E^2}{\sigma_Q^2} \right) \\ &= G_p (SNR)_Q \end{aligned} \quad (3.80)$$

其中， σ_E^2 为预测误差的方差。因子 $(SNR)_Q$ 为信号-量化噪声比，其定义如下：

$$(SNR)_Q = \frac{\sigma_E^2}{\sigma_Q^2} \quad (3.81)$$

另一个因子 G_p 为由差分量化方法得到的处理增益，其定义如下：

$$G_p = \frac{\sigma_M^2}{\sigma_E^2} \quad (3.82)$$

当 G_p 大于 1 时，表示由于采用图 3.82 的差分量化方法，给信噪比带来了增益。对于给定的基本(消息)信号，其方差 σ_M^2 是固定的，所以，只有将预测误差 $e[n]$ 的方差 σ_E^2 最小化，才能使 G_p 最大化。此时的目标变为设计合适的预测滤波器以最小化 σ_E^2 。

在传输语音信号的情况下,DPCM 的最优信号-量化噪声比可比标准 PCM 高出 4~11 dB。最大的改善量发生在从未采用预测到采用一阶预测的过程中。增加预测滤波器的阶数至四阶或五阶,仍会带来附加的增益量,此后就几乎得不到附加的增益量了。由于量化噪声的 6 dB 等价于每样值 1 比特,见式(3.35),因此,在信号-量化噪声比恒定的情况下,假定抽样速率等于 8 kHz,则 DPCM 要比标准 PCM 节省大约 8~16 Kbps(即每样值 1 到 2 比特)。

3.15 自适应差分脉冲编码调制

PCM 语音编码的标准速率为 64 Kbps,传输这样的信号要求信道带宽很高。然而,在某些应用中,例如低容量无线信道中的安全传输,其信道带宽是有限的。在此类应用中,需要在保证可接受的逼真度或再生质量的前提下,以低比特率进行语音编码。

为了以低比特率进行语音编码,可利用语音波形的统计特性和听觉的性质来优化已给出的编码波形。设计目的主要有两点:

1. 尽可能去掉语音信号中的冗余信息。
2. 以有效的方式将可用比特分配给语音信号的非冗余部分以进行编码。

随着比特率从 64 Kbps(用于标准 PCM 中)降到了 32 Kbps,16 Kbps,8 Kbps 和 4 Kbps,去掉冗余信息和分配比特的方法变得越来越复杂。经验表明,在 64 到 8 Kbps 的范围内,为了保证相同的语音质量,当比特率减半时,语音编码的计算复杂度(由加法和乘法运算的数量决定)就增加一个数量级。

本节介绍自适应差分脉冲编码调制(ADPCM)。^[1]这种调制方法采用自适应量化器和自适应预测器,能以 32 Kbps 的速率进行语音编码,即编码速率由标准 PCM 中的每样值 8 比特降到了 4 比特。这里的自适应是指预测器能对输入语音信号的电平和频谱的变化产生响应。由于讲话者、谈话内容以及语普通信过程中信号电平的不断变化,只有结合使用自适应量化和自适应预测的方法,才能在大多数情况下均达到最优的语音质量。

自适应量化是指量化器的步长 $\Delta[n]$ 是时变的。在任何由标识 n 表示的给定抽样瞬间,这里假定自适应量化器具有均匀的转移特性。为了与输入样值 $m[n]$ 的方差 σ_m^2 相匹配,步长 $\Delta[n]$ 的大小不断变化。即

$$\Delta[n] = \phi \hat{\sigma}_m[n] \quad (3.83)$$

其中, ϕ 为常量, $\hat{\sigma}_m[n]$ 为标准偏差 $\sigma_m[n]$ (即方差 σ_m^2 的平方根)的一个估计值。对于非平稳输入, $\sigma_m[n]$ 是时变的。根据式(3.83),自适应量化要解决的问题是连续地计算估计值 $\hat{\sigma}_m[n]$ 。

式(3.83)的实现可采用以下两种方法中的任一种:

1. 前向估计自适应量化(AQF),在这种方法中,由未经量化的输入信号样值向前推导 $\sigma_m[n]$ 的估计值。
2. 后向估计自适应量化(AQB),在这种方法中,由量化器输出的抽样值向后推导 $\sigma_m[n]$ 的估计值。

AQF 方法需用缓冲器存储(用于学习过程的)输入语音信号的未经量化的抽样值。同时,还需将电平信息(每个步长样值通常约为 5~6 比特)准确地传输到距离很远的译码器中。因此,系统必须传送附加的次要信息,这就加重了系统的负担。此外,由于采用 AQF 所带来的编码过

程中的处理时延(对于语音信号,这个时延为是 16 ms)在某些应用中是不允许的,AQB 避免了 AQF 本身固有的电平传输、缓冲和时延问题。在 AQB 中,系统从量化器输出的最近历史中提取信息,以计算步长 $\Delta[n]$ 。因此实际情况中多采用 AQB。

图 3.29 为 AQB 的框图,这是一个非线性反馈系统,系统的稳定趋势不很明显。如果量化器输入的 $m[n]$ 是有界的,接下来的后向估计 $\hat{m}[n]$ 和相应的步长 $\Delta[n]$ 也应当是有界的。在这样的条件下,系统实际上是有稳定的。

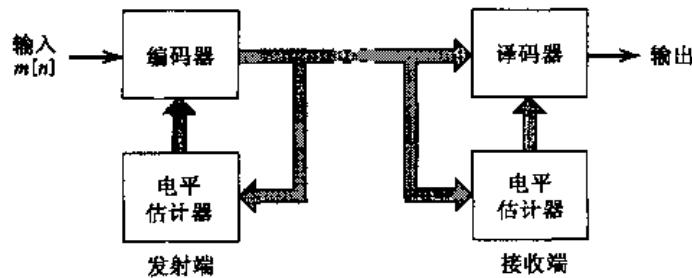


图 3.29 AQB 框图

之所以在 ADPCM 中采用自适应预测,是因为语音信号存在着固有的非平稳性,其表现是语音信号的自相关函数和功率密度谱均为自变量的时变函数。这说明对于这样的输入信号,预测器也应设计为时变的,即自适应的。与自适应量化一样,自适应预测也有两种预测方法:

1. 前向估计自适应预测(APF),用输入信号未经量化的样值去推导预测器系数的估计值。
2. 后向估计自适应预测(APB),用量化器输出的样值和预测误差去推导预测器系数的估计值。

APF 与 AQF 具有同样的缺点(即次要信息、缓冲和时延)。通过采用图 3.30 中的 APB 方法可避免这些缺点。图中标有“自适应预测逻辑”字样的方框表示更新预测器系数的算法。在 APB 中,最优预测器系数是以量化和传输的数据为基础进行估计的,因此,这些系数更新的频率与预计的相同,即从一个样值到另一个样值。所以,在 ADPCM 中多采用 APB 预测方法。

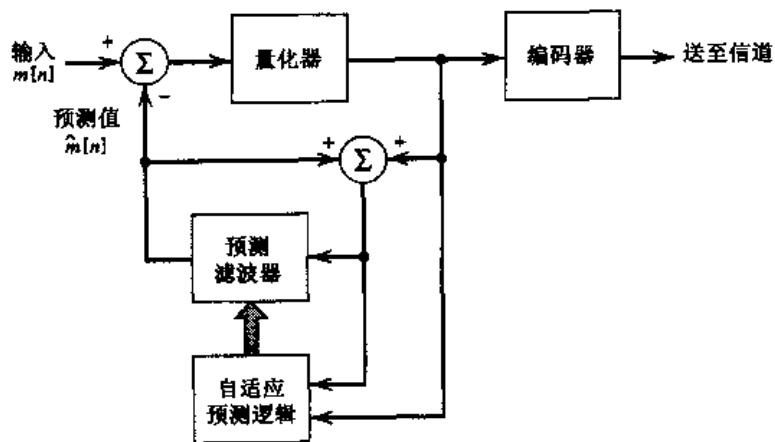


图 3.30 APB 框图

3.13 节描述的预测器最小均方算法,以及基于式(3.83)的量化器自适应方法,在发射机和接收机中均设计为以同步方式结合在一起。

这种结合使系统在 32 Kbps 时具有相当好的性能,正因如此,ADPCM 和 64 Kbps 的标准 PCM 共同成为了国际上语音信号的标准编码技术。

3.16 计算机实验:自适应增量调制

线性增量调制(LDM)应用了 AQB 的简单形式,形成了自适应增量调制(ADM)。所有 ADM 算法均包含两个法则:

1. 如果相继误差的极性相反,那么增量调制器工作在颗粒模式下,此时应减小步长。
2. 如果相继误差的极性相同,那么增量调制器工作在斜率过载模式下,此时应增加步长。

根据这两个法则去改变步长,增量调制器就能够处理输入信号的变化了。

图 3.31 中 ADM 框图的原理是,在自适应过程的每次迭代中,步长增加或减少的幅度均为 50%。调节步长的算法由下式给出:^[12]

$$\Delta[n] = \begin{cases} \frac{|\Delta[n-1]|}{m_q[n]} (m_q[n] + 0.5m_q[n-1]) & \Delta[n-1] \geq \Delta_{\min} \\ \Delta_{\min} & \Delta[n-1] < \Delta_{\min} \end{cases} \quad (3.84)$$

其中, $\Delta[n]$ 为算法在第 n 次迭代(时间阶数)时的步长, $m_q[n]$ 为 1 比特量化器的输出,其值等于 ± 1 。

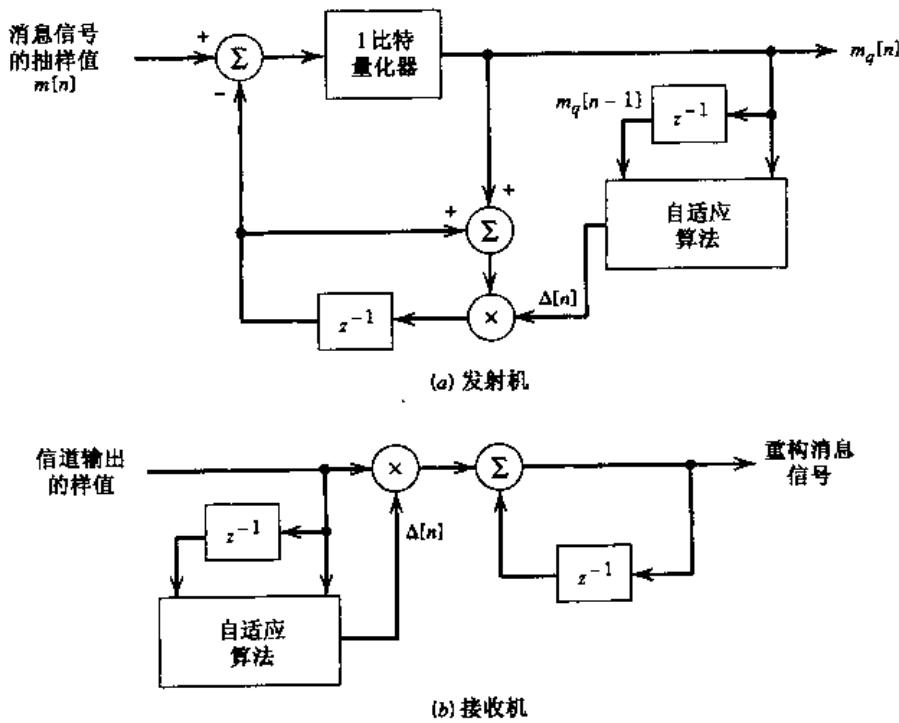


图 3.31 ADM 系统

本实验以正弦输入信号为例,说明基于式(3.84)的 ADM 算法的重构性能,并将它与相应 LDM 的性能进行比较。实验的详细内容如下:

输入信号:

$$m(t) = A \sin(2\pi f_m t)$$

其中,幅度 $A = 10$,频率 $f_m = f_s/100$, f_s 为抽样频率。

LDM:

$$\text{步长 } \Delta[n] = 1 \quad \text{对于所有 } n$$

ADM:

$$\Delta_{\min} = \frac{1}{8}$$

图 3.32 画出了实验结果。图 3.32(a)为 LDM 的结果,3.32(b)为 ADM 的结果。根据画出的波形,可得出以下几点结论:

- ADM 跟踪正弦输入信号变化的能力要远远强于 LDM。这是由于算法在相继迭代中调节了步长。

特别地,在输入信号末端,ADM 步长的减小使它的量化误差小于 LDM 的量化误差。

但是,在输入信号斜率较大的区间内,这两种调制方法产生的量化误差相差不大。

- 由于 ADM 的跟踪性能优于 LDM,因此 ADM 输出信号的平均比特率远低于 LDM。

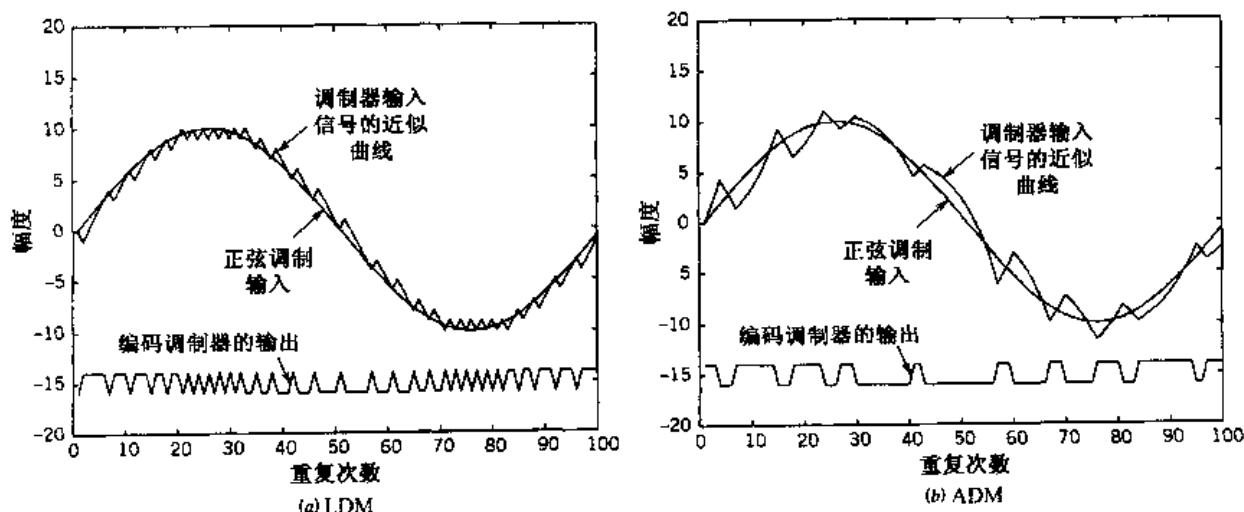


图 3.32 DM 的计算机实验波形

3.17 MPEG 音频编码标准

语音(话音)和音频信号的相似性在于,两者的编码方法的质量都要基于人类的听觉特性来进行衡量。由已有的语音生成模型,已得到关于语音信号的有效编码方法(例如 ADPCM)。但遗憾的是,在音频信号方面并没有获得类似的突破。

本节首先简短地回顾了“背景与预览”中提到的 MPEG-1/音频编码标准。MPEG 代表动态图像专家组,后缀 1 表示它是一系列标准^[13]中的第一个。与 ADPCM 类似,MPEG-1/音频编码标准也是一个有损压缩系统。但从实际应用的观点看,MPEG-1/音频编码标准是不同于 ADPCM 的。即 MPEG-1 标准对于高抽样速率的立体声音频信号来说,能够实现透明的、感觉无

损的压缩。MPEG/音频委员会在非常恶劣的收听条件下进行了主观收听测试。测试结果表明,即使在 6:1 的压缩率下,原始音频信号和编码后的信号在感觉上也并没有什么差别。

MPEG-1/音频编码标准之所以具有这样的性能,是由于其中利用了人类听觉系统的两个心理声学特点:

1. 临界频带

听觉系统的内耳^[14]将非线性范围内的输入信号的功率谱密度用有限的频带表示出来,该频带即称为临界频带。可听到的频带范围扩展至 20 kHz,被 25 个临界频带覆盖,其各自的带宽随频率的增大而增大。可以不严格地说,听觉系统好比一个带通滤波器集,它由 25 个频带相互重叠的带通滤波器组成,该“滤波器”最低可听频率小于 100 Hz,而最高可听频率达到 5 kHz。

2. 听觉屏蔽

听觉屏蔽或噪声屏蔽是一种频域现象,当低电平信号(被屏蔽者)和高电平信号(屏蔽者)同时出现,且二者在频率上足够接近时,就会产生这种现象。如果低电平信号处于屏蔽门限以下,那么当出现较强的信号时,就听不见低电平信号了(即被屏蔽掉了)。当两个信号位于相同的临界频带内时,最易出现这种听觉屏蔽现象。但当两个信号位于相邻临界频带内时,这种现象就不会很明显了。

图 3.33 说明了屏蔽门限的定义并给出了一对相邻频带的相关参数。其中假定屏蔽者(即高电平信号)位于深色阴影临界频带内。位于该深色区间且低于屏蔽门限的低电平信号,被强信号屏蔽了。由图 3.33 可见,屏蔽门限在临界频带上随频率变化。对于临界频带,可定义一个最小屏蔽门限,临界频带内低于此门限的所有低电平信号均被强信号屏蔽。用 dB 表示的屏蔽者和最小屏蔽门限的功率之差,称为信号-屏蔽比(SMR)。图 3.33 也给出了 R 比特量化器的信噪比。信号-屏蔽比和信噪比的差值即为 R 比特量化器的噪声-屏蔽比(NMR),由下式表示:

$$NMR = SMR - SNR \quad (3.85)$$

其中,所有项均用 dB 表示。在临界频带内,只要相关量化器的 NMR 为负,量化噪声就不会被听见。

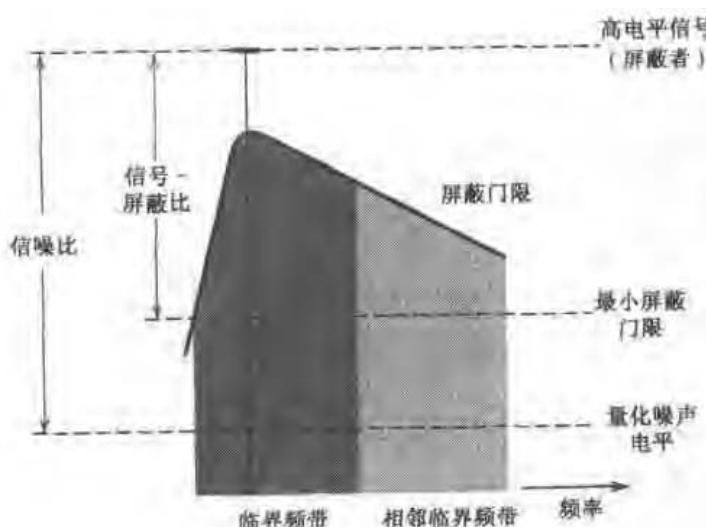


图 3.33 屏蔽门限及相关参数的示意图。高电平信号(屏蔽者)位于深色阴影表示的临界频带内,此频带内的屏蔽影响比浅色阴影内的相邻频带内更为明显

了解了听觉系统的心理声学背景,就能够理解 MPEG-1/音频编码标准的工作原理了。图 3.34 给出了编码器和译码器的基本框图。编码器由四个功能单元组成:时-频映射网络,心理声学模块,量化器和编码器,以及帧-封装单元。译码器由三个功能单元组成:帧-解封单元,频率样值重构网络以及频-时映射网络。只有编码器必须采用心理声学模块。

首先从编码器开始介绍,时-频映射网络的功能是将输入音频信号分解成编码使用的若干子带。映射分为三层,分别标为 I, II 和 III,它们的复杂度、时延和主观感觉性能依次递增。I 层的算法利用一个带通滤波器集将音频信号分为 32 个带宽恒定的子带,II 层和 III 层中也采用这样的滤波器集。根据前面对不等间隔排列的临界频带的讨论,这种滤波器集的设计应兼顾计算效率和感觉性能。II 层的算法只在 I 层算法上进行了简单改进,将数据编码为更大的群,从而提高了压缩性能。III 层的算法则更为周详,因为其设计目标是得到更接近于临界频带划分的频率分解。

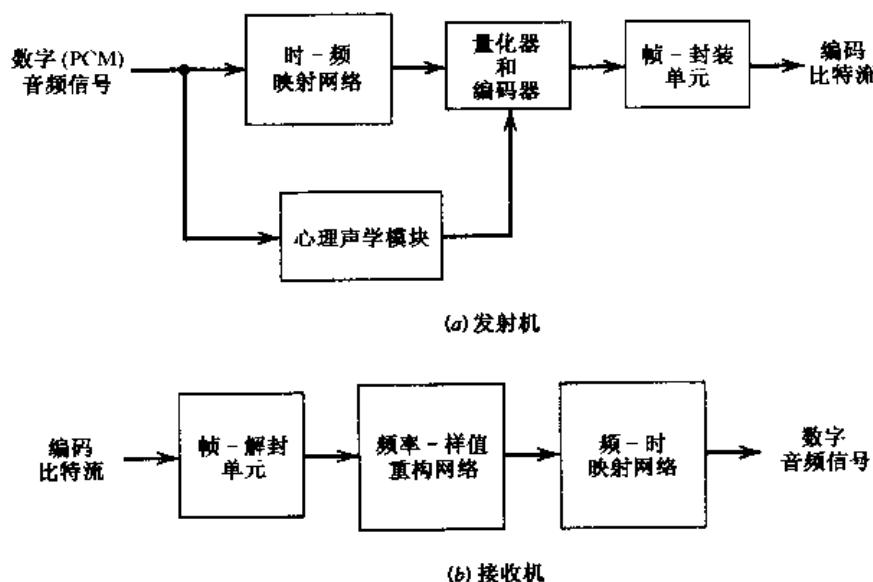


图 3.34 MPEG/音频编码系统

心理声学模块是编码器的核心部分。其功能是对输入音频信号的频谱内容进行分析,然后计算出这三层中每层每个子带的信号-屏蔽比。这一信息被量化器-编码器用于决定如何将可用比特分配给子带信号以进行量化。动态分配比特的目的是使量化噪声的可听性降至最小。最后,帧-封装单元将量化后的音频样值集合为编码比特流。

译码器完成的是编码器所进行的信号处理操作的逆过程,将收到的编码比特流转换为时域音频信号。

总之,MPEG-1/音频编码标准代表了音频信号编码的技术状态。I 层在立体声比特率约为 384 Kbps 且具有透明传输质量时,实现的压缩率为 4。相应地,II 层和 III 层的压缩率分别为 8 和 12,立体声比特率分别近似等于 192 Kbps 和 128 Kbps。MPEG-1/音频编码标准的主观质量等价于多种音乐采用的压缩光盘质量(16 比特 PCM)。压缩光盘(CD)是当今数字音频表示的实际标准。

3.18 总结与讨论

本章介绍了两个基本的互补过程：

- 抽样，在时域中完成。抽样过程将模拟信号波形与它的离散-时间表示形式联系起来。
- 量化，在幅域中完成。量化过程将模拟信号波形与它的离散-幅度表示形式联系起来。

抽样过程建立在抽样定理的基础上。该定理指出，不含高于 W 赫的频率分量的严格带限信号，可由均匀抽样速率等于或大于奈奎斯特速率的样值序列惟一表示。奈奎斯特速率等于每秒 $2W$ 个样值。量化过程利用了一个事实，即作为最后接收者的任何人类感官只能察觉出有限的强度差异。

抽样过程是所有脉冲调制系统的基础。脉冲调制可分为模拟脉冲调制和数字脉冲调制。两者的主要区别在于，模拟脉冲调制系统保持了消息信号幅度的连续性，而数字脉冲调制系统通过量化过程，将消息信号表示为在时间和幅值上均离散的形式。

模拟脉冲调制通过改变传输脉冲的某些参数如幅度、宽度或位置来实现，相应地就分别有 PAM、PDM 或 PPM。在多个信道的时分复用中，信号处理过程经常从 PAM 开始。若在时分复用中采用 PDM 或 PPM，必须要确保在整个调制范围内，某个消息信号的脉冲不会进入到属于另一个消息信号的时隙中。这个约束条件浪费了具有较高峰值因数的电话系统的时间间隔，这也是在电话系统中不采用 PDM 或 PPM 的原因之一。尽管 PPM 的效率比 PDM 要高，但两者都不是用传输带宽换取更高噪声性能的理想系统。

数字脉冲调制系统通过抽样和量化，用编码脉冲序列传输模拟消息信号。PCM 是数字脉冲调制的一种重要形式，它具有很多独特的优点。正是由于这些优点，在传输模拟信号如语音和视频信号时，PCM 成为一种标准的调制方法。这种调制方法的优点包括：很强的抗噪声和抗干扰能力，能够在传输路径中有效再生编码脉冲，以及不同类型的基带信号具有统一的格式等。

实际上，正是由于 PCM 具有这一系列独特的优点，才使它成为了 PSTN 采用的组网方法。从上下文中不难看出，连接 PSTN 和因特网服务提供商的电话信道，由于使用了压缩扩展，因此是非线性的。更为重要的是，这些电话信道同时还是全数字的。这对于计算机用户和服务器之间的高速调制解调器的设计有着重要的影响。第 6 章将讨论这个问题。

DM 和 DPCM 是另外两种有用的数字脉冲调制形式。DM 的主要优点是电路简单。相反地 DPCM 通过增加电路复杂度，降低了信道带宽。这种改进是通过预测的思想实现的，其中去掉了输入数据流中的冗余符号。DPCM 中更进一步的改进，是针对输入数据的统计变化，采用了自适应技术。这样，在系统性能没有严重劣化的前提下，大大降低了带宽需求。

与 ADPCM 不同，MPEG-1 音频编码标准以透明的、感觉无损的方式达到了压缩立体声音频信号的目的。这些卓越的性能是通过利用听觉系统的某些心理声学特性而达到的。

从严格意义上说，术语脉冲调制是用词不当的，因为无论是模拟调制还是数字调制，实际上都是信源编码技术。之所以这样说，是因为消息信号在经历了脉冲调制过程中的一系列变化后，仍然是基带信号。脉冲调制信号的基带本质，可以从如下事实看出，即不管调制后信号的具体形式如何，都可以通过带宽适当的基带信道进行传输。实际上，下一章讲述的正是用脉

冲序列表示的数据的基带传输。

值得重视的是,脉冲调制技术是有损的,因为已调信号新的表示形式将造成一些信息的丢失。例如,在 PAM 中,习惯上在抽样之前使用低通抗混叠滤波器。这样就会滤掉信号中不重要的高频分量,从而损失了相应的信息。从 PCM 产生的量化噪声(即失真)中,可更清楚地看到脉冲调制的有损性质。PCM 所产生的编码脉冲序列并不具备准确表示连续样值的无限精度。但采用脉冲调制过程造成的信息损失是设计者能够控制的,因为设计者能够让信息损失小到终端用户无法察觉的程度。

本章从信号处理的角度讨论了脉冲调制。第 9 章将从通信系统的信息论方面再次讲述 PCM。通过第 9 章的学习,读者将对这一信源编码技术产生更深刻的认识。

注释与参考

- [1] 关于脉冲调制的经典书籍是 Black(1953)。关于此问题,Rowe(1965)中给予了更详细的讲解。对于 PPM 系统的噪声分析,见 Haykin(1994)的第三版。
- [2] Reeves 在 1937 年发明了 PCM。关于此发明的历史性评述见 Reeves(1975)。Jayant 和 Noll(1984)中对 PCM, DPCM, DM 以及它们的变化形式给予了更详细的讲解。Jayant (1976)编辑的书中收集了有关波形量化和编码的早期论文。
- [3] 关于 PCM 系统中量化噪声的讨论细节见 Bennett(1948)以及 Rowe(1965, pp.311 – 321)。
- [4] 非均匀量化器最优化的两个必要条件是式(3.42)和式(3.47)。它们分别由 Lloyd (1957)和 Max(1960)提出,因此命名为“Lloyd-Max 量化器”。关于这两个最优化条件的推导见 Gersho 和 Gray(1992)。
- [5] Smith(1957)讲述了用于信号压缩的 μ 律。美国、加拿大和日本采用的是 μ 律。欧洲采用 A 律进行信号压缩。A 律见 Gattermode(1969, pp.133 – 140)。有关 μ 律和 A 律的讨论见 Kaneko(1970)。
- [6] Fultz 和 Penick(1965)阐述了 T1 载波系统的原始形式。例 3.2 是系统的最新形式,见 Henning 和 Pan(1972)。
- [7] AT&T 最先开始采用关于数字复用的北美/日本标准。CCITT 为世界其他地方选用了另一套标准。除了 5 个数字复接器中信道输入的数量以及它们各自的比特速率有些变动外,CCITT 数字等级与 3.10 节描述的基本相同。
- [8] 关于 DM 的原始论文见 Schouten, DeJager 和 Greefkes(1952)以及 DeJager(1952)。关于 DM 的再次论述见 Schindler(1970)。
- [9] 增量总和调制见 Jayant 和 Noll(1984, pp.399 – 400), 以及 Inose, Yasuda 和 Murakami (1962)。
- [10] Cutler 发明了 DPCM。1952 年的专利发布会公布了此项发明。PCM 和 DPCM 的噪声性能比较见 Jayant(1974), 以及 Rabiner 和 Schafer(1978, Chapter 5)。
- [11] 关于 ADPCM 的讨论见 Jayant 和 Noll(1984)。
- [12] 式(3.84)的 ADM 算法是 Sklar(1988, p.641)算法的改进形式。Sklar 的算法由 Song et al 等人(1971)早期的论文中所述的方法改进得到。Sklar 的论文还提出了最优化的

ADM 系统。为了实际应用的需要,在后期的论文里,描述最优化系统的高阶非线性方程被折线方程近似地代替了。

- [13] 在 Brandenburg 和 Stoll(1994), Pan(1993)以及 Madisetti 和 Williams(1998)编辑的关于数字信号处理的手册中 Peter Noll 的文章里均描述了 MPEG-1/音频编码标准。手册中还讨论了 MPEG-1 的后续标准 MPEG-2。值得注意的是,MPEG-2 提出了以低于 MPEG-1 的抽样频率进行立体声音频编码的技术。
- [14] 听觉器官可对输入声波产生响应,它由三个主要部分组成,这些部分的功能分别如下:
 - 外耳用来收集声音。
 - 中耳为空气与耳蜗流体之间提供声抗匹配,将输入声音造成的鼓膜(耳膜)振动以有效的方式传递给内耳。
 - 内耳将中耳传递过来的机械振动转化为传输给大脑进一步处理的电化或神经中枢信号。

习题

抽样过程

- 3.1 带宽为 10 kHz 的窄带信号,其中心频率等于载频 100 kHz。现分别对该信号的同相分量和正交分量进行抽样,得到它的离散时间表示形式。这种表示形式允许的最小抽样速率等于多少?证明你的结果。如何由同相和正交分量的抽样形式恢复出原始窄带信号?
- 3.2 在自然抽样中,模拟信号 $g(t)$ 和周期性的矩形脉冲串 $c(t)$ 相乘。已知该周期性脉冲串的重复频率为 f_s ,每个矩形脉冲的持续时间为 $T(f_s T \ll 1)$,试求:
 - (a) 信号 $s(t)$ 经自然抽样后的频谱。可假设时刻 $t=0$ 相当于 $c(t)$ 中矩形脉冲的中点。
 - (b) 在满足抽样定理的条件下,证明原始信号 $m(t)$ 能从其自然抽样形式中完全地恢复出来。
- 3.3 指出下列信号的奈奎斯特速率和奈奎斯特间隔:
 - (a) $g(t) = \text{sinc}(200t)$
 - (b) $g(t) = \text{sinc}^2(200t)$
 - (c) $g(t) = \text{sinc}(200t) + \text{sinc}^2(200t)$
- 3.4 (a) 画出调制信号产生的 PAM 信号的频谱。假定调制频率 $f_m = 0.25$ Hz, 抽样周期 $T_s = 1$ s, 脉冲持续时间 $T = 0.45$ s。

$$m(t) = A_m \cos(2\pi f_m t)$$
 (b) 画出理想重构滤波的输出频谱。将此结果与无孔径效应时的相应输出进行比较。

脉冲幅度调制

- 3.5 图 P3.5 给出了消息信号 $m(t)$ 的理想频谱。以 1 kHz 的抽样速率,用幅度为 1、宽度为

0.1 ms 的平顶脉冲对信号进行抽样。计算并粗略画出产生的 PAM 信号的频谱图。

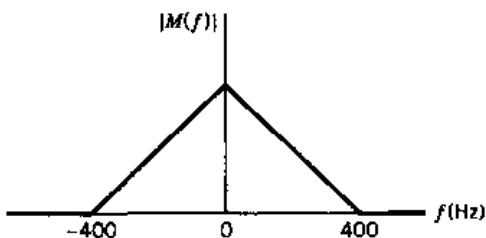


图 P3.5

- 3.6 本题计算 PAM 系统发生孔径效应时需采用的均衡措施。工作频率 $f = f_s/2$, f 对应于消息信号的最高频率分量,且抽样速率等于奈奎斯特速率。画出 $1/\text{sinc}(0.5 T/T_s)$ 对 T/T_s 的曲线,从而找出当 $T/T_s = 0.1$ 时需采取的均衡措施。
- 3.7 考虑通过存在高斯白噪声的信道传输 PAM 信号,信道的最小带宽 $B_T = 1/2 T_s$,其中 T_s 为抽样周期。噪声是零均值的且功率谱密度等于 $N_0/2$ 。PAM 信号使用标准脉冲 $g(t), g(t)$ 的傅里叶变换定义如下:

$$G(f) = \begin{cases} \frac{1}{2B_T}, & |f| < B_T \\ 0, & |f| > B_T \end{cases}$$

对于满幅正弦调制信号,证明:在平均发射功率相等的情况下,PAM 传输和基带信号传输具有相同的信噪比。

- 3.8 对 24 个语音信号均匀抽样后进行时分复用,抽样过程采用宽度为 $1 \mu\text{s}$ 的平顶抽样。为了保持同步,在复用信号中额外加了一个宽度为 $1 \mu\text{s}$,幅值足够大的脉冲。每个语音信号最高频率分量为 3.4 kHz 。
- 假定抽样速率为 8 kHz ,计算复用信号中相继脉冲的时间间隔。
 - 在奈奎斯特抽样速率下,重复(a)中的计算。
- 3.9 12 个不同的消息信号,带宽均为 10 kHz ,将在复用后传输。计算采用下列复用/调制方法时,需要的最小带宽值分别为多少。
- FDM, SSB。
 - TDM, PAM。
- 3.10 PAM 遥感勘测系统中包括 4 个输入信号的复用: $s_i(t), i = 1, 2, 3, 4$ 。其中两个信号 $s_1(t)$ 和 $s_2(t)$ 的带宽均为 80 Hz ,而另两个信号 $s_3(t)$ 和 $s_4(t)$ 的带宽均为 1 kHz 。分别对 $s_3(t)$ 和 $s_4(t)$ 以每秒 2400 个样值的速率进行抽样。将此抽样速率除以 2^R (即 2 的整数次幂)后作为 $s_1(t)$ 和 $s_2(t)$ 的抽样速率。
- 计算 R 的最大值。
 - 用(a)中算出的 R 值设计一个复用系统,首先将 $s_1(t)$ 和 $s_2(t)$ 复合成一个新的序列 $s_5(t)$,然后将 $s_3(t), s_4(t)$ 和 $s_5(t)$ 复用在一起。

线路码

- 3.11 本题推导用于计算3.7节中介绍的5种线路码的功率谱密度的公式。功率谱密度曲线如图3.16所示。对于每种线路码，其比特持续时间为 T_b ，且脉冲幅度A的取值使线路码平均功率的归一化值等于1，如图3.16所示。假定数据流是随机产生的，符号0和1是等概率的。

推导下面总结的线路码的功率谱密度：

(a) 单极性 NRZ 信号：

$$S(f) = \frac{A^2 T_b}{4} \operatorname{sinc}^2(f T_b) \left(1 + \frac{1}{T_b} \delta(f) \right)$$

(b) 极性 NRZ 信号：

$$S(f) = A^2 T_b \operatorname{sinc}^2(f T_b)$$

(c) 单极性 RZ 信号：

$$S(f) = \frac{A^2 T_b}{16} \operatorname{sinc}^2\left(\frac{f T_b}{2}\right) \left[1 + \frac{1}{T_b} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{T_b}\right) \right]$$

(d) 双极性 RZ 信号：

$$S(f) = \frac{A^2 T_b}{4} \operatorname{sinc}^2\left(\frac{f T_b}{2}\right) \sin^2(\pi f T_b)$$

(e) 曼彻斯特编码信号：

$$S(f) = A^2 T_b \operatorname{sinc}^2\left(\frac{f T_b}{2}\right) \sin^2\left(\frac{\pi f T_b}{2}\right)$$

并由此证实图3.16的频谱曲线。

- 3.12 假设对随机二进制数据流(符号是等概率的)进行差分编码后,采用习题3.11中的5种线路码之一进行传输。差分编码给发送数据的功率谱密度带来的影响是什么?证明你的结论。
- 3.13 随机产生的数据流由等概率的二进制符号0和1组成。将其编码为极性NRZ波形,每个二进制符号的定义如下:

$$s(t) = \begin{cases} \cos\left(\frac{\pi t}{T_b}\right), & -\frac{T_b}{2} < t \leq \frac{T_b}{2} \\ 0, & \text{其他} \end{cases}$$

- (a) 假定数据流是00101110,大致画出编码后的波形图。
- (b) 推导该信号的功率谱密度公式,并画出功率谱密度草图。提示:利用式(1.52)。
- (c) 将此随机波形的功率密度谱与习题3.11(b)部分定义的公式进行比较。
- 3.14 已知数据流1110010100,画出下列线路码的发送脉冲序列草图:
- (a) 单极性 NRZ
- (b) 双极性 NRZ
- (c) 单极性 RZ
- (d) 双极性 RZ
- (e) 曼彻斯特码

- 3.15 对习题 3.14 中的二进制数据流进行差分编码,然后用 3.14 题中的 5 种线路码之一传输。假设参考比特为符号 1,画出每个发送数据流的草图。如果假设参考比特为符号 0,对结果会有什么影响?

脉冲编码调制

- 3.16 一个语音信号持续 10 s。以 8 kHz 的速率对其抽样,然后进行编码。需要的信噪比(或信号-量化噪声比)为 40 dB。计算此数字化语音信号需要的最小存储容量值。

- 3.17 考虑输入-输出关系如图 3.10(a)所示的一个均匀量化器。假定量化器的输入服从均值为零、方差为 1 的高斯分布。

- (a) 输入信号的幅值位于 -4 到 +4 范围以外的概率是多少?
 (b) 由(a)的结果,证明量化器的输出信噪比等于下式:

$$(SNR)_o = 6R \approx 7.2 \text{ dB}$$

其中, R 为每样值的比特数。可假定量化器的输入在 -4 到 +4 的范围内。将

(b) 的结果与例 3.1 得到的结果加以比较。

- 3.18 PCM 系统采用的是均匀量化器和 7 比特二进制编码器的级联。系统的比特率等于 $50 \times 10^6 \text{ bps}$ 。

- (a) 系统可正常工作的最大消息带宽是多少?
 (b) 当输入频率 1 MHz 的满幅正弦调制信号时,计算输出信-噪比(信号-量化噪声比)。

- 3.19 证明对于非均匀量化器,其量化误差的均方值近似等于 $(1/12) \sum_i \Delta_i^2 p_i$, 其中 Δ_i 为第 i 个量化间隔的大小, p_i 为输入信号幅值落在第 i 个量化间隔内的概率。假定量化间隔 Δ_i 小于输入信号的偏移量。

- 3.20 (a) 将幅值为 3.25 V 的正弦信号加在输出为 0, ± 1 , ± 2 , ± 3 V 的中平型均匀量化器上。当输入为一个完整周期时,画出量化器的输出信号波形草图。

- (b) 当采用输出为 0.5, ± 1.5 , ± 2.5 , ± 3.5 V 的中升型量化器时,重复(a)中的计算。

- 3.21 采用 4 比特二进制 PCM 系统传输信号。量化器是中升型,量化间隔为 1 V,画出输入为一个完整周期时产生的 PCM 信号波形草图。假定抽样速率等于每秒 4 个样值,抽样时刻在 $t = \pm 1/8, \pm 3/8, \pm 5/8, \dots$ s 处。

$$m(t) = 6 \sin(2\pi t) \text{ V}$$

- 3.22 图 P3.22 所示为 PCM 信号。分别用 +1 V 和 -1 V 的幅度电平表示二进制符号 1 和 0。码字由 3 个比特组成。确定产生此 PCM 信号的模拟信号的抽样。

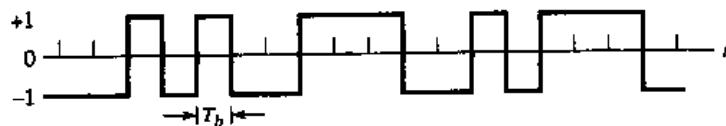


图 P3.22

- 3.23 考虑一串($n-1$ 个)再生中继器,对二进制 PCM 信号共有 n 次顺序的判决。其中,最后一次判决在接收机中进行。假设通过系统传输的任意二进制符号被中继器反

转的概率彼此独立,均为 p_1 。用 p_n 代表二进制符号经整个系统传输后出错的概率。

(a) 证明

$$p_n = \frac{1}{2}[1 - (1 - 2p_1)^n]$$

(b) 如果 p_1 很小,而 n 不是太大,那么相应的 p_n 值等于多少?

3.24 讨论设计 PCM 的再生中继器的基本问题。

增量调制

3.25 考虑由双曲正切函数定义的测试信号 $m(t)$:

$$m(t) = A \tanh(\beta t)$$

其中, A 和 β 均为常量。计算此信号进行增量调制 Δ 时,为避免斜率过载,所需要的最小步长。

3.26 考虑将频率为 f_m ,幅度为 A_m 的正弦信号,加在步长为 Δ 的增量调制器中。证明斜率过载失真将在下述情况发生:

$$A_m > \frac{\Delta}{2\pi f_m T_s}$$

这里的 T_s 为抽样周期。当无斜率过载失真时,最大的发射功率是多少?

3.27 设计一个增量调制器,用于处理频带限制在 3.4 kHz 内的语音信号。调制器参数如下:

- 抽样速率 = $10 f_{Nyquist}$, 其中 $f_{Nyquist}$ 是语音信号的奈奎斯特速率。
- 步长 $\Delta = 100$ mV。

用 1 kHz 的正弦信号测试此调制器。确定为避免斜率过载,测试信号的最大幅值。

3.28 本题推导 DM 系统平均信噪比(信号-量化噪声比)的经验公式,测试信号是幅度为 A ,频率为 f_m 的正弦信号。假定系统产生的颗粒噪声的功率谱密度由下式给出:

$$S_N(f) = \frac{\Delta^2}{6f_s}$$

其中, f_s 为抽样速率, Δ 为步长(注意,此公式与 PCM 系统量化噪声的功率谱密度公式基本相同,只是 PCM 系统中的 $\Delta/2$ 由 DM 系统中的 Δ 取代了)。DM 系统是为了处理频率限制在带宽 W 以内的模拟消息信号而设计的。

(a) 证明系统产生的平均量化噪声功率等于

$$N = \frac{4\pi^2 A^2 f_m^2 W}{3f_s^3}$$

这里假设步长 Δ 的选取符合习题 3.27 的公式,以避免斜率过载。

(b) 计算正弦输入下,DM 系统的信噪比(信号-量化噪声比)。

3.29 考虑适用于带宽限制在 $W = 5$ Hz 内的模拟消息信号的 DM 系统。将幅度 $A = 1$ V, 频率 $f_m = 1$ kHz 的正弦测试信号加在该系统上。系统的抽样速率等于 50 kHz。

(a) 计算最小化斜率过载所需要的步长 Δ 。

- (b) 对于指定的正弦测试信号,计算系统信噪比(信号-量化噪声比)。计算中可采用习题 3.27 和习题 3.28 推导出的公式。
- 3.30 考虑带宽为 3 kHz 的低通信号。采用步长 $\Delta = 0.1 \text{ V}$ 的线性 DM 系统处理该信号,抽样速率等于奈奎斯特速率的 10 倍。
- 当系统不发生斜率过载失真时,计算频率为 1 kHz 的正弦测试信号的最大幅值。
 - 在条件(a)下,分别计算滤波前和滤波后的输出信噪比。

线性预测

- 3.31 用一阶线性预测器处理正弦信号的抽样。抽样速率等于 $10f_0$, f_0 为正弦信号的频率。预测器有一个相关系数,用 w_1 表示。
- 确定最小化预测误差方差所需最优 w_1 值。
 - 确定预测误差方差的最小值。
- 3.32 稳态过程 $X(t)$ 的自相关函数取值如下:

$$R_X(0) = 1$$

$$R_X(1) = 0.8$$

$$R_X(2) = 0.6$$

$$R_X(3) = 0.4$$

- 计算含有 3 个单位时延的最优线性预测器的相关系数。
 - 计算相应预测误差的方差。
- 3.33 采用含有两个单位时延的线性预测器,重复习题 3.32 的计算。比较这两个最优线性预测器的性能。

差分脉冲编码调制

- 3.34 DPCM 系统采用一个时延的线性预测器。延迟了一个抽样间隔的输入信号的归一化自相关函数值为 0.75。对此预测器进行设计,使它的预测误差的方差达到最小。计算此预测器获得的处理增益值。
- 3.35 计算采用习题 3.32 中的三时延最优线性预测器的 DPCM 系统的处理增益,与采用习题 3.33 中的二时延最优线性预测器的 DPCM 系统的处理增益相比,有多大的改进值。计算中可使用习题 3.32 给出的输入信号自相关函数值。
- 3.36 本题比较 DPCM 系统与采用压扩的普通 PCM 系统的性能。
- 在量化级数足够大的情况下,PCM 系统的信噪比(信号-量化噪声比)一般由下式定义:

$$10 \log_{10}(\text{SNR})_0 = \alpha + 6n \text{ dB}$$

其中, 2^n 为量化级数。采用 μ 律压扩的 PCM 系统的常量 α 定义如下:

$$\alpha \approx 4.77 - 20 \log_{10}(\log(1 + \mu)) \text{ dB}$$

另一方面,对于 DPCM 系统,常量 α 位于 $-3 < \alpha < 15 \text{ dB}$ 的范围内。这里引用的公式适用于电话质量语音信号。

在下列假定下,比较 DPCM 系统和采用 $\mu = 255$ 的 μ 律压扩 PCM 系统的性能:

- 在每样值的比特数相同的情况下,与压扩 PCM 的 $(SNR)_o$ 相比,DPCM 带来的 $(SNR)_o$ 的改善值。
- 在同样的 $(SNR)_o$ 情况下,与压扩 PCM 相比,DPCM 每样值需要的比特数减少的数量。

3.37 在图 P3.37 描述的 DPCM 系统中,证明:在没有信道噪声的情况下,发射预测滤波器和接收预测滤波处理的输入信号之间差别很小。

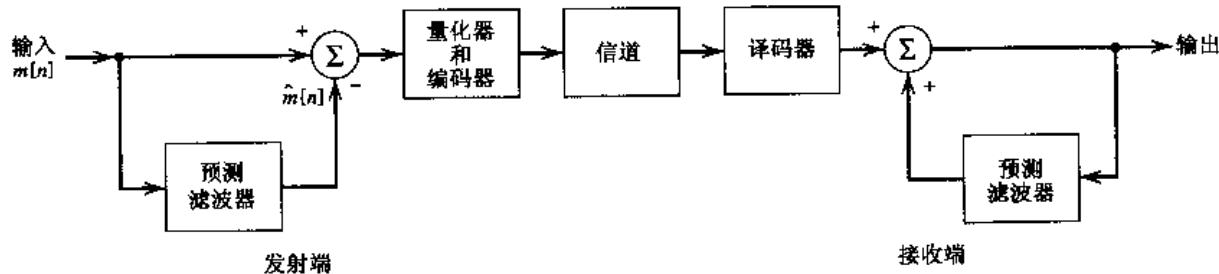


图 P3.37

计算机实验

3.38 以 8 kHz 的速率对频率 $f_0 = 10^4/2\pi$ Hz 的正弦信号进行抽样,然后送入抽样-保持电路中,得到脉冲宽度 $T = 500 \mu s$ 的平顶 PAM 信号 $s(t)$ 。

- 给出 PAM 信号 $s(t)$ 的波形。
- 计算 PAM 信号 $s(t)$ 的幅度谱 $|S(f)|$ 。
- 计算 $|S(f)|$ 的包络。从而证明包络第一次过零时的频率等于 $(1/T) = 20$ kHz。

3.39 本题采用计算机仿真,比较使用 μ 律的压扩 PCM 系统与使用均匀量化的相应系统的性能。仿真使用幅度可变的正弦输入信号。

- 采用表 3.4 中的 μ 律,画出输出信噪比对输入信噪比的变化曲线,所有比例值均用 dB 表示。
- 将(a)中的计算结果与有 256 个量化级的均匀量化器的计算结果进行比较。

3.40 本实验研究具有下面递归形式的信号 $x[n]$ 的线性自适应预测:

$$x[n] = 0.8x[n - 1] - 0.1x[n - 2] + 0.1v[n]$$

其中, $v[n]$ 来自于均值为零、方差为 1 的离散时间白噪声过程。(将这种方式产生的过程称为二阶自回归过程。)采用下面定义的标准 LMS 算法进行自适应预测:

$$\begin{aligned}\hat{x}[n] &= \sum_{k=1}^p w_k[n]x[n - k] \\ e[n] &= x[n] - \hat{x}[n] \\ w_k[n + 1] &= w_k[n] + \left(\frac{\mu}{\sum_{k=1}^p x^2[n - k]} \right) x[n - k]e[n], \quad k = 1, 2, \dots, p\end{aligned}$$

其中, p 为预测阶数, μ 为归一化的步长参数。需注意 μ 是无量纲的。按下式选择

μ 时, 可以保证算法的稳定性:

$$0 < \mu < 2$$

对于所有

$$w_k[0] = 0 \quad \text{对于所有 } k$$

算法的学习曲线的定义为参数值指定时, 均方误差相对于迭代次数 n 的变化曲线。将算法运行多次, 并将各次得到的 $e^2[n]$ 关于 n 的曲线取平均, 就可以得到学习曲线。

- (a) 在固定的预测阶数 $p = 5$ 和三个不同的步长参数值: $\mu = 0.0075, 0.05$ 和 0.5 下, 画出关于 $x[n]$ 的自适应预测的学习曲线。
- (b) 由(a)部分的学习曲线能得出什么结论?

第4章 基带脉冲传输

本章讨论通过基带信道传输数字数据的问题,重点讨论以下几个问题:

- 匹配滤波器,即在加性高斯白噪声下检测已知信号的最佳系统。
- 计算由信道噪声所引起的误比特率。
- 符号间干扰,是由于信道在实际应用中通常呈弥散特性而引起的。
- 用于无失真基带数据传输的奈奎斯特准则。
- 用于消除符号间干扰影响的相关编码或部分响应信号。
- 数字用户线。
- 弥散基带信道的均衡。
- 用于显示数据传输中符号间干扰和信道噪声影响的眼图。

4.1 引言

在第3章中,我们讨论了将承载消息的模拟信号转换为数字形式的技术。在实际应用中,还存在另外一种数字数据的产生方式:数据可以代表离散信息源(如数字计算机)的输出。在本章中,我们研究数字数据(任何信源)通过基带信道的传输。^[1]使用调制在带通信道进行的数据传输在第6章中讨论。

数字数据具有很宽的频谱,其中的低频分量十分重要。因此,数字数据的基带传输需要其带宽足以容纳数据流基本频率分量的低通信道。但由于信道频率响应往往是偏离理想低通滤波器频率响应的,所以信道通常是弥散的。这样的信道使得数据传输的结果,即每个接收到的脉冲都会受到相邻脉冲的影响,从而,产生一种称为码间干扰(ISI)的常见干扰形式。ISI是接收机输出端的重构数据流中产生比特差错的一个主要原因。要纠正ISI,就必须对整个系统的脉冲波形进行控制。因此,本章的大部分内容将用于介绍各种脉冲波形。

引起基带数据传输系统产生比特差错的另一个重要原因是普遍存在的信道噪声。显然,信道噪声和ISI在系统中是同时存在的。但为了理解它们是如何影响系统性能的,我们首先将它们分别进行考虑。在本章的后面部分,我们将研究二者混合作用的影响。

本章首先介绍通信理论中的一个基本结论。该结论用于检测淹没在加性高斯白噪声中的波形已知的脉冲信号。检测脉冲的最佳设备包括一个称为匹配滤波器^[2]的线性时不变滤波器。这样命名的原因是该滤波器的冲激响应与所检测的脉冲信号是相匹配的。

4.2 匹配滤波器

在通信系统研究中,经常出现的一个基本问题就是检测一个通过信道传输并受到信道噪声(如接收机前端的加性噪声)干扰的脉冲信号。为讨论方便,我们假设系统的主要干扰源为信道噪声。

接下来考虑图 4.1 所示的接收机模型。该模型包括了一个冲激响应为 $b(t)$ 的线性时不变滤波器。滤波器输入 $x(t)$ 由一个被加性信道噪声 $w(t)$ 破坏的脉冲信号 $g(t)$ 组成, 即

$$x(t) = g(t) + w(t), \quad 0 \leq t \leq T \quad (4.1)$$

其中, T 为任意的观察间隔。脉冲信号 $g(t)$ 可以是数字通信系统中的一个二进制符号 1 或 0。 $w(t)$ 是一个均值为零、功率谱密度为 $N_0/2$ 的白噪声过程的抽样函数。假设对于接收机来说, 脉冲信号 $g(t)$ 的波形是已知的。因此不确定性来自于噪声 $w(t)$ 。接收机的功能是在给定接收信号 $x(t)$ 的情况下, 以最优的方式检测出脉冲信号 $g(t)$ 。要满足这样的要求, 必须对滤波器进行最优化设计, 使得从统计角度来说, 噪声对滤波器输出的影响最小, 从而提高脉冲信号 $g(t)$ 的检测效率。

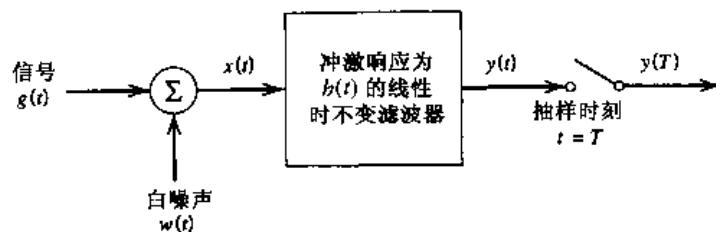


图 4.1 线性接收机

既然滤波器是线性的, 则其输出 $y(t)$ 就可以表示为

$$y(t) = g_o(t) + n(t) \quad (4.2)$$

其中, $g_o(t)$ 和 $n(t)$ 由输入 $x(t)$ 中的信号分量和噪声分量分别产生。要使输出信号分量 $g_o(t)$ 远大于输出噪声分量 $n(t)$, 一个简单的办法就是使滤波器的输出信号分量 $g_o(t)$ 在时刻 $t = T$ 的瞬时功率值, 与输出噪声 $n(t)$ 的平均功率相比尽可能大。这等价于将峰值脉冲信噪比最大化。峰值脉冲信噪比的定义为

$$\eta = \frac{|g_o(T)|^2}{E[n^2(t)]} \quad (4.3)$$

其中, $|g_o(T)|^2$ 为输出信号的瞬时功率值, E 为期望运算符, $E[n^2(t)]$ 为平均输出噪声功率的度量值。要求具体设计滤波器的冲激响应 $b(t)$, 使得式(4.3)中的输出信噪比达到最大值。

用 $G(f)$ 表示已知信号 $g(t)$ 的傅里叶变换, $H(f)$ 表示该滤波器的频率响应。那么, 输出信号 $g_o(t)$ 的傅里叶变换就等于 $H(f)G(f)$, 则 $g_o(t)$ 可从下面的傅里叶反变换得出:

$$g_o(t) = \int_{-\infty}^{\infty} H(f)G(f) \exp(j2\pi ft) df \quad (4.4)$$

因此, 当对滤波器输出在时刻 $t = T$ 进行抽样时, 有(在不考虑信道噪声的情况下):

$$|g_o(T)|^2 = \left| \int_{-\infty}^{\infty} H(f)G(f) \exp(j2\pi fT) df \right|^2 \quad (4.5)$$

下面考虑噪声 $w(t)$ 单独作用时对滤波器输出的影响。输出噪声 $n(t)$ 的功率谱密度 $S_n(f)$ 等于输入噪声 $w(t)$ 的功率谱密度乘以平方幅度响应 $|H(f)|^2$ (见 1.7 节)。由于 $w(t)$ 是具有常数功率谱密度 $N_0/2$ 的白噪声, 因此

$$S_N(f) = \frac{N_0}{2} |H(f)|^2 \quad (4.6)$$

从而,输出噪声 $n(t)$ 的平均功率为

$$\begin{aligned} E[n^2(t)] &= \int_{-\infty}^{\infty} S_N(f) df \\ &= \frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 df \end{aligned} \quad (4.7)$$

将式(4.5)和式(4.7)代入式(4.3),可以将峰值脉冲信噪比的表达式重写为

$$\eta = \frac{\left| \int_{-\infty}^{\infty} H(f)G(f) \exp(j2\pi fT) df \right|^2}{\frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 df} \quad (4.8)$$

现在的问题就是在给定 $G(f)$ 的情况下,寻找滤波器频率响应 $H(f)$ 的特定表达式,使得 η 取最大值。为了找到解决办法,我们在式(4.8)的分子中应用了一个称为施瓦茨不等式(Schwarz's inequality)的数学结论。

施瓦茨不等式的推导可参见第5章。假设有两个复函数 $\phi_1(x)$ 和 $\phi_2(x)$ (其中 x 为实变量)分别满足下列条件:

$$\int_{-\infty}^{\infty} |\phi_1(x)|^2 dx < \infty$$

和

$$\int_{-\infty}^{\infty} |\phi_2(x)|^2 dx < \infty$$

则

$$\left| \int_{-\infty}^{\infty} \phi_1(x)\phi_2(x) dx \right|^2 \leq \int_{-\infty}^{\infty} |\phi_1(x)|^2 dx \int_{-\infty}^{\infty} |\phi_2(x)|^2 dx \quad (4.9)$$

式(4.9)中的等号,当且仅当

$$\phi_1(x) = k\phi_2^*(x) \quad (4.10)$$

时成立。其中, k 为一个任意常量, * 代表取复共轭。

回到当前的问题。应用施瓦茨不等式(4.9),并令 $\phi_1(x) = H(f)$, $\phi_2(x) = G(f)\exp(j2\pi fT)$, 则式(4.8)的分子可写为

$$\left| \int_{-\infty}^{\infty} H(f)G(f) \exp(j2\pi fT) df \right|^2 \leq \int_{-\infty}^{\infty} |H(f)|^2 df \int_{-\infty}^{\infty} |G(f)|^2 df \quad (4.11)$$

代入式(4.8),可将峰值脉冲信噪比重新定义为

$$\eta \leq \frac{2}{N_0} \int_{-\infty}^{\infty} |G(f)|^2 df \quad (4.12)$$

该关系式的右边部分并不依赖于滤波器的频率响应 $H(f)$,而只由信号能量和噪声功率谱密度决定。所以,只要 $H(f)$ 取适当值,使得等号成立,峰值脉冲信噪比 η 就会达到最大值。即

$$\eta_{\max} = \frac{2}{N_0} \int_{-\infty}^{\infty} |G(f)|^2 df \quad (4.13)$$

相应地,设 $H(f)$ 的最优值为 $H_{\text{opt}}(f)$ 。要找到该最优值,需要用到式(4.10),此时有

$$H_{\text{opt}}(f) = k G^*(f) \exp(-j2\pi f T) \quad (4.14)$$

其中, $G^*(f)$ 为输入信号 $g(t)$ 的傅里叶变换的复共轭, k 为适当维度的比例因子。这个关系式表明,除了因子 $k \exp(-j2\pi f T)$ 外, 最优滤波器的频率响应与输入信号傅里叶变换的复共轭是相同的。

式(4.14)具体规定了频域中的最优滤波器。为了在时域中描述最优滤波器的特性, 对式(4.14)中的 $H_{\text{opt}}(f)$ 进行傅里叶反变换, 得到最优滤波器的冲激响应为

$$h_{\text{opt}}(t) = k \int_{-\infty}^{\infty} G^*(f) \exp[-j2\pi f(T-t)] df \quad (4.15)$$

对于一个实信号 $g(t)$ 来说, 有 $G^*(f) = G(-f)$, 因此可将式(4.15)重写为

$$\begin{aligned} h_{\text{opt}}(t) &= k \int_{-\infty}^{\infty} G(-f) \exp[-j2\pi f(T-t)] df \\ &= k \int_{-\infty}^{\infty} G(f) \exp[j2\pi f(T-t)] df \\ &= kg(T-t) \end{aligned} \quad (4.16)$$

式(4.16)说明, 如果去掉比例因子 k , 最优滤波器的冲激响应就是输入信号 $g(t)$ 的时间反转和延迟, 也就是说它与输入信号相“匹配”。用这种方式定义的线性时不变滤波器称为匹配滤波器。值得注意的是, 在推导匹配滤波器的过程当中, 我们对信道噪声 $w(t)$ 所做的惟一假设就是它是一个具有零均值且功率谱密度为 $N_0/2$ 的平稳白噪声。换言之, 我们没有对信道噪声 $w(t)$ 的统计特性做任何假设。

4.2.1 匹配滤波器的性质

我们注意到, 一个在持续时间 T 内与脉冲信号 $g(t)$ 相匹配的滤波器, 其特性是冲激响应为输入 $g(t)$ 的时间反转和延迟, 如下式所示:

$$h_{\text{opt}}(t) = kg(T-t)$$

换言之, 除了时延 T 及比例因子 k 之外, 冲激响应 $h_{\text{opt}}(t)$ 是由滤波器所匹配的脉冲信号 $g(t)$ 的波形惟一确定的。在频域里, 除延迟因子之外, 匹配滤波器的频率响应, 是输入信号 $g(t)$ 的傅里叶变换的复共轭, 即

$$H_{\text{opt}}(f) = k G^*(f) \exp(-j2\pi f T)$$

对采用匹配滤波器的信号处理系统进行性能分析, 所得到的最重要的结果为:

匹配滤波器的峰值脉冲信噪比仅由滤波器输入端的信号能量与白噪声的功率谱密度之比决定。

为了证明这个性质, 假设有一个与已知信号 $g(t)$ 相匹配的滤波器。匹配滤波器的输出信号 $g_o(t)$ 的傅里叶变换为

$$\begin{aligned} G_o(f) &= H_{\text{opt}}(f)G(f) \\ &= kG^*(f)G(f) \exp(-j2\pi fT) \\ &= k|G(f)|^2 \exp(-j2\pi fT) \end{aligned} \quad (4.17)$$

将式(4.17)代入傅里叶反变换的公式,得到匹配滤波器在时刻 $t = T$ 的输出为

$$\begin{aligned} g_o(T) &= \int_{-\infty}^{\infty} G_o(f) \exp(j2\pi fT) df \\ &= k \int_{-\infty}^{\infty} |G(f)|^2 df \end{aligned}$$

根据瑞利能量定理(Rayleigh's energy theorem),一个脉冲信号的平方幅谱关于频率的积分等于信号能量 E :

$$E = \int_{-\infty}^{\infty} g^2(t) dt = \int_{-\infty}^{\infty} |G(f)|^2 df$$

因此

$$g_o(T) = kE \quad (4.18)$$

将式(4.14)代入式(4.7),得到平均输出噪声功率为

$$\begin{aligned} E[n^2(t)] &= \frac{k^2 N_0}{2} \int_{-\infty}^{\infty} |G(f)|^2 df \\ &= k^2 N_0 E / 2 \end{aligned} \quad (4.19)$$

这里再次用到了瑞利能量定理。因此,峰值脉冲信噪比有如下最大值:

$$\eta_{\max} = \frac{(kE)^2}{(k^2 N_0 E / 2)} = \frac{2E}{N_0} \quad (4.20)$$

由式(4.20)可见,峰值脉冲信噪比对输入信号 $g(t)$ 波形的依赖性,已经完全被匹配滤波器去除了。相应地,在评价匹配滤波器抗加性白噪声的能力时,我们发现,具有同样能量的信号都是等效的。需注意,信号能量 E 的单位是 J(焦),噪声功率谱密度 $N_0/2$ 的单位是 W/Hz,因此比值 $2E/N_0$ 是无量纲的。我们称 E/N_0 为信号能量对噪声功率的比值。

例 4.1 矩形脉冲的匹配滤波器

如图 4.2(a)所示,设有一个幅值为 A ,持续时间为 T 的矩形脉冲信号 $g(t)$ 。匹配滤波器的冲激响应 $h(t)$ 具有与输入信号完全相同的波形。该匹配滤波器对输入信号 $g(t)$ 的响应即输出信号 $g_o(t)$,具有三角形的波形,如图 4.2(b)所示。

输出信号 $g_o(t)$ 的最大值等于 $kA^2 T$,即因子 k 乘上输入信号 $g(t)$ 的能量。最大输出出现在时刻 $t = T$,如图 4.2(b)所示。

对于矩形脉冲这样的特殊例子,匹配滤波器可以用称为积分-清除电路的电路来实现,其框图见图 4.3。这种积分器在矩形脉冲下计算面积,输出结果在时刻 $t = T$ 上进行抽样,其中 T 为脉冲持续时间。在 $t = T$ 之后的瞬间,积分器恢复到初始状态,电路由此得名。图 4.2(c)显示了图 4.2(a)中的矩形脉冲通过该积分-清除电路后的输出波形。由图可见,在 $0 \leq t \leq T$ 内,该电路的输出与匹配滤波器的输出具有相同的波形。

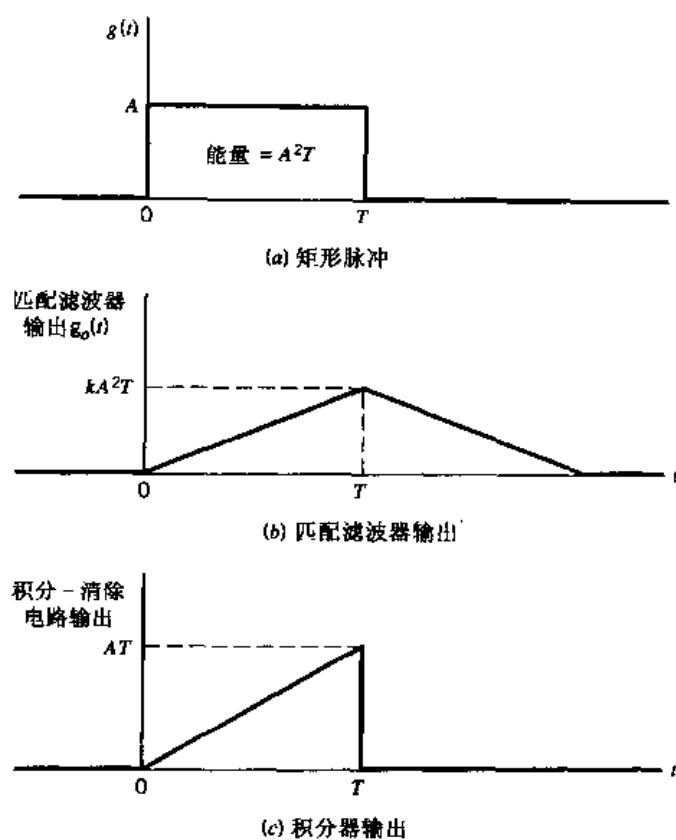


图 4.2 波形图



图 4.3 积分 - 清除电路

4.3 噪声引起的差错率

在 3.8 节中, 我们定性地讨论了信道噪声对二进制 PCM 系统性能的影响。现在我们用匹配滤波器作为加性白噪声下脉冲已知时的最优检测器, 来推导该系统中噪声引起的差错率的计算公式。

考虑一个基于 NRZ 信号的二进制 PCM 系统。在该信号形式里, 符号 1 和 0 分别由幅度和持续时间均相等的正、负矩形脉冲表示。信道噪声模型为具有零均值且功率谱密度为 $N_0/2$ 的加性高斯白噪声 $w(t)$ 。这样假设是为了以后计算的需要。因此, 在信号间隔 $0 \leq t \leq T_b$ 内, 接收信号可写为

$$x(t) = \begin{cases} +A + w(t), & \text{发送符号 1 时} \\ -A + w(t), & \text{发送符号 0 时} \end{cases} \quad (4.21)$$

其中, T_b 为比特持续时间, A 为发送脉冲幅度。假设接收端已知每个发送脉冲的起止时间。换言之, 接收端具有脉冲波形的先验知识, 但不知道其极性。已知含噪声的信号 $x(t)$, 接收机要在每个信号时间间隔内判决发送符号是 1 还是 0。

用于完成判决过程的接收机结构见图 4.4。它由一个匹配滤波器、一个抽样器和一个判决器组成。匹配滤波器与幅度为 A 、持续时间为 T_b 的矩形脉冲相匹配，并利用了接收机可用的位定时信息。在匹配滤波器的每个输出信号间隔的末端进行抽样。信道噪声 $w(t)$ 的存在增加了滤波器输出的随机性。

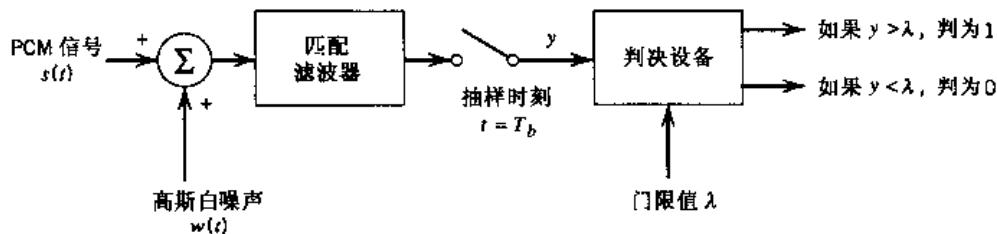


图 4.4 采用双极性 NRZ 信号的二进制编码 PCM 波形的基带传输接收机

设 y 表示在信号间隔末端得到的抽样值，在判决设备里，样值 y 要与预设的门限值 λ 进行比较。如果 y 超过门限值，那么，接收机判决输出信号为 1；否则，判决为 0。当样值 y 与门限值 λ 相等时，我们习惯上让接收机对发送符号进行猜测。这样的判决就如同扔一枚硬币的结果，并不会影响平均差错概率。

这里考虑两种可能的错误：

1. 实际发送为 0 而判决为 1，我们把这类差错称为第一类差错。
2. 实际发送为 1 而判决为 0，我们把这类差错称为第二类差错。

为了计算平均差错概率，我们分别考虑这两种情况。

假设发送符号 0。那么，根据式(4.21)，接收信号为

$$x(t) = -A + w(t), \quad 0 \leq t \leq T_b \quad (4.22)$$

相应地，匹配滤波器在抽样时刻 $t = T_b$ 的输出为（为了表达方便，设例 4.1 中的 kAT_b 等于 1）：

$$\begin{aligned} y &= \int_0^{T_b} x(t) dt \\ &= -A + \frac{1}{T_b} \int_0^{T_b} w(t) dt \end{aligned} \quad (4.23)$$

这代表了随机变量 Y 的抽样值。由于噪声 $w(t)$ 为高斯白噪声，因此可将随机变量 Y 的特性描述如下：

- 随机变量 Y 服从均值为 $-A$ 的高斯分布。
- 随机变量 Y 的方差为

$$\begin{aligned} \sigma_Y^2 &= E[(Y + A)^2] \\ &= \frac{1}{T_b^2} E\left[\int_0^{T_b} \int_0^{T_b} w(t)w(u) dt du\right] \\ &= \frac{1}{T_b^2} \int_0^{T_b} \int_0^{T_b} E[w(t)w(u)] dt du \\ &= \frac{1}{T_b^2} \int_0^{T_b} \int_0^{T_b} R_w(t, u) dt du \end{aligned} \quad (4.24)$$

其中, $R_w(t, u)$ 为白噪声 $w(t)$ 的自相关函数。由于 $w(t)$ 是功率谱密度为 $N_0/2$ 的白噪声, 所以

$$R_w(t, u) = \frac{N_0}{2} \delta(t - u) \quad (4.25)$$

其中, $\delta(t - u)$ 是一个时移的 delta 函数。将式(4.25)代入式(4.24)得到

$$\begin{aligned} \sigma_Y^2 &= \frac{1}{T_b^2} \int_0^{T_b} \int_0^{T_b} \frac{N_0}{2} \delta(t - u) dt du \\ &= \frac{N_0}{2T_b} \end{aligned} \quad (4.26)$$

此处利用了 delta 函数的延时特性以及其面积为 1 的性质。已知发送符号为 0, 则随机变量 Y 的条件概率密度函数为

$$f_Y(y|0) = \frac{1}{\sqrt{\pi N_0/T_b}} \exp\left(-\frac{(y + A)^2}{N_0/T_b}\right) \quad (4.27)$$

如图 4.5(a)所示。设 p_{10} 表示已知发送符号为 0 时的条件差错概率。该概率值由图中从门限值 λ 到正无穷大时曲线 $f_Y(y|0)$ 下方的阴影区域所表示, 对应于判决为符号 1 时 y 的取值范围。当不考虑噪声时, 匹配滤波器的输出在时刻 $t = T_b$ 的抽样值 y 等于 $-A$ 。当存在噪声时, y 偶尔会大于 λ , 此时将出错。在发送符号为 0 的条件下, 该差错概率定义为

$$\begin{aligned} p_{10} &= P(y > \lambda | \text{发送符号 } 0) \\ &= \int_{\lambda}^{\infty} f_Y(y|0) dy \\ &= \frac{1}{\sqrt{\pi N_0/T_b}} \int_{\lambda}^{\infty} \exp\left(-\frac{(y + A)^2}{N_0/T_b}\right) dy \end{aligned} \quad (4.28)$$

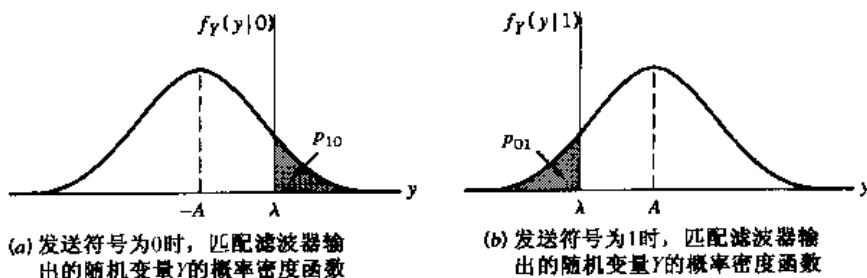


图 4.5 PCM 系统噪声分析

下面, 介绍一下所谓的互补误差函数:^[3]

$$\operatorname{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^{\infty} \exp(-z^2) dz \quad (4.29)$$

该函数与高斯分布密切相关。对一个较大的正数 u 来说, 互补误差函数的上界为

$$\operatorname{erfc}(u) < \frac{\exp(-u^2)}{\sqrt{\pi u}} \quad (4.30)$$

为了用互补误差函数来表示条件差错概率 p_{10} , 先定义一个新的变量

$$z = \frac{y + A}{\sqrt{N_0/T_b}}$$

因此,可将式(4.28)写为紧凑形式:

$$\begin{aligned} p_{10} &= \frac{1}{\sqrt{\pi}} \int_{(A+\lambda)/\sqrt{N_0/T_b}}^{\infty} \exp(-z^2) dz \\ &= \frac{1}{2} \operatorname{erfc}\left(\frac{A + \lambda}{\sqrt{N_0/T_b}}\right) \end{aligned} \quad (4.31)$$

接下来假设发送符号1。此时,由匹配滤波器输出样值 y 所代表的高斯随机变量 Y 具有均值 $+A$,方差 $N_0/2T_b$ 。需注意,与发送符号0时的情况相比,随机变量 Y 的均值改变了,但其方差与前面相同。因此,在发送符号1时, Y 的条件概率密度函数为

$$f_Y(y|1) = \frac{1}{\sqrt{\pi N_0/T_b}} \exp\left(-\frac{(y - A)^2}{N_0/T_b}\right) \quad (4.32)$$

如图4.5(b)所示。设 p_{01} 表示已知发送符号为1时的条件差错概率。该概率值由图中从负无穷大到门限值 λ 的曲线 $f_Y(y|1)$ 下的阴影区域所表示,对应于判决为符号0时的 y 的取值范围。当不考虑噪声时,匹配滤波器的输出在时刻 $t = T_b$ 的抽样值 y 就等于 $+A$ 。当存在噪声时, y 偶尔会小于 λ ,此时将出错。在发送符号为1的条件下,该差错概率定义为

$$\begin{aligned} p_{01} &= P(y < \lambda | \text{发送符号 } 1) \\ &= \int_{-\infty}^{\lambda} f_Y(y|1) dy \\ &= \frac{1}{\sqrt{\pi N_0/T_b}} \int_{-\infty}^{\lambda} \exp\left(-\frac{(y - A)^2}{N_0/T_b}\right) dy \end{aligned} \quad (4.33)$$

为了用互补误差函数来表示 p_{01} ,定义一个新变量

$$z = \frac{A - y}{\sqrt{N_0/T_b}}$$

因此,可将式(4.33)写为紧凑形式:

$$\begin{aligned} p_{01} &= \frac{1}{\sqrt{\pi}} \int_{(A-\lambda)/\sqrt{N_0/T_b}}^{\infty} \exp(-z^2) dz \\ &= \frac{1}{2} \operatorname{erfc}\left(\frac{A - \lambda}{\sqrt{N_0/T_b}}\right) \end{aligned} \quad (4.34)$$

确定了条件差错概率 p_{10} 和 p_{01} 后,下一个任务就是推导平均符号差错概率的计算公式,记为 P_e 。值得注意的是,这两种可能的差错是互斥事件,因为接收机在特定的抽样瞬间,如果选择符号1,就不会出现符号0,反之亦然。设 p_0, p_1 分别代表发送符号为0和1的先验概率。因此,接收机的平均符号差错概率 P_e 可由下式给出:

$$\begin{aligned} P_e &= p_0 p_{10} + p_1 p_{01} \\ &= \frac{p_0}{2} \operatorname{erfc}\left(\frac{A + \lambda}{\sqrt{N_0/T_b}}\right) + \frac{p_1}{2} \operatorname{erfc}\left(\frac{A - \lambda}{\sqrt{N_0/T_b}}\right) \end{aligned} \quad (4.35)$$

由式(4.35)可见, P_e 实际上是门限值 λ 的函数,这就意味着需要求出最优点限值,使得 P_e

最小。为此使用莱布尼兹准则(Leibniz's rule)。

考虑积分

$$\int_{a(u)}^{b(u)} f(z, u) dz$$

莱布尼兹准则表明,这个积分函数对 u 的导数为

$$\frac{d}{du} \int_{a(u)}^{b(u)} f(z, u) dz = f(b(u), u) \frac{db(u)}{du} - f(a(u), u) \frac{da(u)}{du} + \int_{a(u)}^{b(u)} \frac{\partial f(z, u)}{\partial u} dz$$

此时,由式(4.29)的互补误差函数定义可得

$$\begin{aligned} f(z, u) &= \frac{2}{\sqrt{\pi}} \exp(-z^2) \\ a(u) &= u \\ b(u) &= \infty \end{aligned}$$

将莱布尼兹准则应用于互补误差函数,得

$$\frac{d}{du} \operatorname{erfc}(u) = -\frac{1}{\sqrt{\pi}} \exp(-u^2) \quad (4.36)$$

应用式(4.36),将式(4.35)对 λ 求导,然后令结果为零并化简,得到最优门限为

$$\lambda_{\text{opt}} = \frac{N_0}{4AT_b} \log\left(\frac{p_0}{p_1}\right) \quad (4.37)$$

对于符号 1 和 0 等概率的特例,有

$$p_1 = p_0 = \frac{1}{2}$$

则式(4.37)变为

$$\lambda_{\text{opt}} = 0$$

这个结果与直观情况是相符的。该式说明,对于等概率二进制符号的传输,我们应当选择两个符号 0 和 1 的脉冲峰值 $-A$ 和 $+A$ 的中点为门限值。注意,在这个特例中,还有

$$p_{01} = p_{10}$$

条件差错概率 p_{10} 和 p_{01} 相等的信道被称为是二进制对称的。因此,式(4.35)中的平均符号差错概率简化为

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\frac{A}{\sqrt{N_0/T_b}}\right) \quad (4.38)$$

则每比特发射信号能量可定义为

$$E_b = A^2 T_b \quad (4.39)$$

因此,我们可以最终给出图 4.4 所示接收机的平均符号差错概率为

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \quad (4.40)$$

这说明一个二进制对称信道中的平均符号差错概率仅取决于每比特发射信号的能量与噪声功

率谱密度之比 E_b/N_0 。

根据式(4.30)中互补误差函数的上界,我们可相应地将 PCM 接收机的平均符号差错概率界定为

$$P_e < \frac{\exp(-E_b/N_0)}{2\sqrt{\pi E_b/N_0}} \quad (4.41)$$

因此,图 4.4 所示的 PCM 接收机表明了随着 E_b/N_0 的增加,平均符号差错概率会有一个指数级的改善。

这个重要结论在图 4.6 中得到进一步说明。该图描绘了平均符号差错概率 P_e 随无量纲比值 E_b/N_0 的变化特性。特别地,当比值 E_b/N_0 增加时, P_e 迅速下降。因此,最终发射信号能量的“非常小的增加”将使得二进制脉冲的接收几乎没有差错,这正如前面在 3.8 节中讨论过的一样。值得注意的是,在实际应用中,必须相对地看待信号能量的增加。例如,在 E_b 较小时,要使 E_b/N_0 增加 3 dB,就比 E_b 较大时容易实现得多。

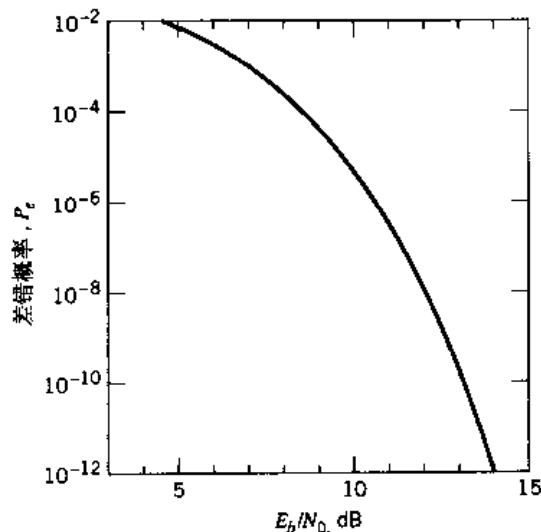


图 4.6 PCM 接收机中的差错概率

4.4 码间干扰

在基带脉冲传输系统中,另一个造成比特差错的原因就是 ISI。ISI 在通信信道呈弥散特性时就会出现。这里首先需要强调一个关键问题:已知一个脉冲波形,怎样才能利用它以 M 进制形式传输数据?答案是使用离散脉冲调制。传输脉冲的幅度、宽度以及位置都根据要传送的数据流以离散方式变化。对于数字数据的基带传输,就功率和带宽利用率来说,离散 PAM 是最有效的方法之一。因此,我们讨论离散 PAM 系统。首先考虑二进制数据的情况。本章的后面部分中,将考虑更常用的数据情况。

图 4.7 给出了基带二进制 PAM 系统的一般形式。输入二进制序列 $\{b_k\}$ 由符号 1 和 0 组成,每个符号的持续时间都是 T_b 。脉冲幅度调制器将这组二进制序列转换为一组新的窄脉冲(近似为单位脉冲)序列,其幅度 a_k 以极性形式表达:

$$a_k = \begin{cases} +1 & b_k = 1 \\ -1 & b_k = 0 \end{cases} \quad (4.42)$$

将生成的窄脉冲序列送至冲激响应为 $g(t)$ 的发射滤波器中, 得到发射信号

$$s(t) = \sum_k a_k g(t - kT_b) \quad (4.43)$$

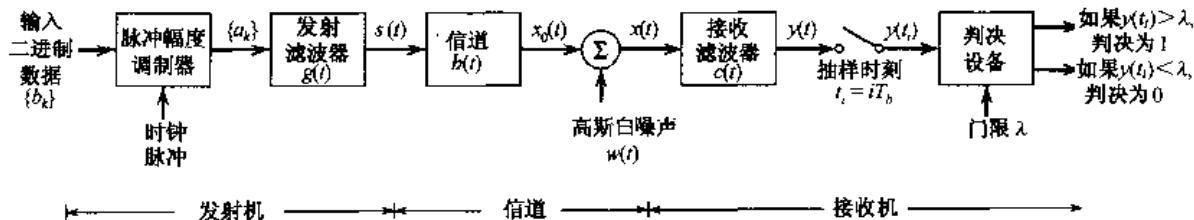


图 4.7 基带二进制数据传输系统

信号 $s(t)$ 在通过冲激响应为 $b(t)$ 的信道传输后发生改变。此外, 在接收机输入端, 信道也将随机噪声带入到了信号中。接着, 含噪声的信号 $x(t)$ 通过冲激响应为 $c(t)$ 的接收滤波器。对接收滤波器的输出 $y(t)$ 按与发射机同步的方式进行抽样, 抽样时刻由从接收滤波器输出信号中提取的定时信号或时钟来决定。最后, 所获得的抽样序列通过判决器来重构原始数据序列。具体地说, 每个抽样值的幅度都要与门限 λ 进行比较。如果抽样值幅度超过门限 λ , 判决器就判为 1, 反之则为 0。如果抽样值幅度等于门限值, 则像抛硬币一样, 由接收机简单地进行随机猜测来判决发送的符号。

接收滤波器的输出记为

$$y(t) = \mu \sum_k a_k p(t - kT_b) + n(t) \quad (4.44)$$

其中, μ 为比例因子, 脉冲 $p(t)$ 如定义。为了进行准确的描述, 式(4.44)中的脉冲 $p(t - kt_b)$ 的自变量中应包含时延 t_0 (t_0 为任意值), t_0 代表脉冲通过系统的传输时延。为了简化表达并不失一般性, 在式(4.44)中, 将时延 t_0 设为零。

带有比例因子的脉冲 $\mu p(t)$ 由发射滤波器的冲激响应 $g(t)$, 信道的冲激响应 $b(t)$ 和接收滤波器的冲激响应 $c(t)$ 进行两次卷积而成, 即

$$\mu p(t) = g(t) \star b(t) \star c(t) \quad (4.45)$$

其中 \star 代表卷积。对脉冲 $p(t)$ 进行归一化假设:

$$p(0) = 1 \quad (4.46)$$

因此, 可将 μ 作为比例因子来说明信号通过系统传输后发生的幅度变化。

由于时域中的卷积对应于频域中的乘法, 因此可用傅里叶变换将式(4.45)转变为等价形式:

$$\mu P(f) = G(f)H(f)C(f) \quad (4.47)$$

其中, $P(f)$, $G(f)$, $H(f)$ 和 $C(f)$ 分别为 $p(t)$, $g(t)$, $b(t)$ 和 $c(t)$ 的傅里叶变换。

式(4.44)中的 $n(t)$ 是由信道噪声 $w(t)$ 所带来的接收滤波器输出端的噪声。习惯上认为 $w(t)$ 为零均值的高斯白噪声。

对接收滤波器输出信号 $y(t)$ 在时刻 $t_i = iT_b$ 进行抽样(i 取整数), 根据式(4.46), 得

$$\begin{aligned}
 y(t_i) &= \mu \sum_{k=-\infty}^{\infty} a_k p[(i-k)T_b] + n(t_i) \\
 &= \mu a_i + \mu \sum_{\substack{k=-\infty \\ k \neq i}}^{\infty} a_k p[(i-k)T_b] + n(t_i)
 \end{aligned} \tag{4.48}$$

上式第一项 μa_i 代表了第 i 个传输比特的对应值。第二项代表了对第 i 个比特进行译码时所有其他传输比特的残留影响。这种由抽样时刻 t_i 前后的脉冲产生的残留影响就是 ISI。最后一项 $n(t_i)$ 代表了 t_i 时刻的噪声抽样。

若不考虑 ISI 和噪声的影响, 可从式(4.48)中得到

$$y(t_i) = \mu a_i$$

这说明在理想情况下, 第 i 个传输比特得到了正确译码。然而, 系统中不可避免地存在着 ISI 和噪声, 从而在接收机输出端的判决器中引入了误差。因此, 在设计发射和接收滤波器时, 目的就是使噪声和 ISI 的影响最小化, 从而以最小的差错率将数字数据传送到目的地。

当信噪比较高时(如在电话系统中), 系统性能更多地受到 ISI 而不是噪声的影响。换言之, 此时可以忽略 $n(t_i)$ 。在接下来的几节中, 我们假设此条件成立, 以便能将注意力集中到 ISI 以及相应的控制技术上。特别地, 我们最希望的是得到能够完全消除 ISI 的脉冲波形 $p(t)$ 。

4.5 无失真基带二进制传输的奈奎斯特准则

信道的频率响应和发送脉冲波形通常都是给定的, 问题在于如何确定发射和接收滤波器的频率响应, 以便重构原始二进制数据序列 $\{b_k\}$ 。接收机是通过从输出信号 $y(t)$ 中提取并译码出相应的系数序列 $\{a_k\}$ 来实现这一点的。这种提取包括了对输出 $y(t)$ 在时刻 $t = iT_b$ 进行抽样。译码过程要求加权脉冲 $a_k p(iT_b - kT_b)$ 在 $k = i$ 时不受由 $k \neq i$ 的其他加权脉冲的尾部重叠所带来的 ISI 的影响。这就要求我们能够控制整个脉冲 $p(t)$, 即

$$p(iT_b - kT_b) = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \tag{4.49}$$

其中, 归一化 $p(0) = 1$ 。如果 $p(t)$ 满足式(4.49)的条件, 则式(4.48)中的接收机输出 $y(t_i)$ 可简化为(忽略噪声项):

$$y(t_i) = \mu a_i \quad \text{对于所有 } i$$

这意味着没有 ISI。因此, 式(4.49)的两个条件保证了无噪声下的理想接收。

从设计角度讲, 将式(4.49)映射到频域是很有意义的。接下来, 我们考虑样值序列 $\{p(nT_b)\}$, 其中 $n = 0, \pm 1, \pm 2, \dots$ 。由第 3 章中有关抽样过程的讨论可知, 时域上的抽样会产生频域上的周期信号。因此有

$$P_\delta(f) = R_b \sum_{n=-\infty}^{\infty} P(f - nR_b) \tag{4.50}$$

其中, $R_b = 1/T_b$ 为以 bps 为单位的比特率。 $P_\delta(f)$ 是周期为 T_b 的 delta 函数的无限周期序列的傅里叶变换。 $P_\delta(f)$ 在每个区间上都用 $p(t)$ 的对应样值进行了加权。即

$$P_s(f) = \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} [p(mT_b) \delta(t - mT_b)] \exp(-j2\pi ft) dt \quad (4.51)$$

设整数 $m = i - k$, 则 $i = k$ 对应于 $m = 0$ 。同样地, $i \neq k$ 对应于 $m \neq 0$ 。将式(4.49)的条件应用于式(4.51)的积分样值 $p(t)$, 有

$$P_s(f) = \int_{-\infty}^{\infty} p(0) \delta(t) \exp(-j2\pi ft) dt = p(0) \quad (4.52)$$

其中用到了 delta 函数的延时特性。由式(4.46)可得 $p(0) = 1$ 。因此, 由式(4.50)和式(4.52)可以得出结论: 当下式成立时, 系统满足无 ISI 的条件:

$$\sum_{n=-\infty}^{\infty} P(f - nR_b) = T_b \quad (4.53)$$

由此可以阐述在无噪声时的无失真基带传输的奈奎斯特准则^[4]: 如果频率函数 $P(f)$ 满足式(4.53), 那么就能消除以时间间隔 T_b 进行抽样的 ISI。注意 $P(f)$ 对应的是整个系统, 包括对应于式(4.47)的发射滤波器、信道和接收滤波器。

4.5.1 理想奈奎斯特信道

满足式(4.53)的最简单办法是规定频率函数 $P(f)$ 为矩形函数的形式, 即

$$P(f) = \begin{cases} \frac{1}{2W}, & -W < f < W \\ 0, & |f| > W \end{cases} \quad (4.54)$$

$$= \frac{1}{2W} \text{rect}\left(\frac{f}{2W}\right)$$

其中, $\text{rect}(f)$ 代表一个以 $f = 0$ 为中心的单位幅度单位持续时间的矩形函数, 整个系统带宽 W 定义为

$$W = \frac{R_b}{2} = \frac{1}{2T_b} \quad (4.55)$$

根据式(4.54)和式(4.55)所描述的解决方案, 绝对值超过比特率一半的频率就不需要了。由表 A6.3 中的傅里叶变换对 2 可知, 无 ISI 时的信号波形可以由 sinc 函数定义:

$$p(t) = \frac{\sin(2\pi Wt)}{2\pi Wt} \quad (4.56)$$

$$= \text{sinc}(2Wt)$$

比特率的特殊值 $R_b = 2W$ 称为奈奎斯特速率, W 称为奈奎斯特带宽。相应地, 由式(4.54)描述的频域中的理想基带脉冲传输系统, 或由式(4.56)描述的与之等价的时域系统, 就称为理想奈奎斯特信道。

图 4.8(a)和图 4.8(b)分别为 $P(f)$ 和 $p(t)$ 的图形。图 4.8(a)分别在正、负频率处给出了频率函数 $P(f)$ 的归一化形式。图 4.8(b)包括了信号间隔以及相应的以抽样时刻为中心的波形。函数 $p(t)$ 可视为通带幅度响应为 $1/2W$, 带宽为 W 的理想低通滤波器的冲激响应。函数 $p(t)$ 在原点有峰值, 且在比特持续时间 T_b 的整数倍时通过零点。显然, 如果对接收波形 $y(t)$

在时刻 $t = 0, \pm T_b, \pm 2T_b, \dots$ 上进行抽样，则由 $\mu p(t - iT_b)$ 定义的脉冲彼此间将不存在干扰。其中， μ 为任意幅度， $i = 0, \pm 1, \pm 2, \dots$ 为下标。图 4.9 以二进制序列 1011010 为例说明了这个结论。

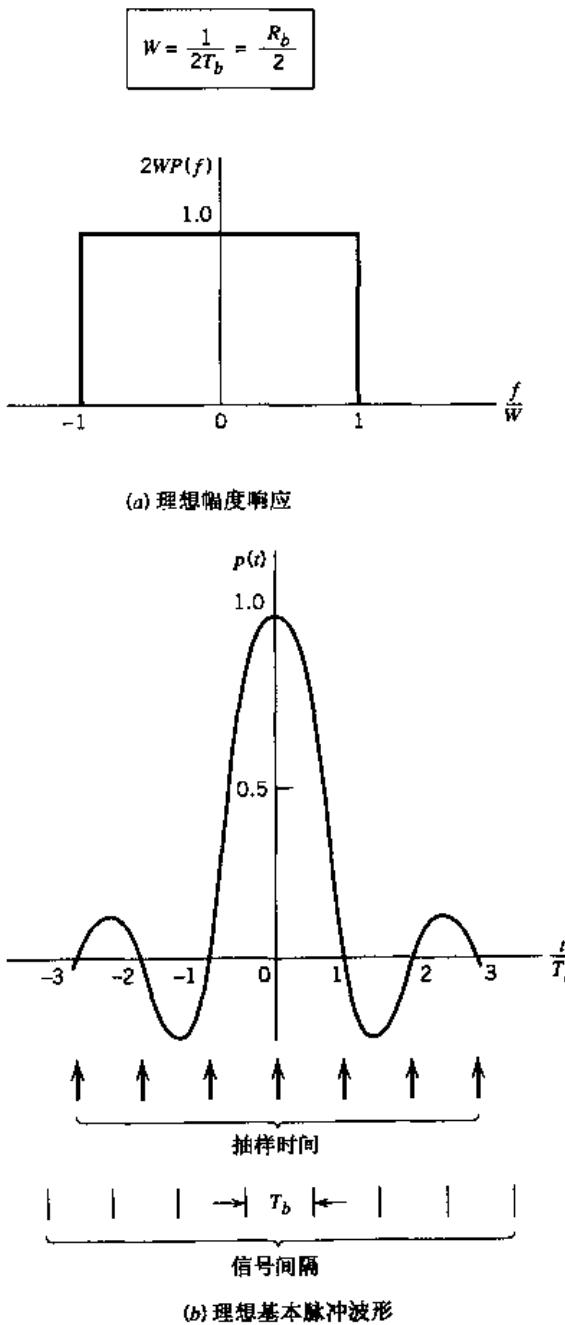


图 4.8 理想幅度响应和脉冲波形

采用理想奈奎斯特信道，以最小可能带宽实现了无 ISI，从而达到了节约带宽的目的。但其中存在的两个实际困难，成为系统设计中一个不可实现的目标：

1. 采用这种方法，要求 $P(f)$ 的幅度特性从 $-W$ 到 W 是平坦的，而在其他地方为零。即在边界频率 $\pm W$ 上存在突变，这在物理上是不可实现的。

2. 当 $|t|$ 较大时, 函数 $p(t)$ 以 $1/|t|$ 减少, 衰减速率缓慢。这也是由 $P(f)$ 在 $\pm W$ 上的不连续性造成的。因此, 实际上在接收机的抽样时刻并没有错误容限。

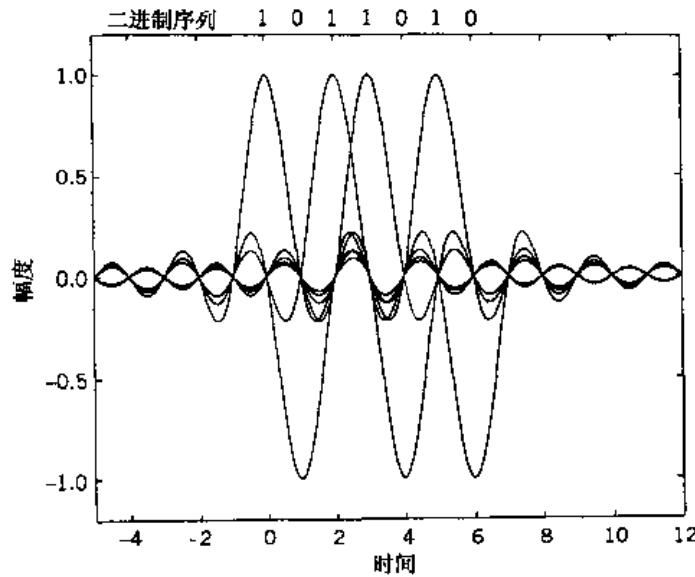


图 4.9 对应于序列 1011010 的 sinc 脉冲序列

为了估算这种定时误差的影响, 考虑在时刻 $t = \Delta t$ 的 $y(t)$ 样值。其中, Δt 为定时误差。为简便起见, 设正确抽样时刻 t_i 为零。在不考虑噪声的情况下, 有(由式(4.48)):

$$\begin{aligned} y(\Delta t) &= \mu \sum_k a_k p(\Delta t - kT_b) \\ &= \mu \sum_k a_k \frac{\sin[2\pi W(\Delta t - kT_b)]}{2\pi W(\Delta t - kT_b)} \end{aligned} \quad (4.57)$$

由定义 $2WT_b = 1$, 可将式(4.57)写为

$$y(\Delta t) = \mu a_0 \operatorname{sinc}(2W\Delta t) + \frac{\mu \sin(2\pi W\Delta t)}{\pi} \sum_{k \neq 0} \frac{(-1)^k a_k}{(2W\Delta t - k)} \quad (4.58)$$

上式右边第一项定义了所需的符号, 第二项的级数代表对输出信号 $y(t)$ 抽样时由于定时误差 Δt 所带来的 ISI。遗憾的是, 由于无法消除这个级数, 造成了接收机的判决错误。

4.5.2 升余弦频谱

通过将信道带宽从最小值 $W = R_b/2$ 拓展到界于 W 和 $2W$ 之间的可调值, 可有效克服理想奈奎斯特信道所遇到的实际困难。假定整个频率响应 $P(f)$ 满足一个比理想奈奎斯特信道更为严格的条件。具体地, 保留式(4.53)中的三项并将频率带宽限制在 $[-W, W]$, 表示为

$$P(f) + P(f - 2W) + P(f + 2W) = \frac{1}{2W}, \quad -W \leq f \leq W \quad (4.59)$$

不难找到满足式(4.59)的带限函数。升余弦频谱就是其中一个。该频率响应 $p(f)$ 由一个平坦部分和一个有正弦形式的滚降部分组成, 即

$$P(f) = \begin{cases} \frac{1}{2W}, & 0 \leq |f| < f_1 \\ \frac{1}{4W} \left\{ 1 - \sin \left[\frac{\pi(|f| - W)}{2W - 2f_1} \right] \right\}, & f_1 \leq |f| < 2W - f_1 \\ 0, & |f| \geq 2W - f_1 \end{cases} \quad (4.60)$$

频率参数 f_1 与带宽 W 之间有下述关系:

$$\alpha = 1 - \frac{f_1}{W} \quad (4.61)$$

参数 α 称为滚降因子, 表示理想带宽 W 的额外带宽。传输带宽 B_T 的定义为

$$\begin{aligned} B_T &= 2W - f_1 \\ &= W(1 + \alpha) \end{aligned} \quad (4.61)$$

将 α 的三个值 0, 0.5 和 1 与 $2W$ 相乘可以进行归一化。归一化后的频率响应 $P(f)$ 如图 4.10(a) 所示。由图可知, 当 $\alpha = 0.5$ 或 1 时, 与理想奈奎斯特信道(即 $\alpha = 0$)相比较而言, 函数 $P(f)$ 是逐渐截止的, 因而更易于实现。函数 $P(f)$ 关于奈奎斯特带宽 W 呈奇对称, 因此能够满足式(4.59)的条件。

时间响应 $p(t)$ 是频率响应 $P(f)$ 的傅里叶反变换。利用式(4.60)定义的 $P(f)$, 可得(见习题 4.13):

$$p(t) = (\text{sinc}(2Wt)) \left(\frac{\cos(2\pi\alpha Wt)}{1 - 16\alpha^2 W^2 t^2} \right) \quad (4.62)$$

图 4.10(b) 为 $\alpha = 0, 0.5$ 和 1 时的情况。

时间响应 $p(t)$ 由两个因子的乘积组成: 描述理想奈奎斯特信道的因子 $\text{sinc}(2Wt)$, 对较大 $|t|$ 值以 $1/|t|^2$ 衰减的因子。第一个因子保证了 $p(t)$ 时刻 $t = iT$ 对应的期望抽样瞬时通过零点, 其中, i 取整数(正或负)。第二个因子使该脉冲的尾部衰减大大低于从理想奈奎斯特信道获得的尾部衰减, 因此, 采用这种脉冲进行的二进制波形传输对抽样定时误差就不那么敏感。实际上, 当 $\alpha = 1$ 时, 由于 $p(t)$ 尾部振荡幅度最小, 滚降也最为平缓。因此, 随着滚降因子 α 从 0 增大到 1, 由定时误差产生的 ISI 值就会逐渐减小。

$\alpha = 1$ (即 $f_1 = 0$) 时的特例称为完全余弦滚降特性, 此时式(4.60)的频率响应可化简为

$$P(f) = \begin{cases} \frac{1}{4W} \left[1 + \cos \left(\frac{\pi f}{2W} \right) \right], & 0 < |f| < 2W \\ 0, & |f| \geq 2W \end{cases} \quad (4.63)$$

相应地, 时间响应 $p(t)$ 可化简为

$$p(t) = \frac{\text{sinc}(4Wt)}{1 - 16W^2 t^2} \quad (4.64)$$

时间响应具有如下两个性质:

- 在 $t = \pm T_b/2 = \pm 1/4W$ 时, 有 $p(t) = 0.5$ 。这即是说, 在脉冲的 1/2 幅度处测量得到的脉冲宽度正好等于比特持续时间 T_b 。

2. 除了抽样时刻为 $t = \pm T_b, \pm 2T_b, \dots$ 的这些通常的过零点外, 在时刻 $t = \pm 3T_b/2, \pm 5T_b/2, \dots$ 也有过零点。

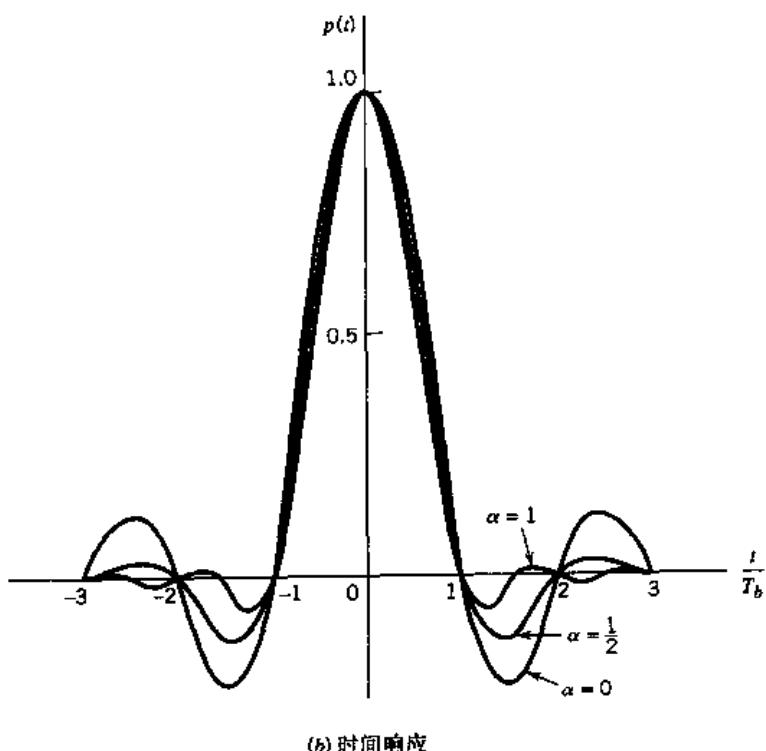
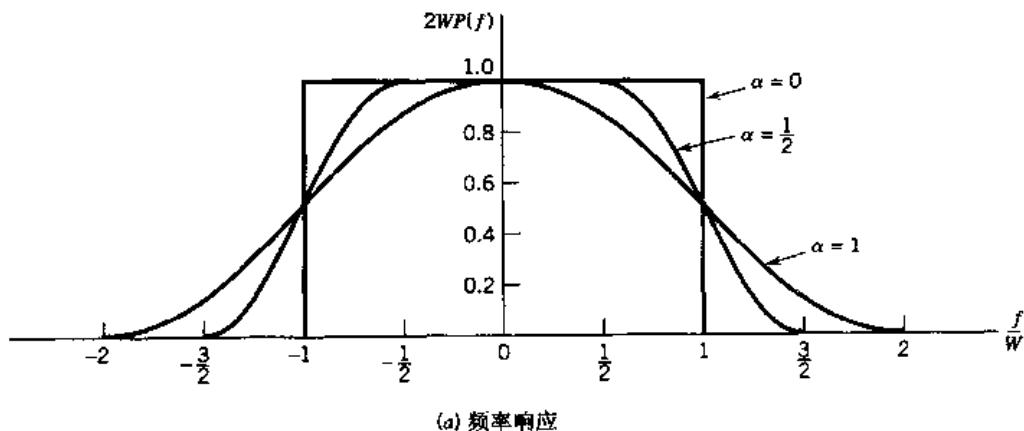


图 4.10 对于不同滚降因子的响应

当需要从接收信号中提取定时信号以进行同步时,这两条性质就非常有用。然而,为了获得这种性质而付出的代价就是,所需的信道带宽是对应于 $\alpha = 0$ 的理想奈奎斯特信道带宽的两倍。

例 4.2 T1 系统的带宽要求

第 3 章的例 3.2 中给出了 T1 载波系统的信号形式。该系统用于复用基于 8 比特 PCM 码字的 24 路独立语音输入。最终的时分复用信号(包括一个帧比特)的比特持续时间为

$$T_b = 0.647 \mu\text{s}$$

假设采用理想奈奎斯特信道，则 T1 系统的最小传输带宽 B_T 为 ($\alpha = 0$)：

$$B_T = W = \frac{1}{2T_b} = 772 \text{ kHz}$$

利用 $\alpha = 1$ 的完全余弦滚降特性，可以得到传输带宽的更为现实的数值。由这个例子，可发现

$$B_T = W(1 + \alpha) = 2W = \frac{1}{T_b} = 1.544 \text{ MHz}$$

4.6 相关电平编码

ISI 的存在往往导致系统性能的恶化。因此，ISI 问题成为每个系统设计者都必须重点考虑的问题。通过以可控方式将 ISI 引入发射信号中，就有可能在一个带宽为 W 赫的信道中，使信号速率达到每秒 $2W$ 个符号（即奈奎斯特速率）。这种方法称为相关电平编码或部分响应技术^[5]。这种方法的设计基于如下前提：既然引入到发射信号中的 ISI 是已知的，则其影响就可以在接收机中给予确定的描述。采用抗扰动滤波器的相关电平编码，可以看成是在带宽为 W 赫时获得奈奎斯特速率的一种实用方法。

4.6.1 双二进制信号

下面通过讨论双二进制信号的具体例子来阐述相关电平编码的基本思想。其中“双”代表对一个直接的二进制系统的传输容量加倍。这种相关电平编码的特定形式也称为第 I 类部分响应。

假设有一个二进制输入序列 $\{b_k\}$ ，由不相关的二进制符号 1 和 0 组成，每个符号的持续时间为 T_b 。和前面一样，使该序列通过一个能够生成 2 电平窄脉冲序列（近似于一个单位脉冲）的脉冲幅度调制器，该脉冲序列的幅度 a_k 定义为

$$a_k = \begin{cases} +1 & \text{当 } b_k = 1 \\ -1 & \text{当 } b_k = 0 \end{cases} \quad (4.65)$$

使该序列通过双二进制编码器，得到一个 3 电平输出。即 -2, 0 和 +2。这种变换可通过采用图 4.11 所示的方案来实现。2 电平序列 $\{a_k\}$ 首先通过带有一个延迟器和求和器的简单滤波器。对于每个加到该滤波器输入端的单位脉冲，都能在滤波器输出端得到两个间隔 T_b 秒的单位脉冲。可将双二进制编码器的输出 c_k 表示为当前输入脉冲 a_k 和其前一个输入脉冲 a_{k-1} 之和，即

$$c_k = a_k + a_{k-1} \quad (4.66)$$

这种转变的一个作用就是把不相关的 2 电平脉冲输入序列 $\{a_k\}$ 变为了相关的 3 电平脉冲序列 $\{c_k\}$ 。这种相邻脉冲间的相关性可以看成是以人为的方式将 ISI 引入到了发射信号中。但通过这种方式引入的 ISI 是在设计者的控制之下的，这就是相关编码的基础。

产生 T_b 秒时延的理想延迟单元具有频率响应 $\exp(-j2\pi f T_b)$ ，因此图 4.11 所示简单延迟滤波器的频率响应为 $1 + \exp(-j2\pi f T_b)$ 。该滤波器与一个理想奈奎斯特信道级联后的整体

频率响应则为

$$\begin{aligned} H_I(f) &= H_{\text{Nyquist}}(f)[1 + \exp(-j2\pi f T_b)] \\ &= H_{\text{Nyquist}}(f)[\exp(j\pi f T_b) + \exp(-j\pi f T_b)] \exp(-j\pi f T_b) \\ &= 2H_{\text{Nyquist}}(f) \cos(\pi f T_b) \exp(-j\pi f T_b) \end{aligned} \quad (4.67)$$

其中, $H_I(f)$ 的下标 I 表示部分响应的相关类型。对于一个带宽为 $W = 1/2T_b$ 的理想奈奎斯特信道, 有(忽略比例因子 T_b):

$$H_{\text{Nyquist}}(f) = \begin{cases} 1, & |f| \leq 1/2T_b \\ 0, & \text{其他} \end{cases} \quad (4.68)$$

因此, 双二进制信号方案的整个频率响应具有半周期余弦函数的形式, 即

$$H_I(f) = \begin{cases} 2 \cos(\pi f T_b) \exp(-j\pi f T_b), & |f| \leq 1/2T_b \\ 0, & \text{其他} \end{cases} \quad (4.69)$$

其幅度响应和相位响应分别如图 4.12(a) 和图 4.12(b) 所示。这种频率响应的一个优点是由于其在频带边缘是连续的, 因此在实际应用中易于进行近似。

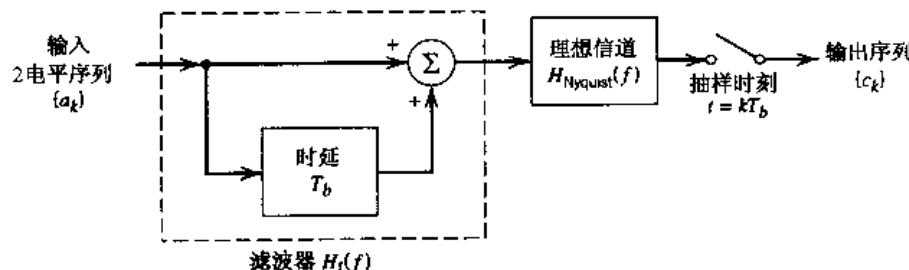


图 4.11 双二进制信号方案

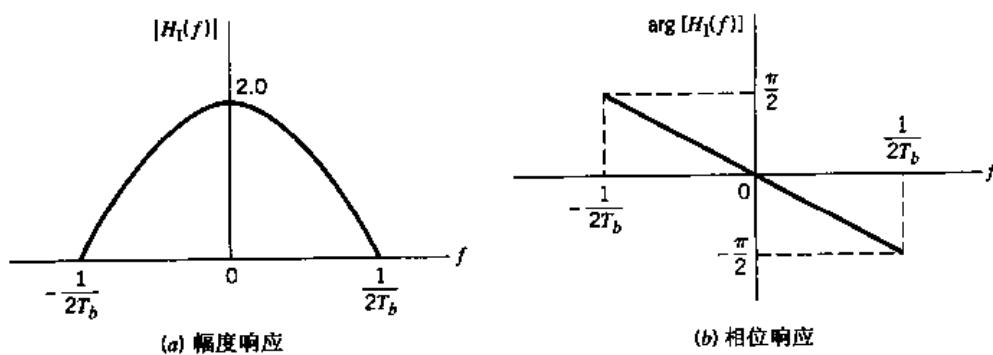


图 4.12 双二进制变换滤波器的频率响应

由式(4.67)的第一行和式(4.68)中 $H_{\text{Nyquist}}(f)$ 的定义, 不难发现对应于频率响应 $H_I(f)$ 的冲激响应由两个间隔 T_b 秒的 sinc(奈奎斯特)脉冲组成, 即(其中去掉了比例因子):

$$\begin{aligned} h_I(t) &= \frac{\sin(\pi t/T_b)}{\pi t/T_b} + \frac{\sin[\pi(t-T_b)/T_b]}{\pi(t-T_b)/T_b} \\ &= \frac{\sin(\pi t/T_b)}{\pi t/T_b} - \frac{\sin(\pi t/T_b)}{\pi(t-T_b)/T_b} \\ &= \frac{T_b^2 \sin(\pi t/T_b)}{\pi t(T_b-t)} \end{aligned} \quad (4.70)$$

冲激响应 $b_1(t)$ 如图 4.13 所示,由图可见,在抽样瞬间 $b_1(t)$ 只有两个可分辨的值。图 4.13 解释了将这种相关编码类型称为部分响应信号的原因,即在这种编码方法中,输入脉冲的响应覆盖范围大于一个信号间隔。换言之,在任何一个信号间隔中的响应都是“部分”的。值得注意的是, $b_1(t)$ 的尾部衰减为 $1/(|t|)^2$,这比理想奈奎斯特信道中的衰减速率 $1/|t|$ 更快。

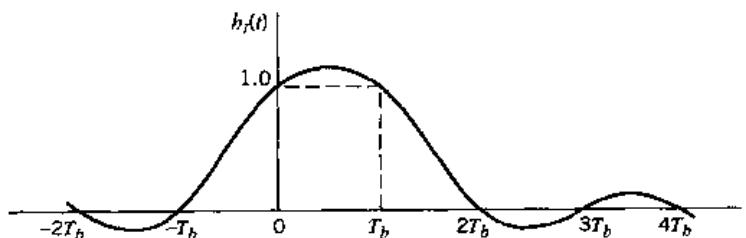


图 4.13 双二进制变换滤波器的冲激响应

原始的 2 电平序列 $\{a_k\}$ 可利用式(4.66)从双二进制编码序列 $\{c_k\}$ 中检测得到。假定 \hat{a}_k 代表接收机在时刻 $t = kT_b$ 对原始脉冲 a_k 的估计值。从 c_k 中减去前一个估计值 \hat{a}_{k-1} , 得

$$\hat{a}_k = c_k - \hat{a}_{k-1} \quad (4.71)$$

显然,如果 c_k 接收无差错,且在时刻 $t = (k-1)T_b$ 的前一个估计值 \hat{a}_{k-1} 对应于一个正确的判决,则当前估计值 \hat{a}_k 也将是正确的。这种利用先前符号存储估计值来得到当前符号估计值的技术称为判决反馈。

上述判决过程实质上是发射机中简单延迟滤波器的逆操作。该判决过程的一个主要缺点是,一旦出错,错误将在输出中不断传播,因为对当前输入 a_k 的判决取决于先前输入 a_{k-1} 判决的正确性。

一个避免错误传播现象的实用方法是在双二进制编码之前采用预编码,如图 4.14 所示。对二进制数据序列 $\{b_k\}$ 进行预编码操作,将其转变为定义如下的二进制序列 $\{d_k\}$:

$$d_k = b_k \oplus d_{k-1} \quad (4.72)$$

其中,符号 \oplus 表示二进制数字 b_k 和 d_{k-1} 的模 2 加。该加法等价于一个两输入的异或操作,其执行如下:

$$d_k = \begin{cases} \text{符号 1} & \text{当 } b_k \text{ 和 } d_{k-1} \text{ 中有且仅有一个是 1 时} \\ \text{符号 0} & \text{其他} \end{cases} \quad (4.73)$$

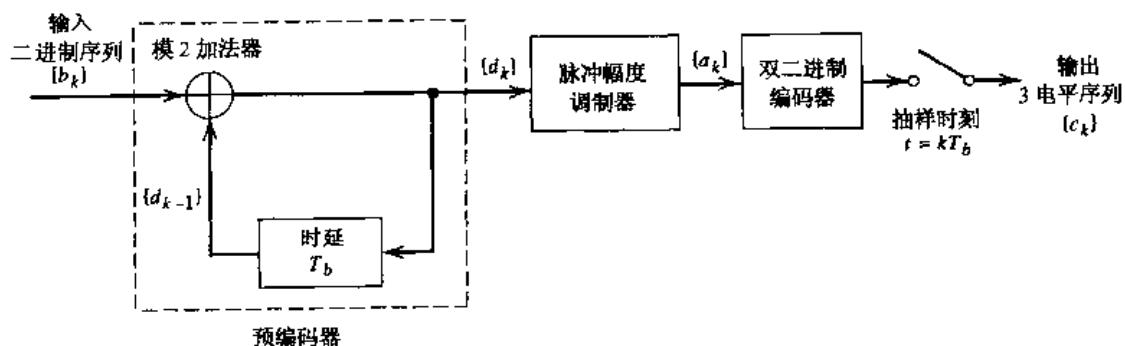


图 4.14 预编码双二进制方案;双二进制编码器的详细情况见图 4.11

使预编码二进制序列 $\{d_k\}$ 通过一个脉冲幅度调制器, 得到相应的2电平窄脉冲序列 $\{a_k\}$, 其中 $a_k = \pm 1$ 。接下来使该窄脉冲序列 $\{a_k\}$ 通过双二进制编码器, 产生序列 $\{c_k\}$:

$$c_k = a_k + a_{k-1} \quad (4.74)$$

值得注意的是, 和双二进制编码的线性操作不同, 式(4.72)描述的预编码是一个非线性操作。

结合式(4.72)和式(4.74)可得

$$c_k = \begin{cases} 0 & \text{当 } b_k = 1 \\ \pm 2 & \text{当 } b_k = 0 \end{cases} \quad (4.75)$$

见例 4.3。由式(4.75), 可推导出用于从序列 $\{c_k\}$ 中检测原始二进制序列 $\{b_k\}$ 的判决准则如下:

$$\begin{aligned} \text{如果 } |c_k| < 1, & \text{ 则 } b_k = 1 \\ \text{如果 } |c_k| > 1, & \text{ 则 } b_k = 0 \end{aligned} \quad (4.76)$$

当 $|c_k| \approx 1$ 时, 接收机就简单地在符号 1 和 0 之间进行一个随机猜测。按照该判决准则, 检测器由一个整流器组成, 检测器的输出在判决设备中与门限值 1 进行比较。该检测器的框图如图 4.15。该检测器的一个有用特征是除了当前样值之外, 不需要知道任何其他输入样值。因此, 错误传播现象是不可能在图 4.15 所示的检测器中发生的。

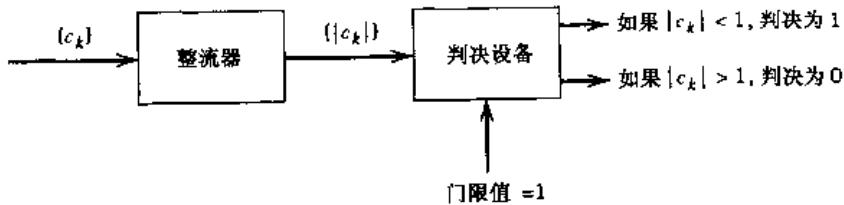


图 4.15 用于从预编码双二进制编码器输出中恢复原始二进制序列的检测器

例 4.3 使用预编码的双二进制编码

设一个二进制数据序列 0010110。为了对这个序列进行预编码(包括将预编码器的输出反馈到输入端), 在预编码的输出中增加一个额外比特。该额外比特取值任意, 如取 1。由式(4.73)不难发现, 预编码器的输出序列 $\{d_k\}$ 如表 4.1 第 2 行所示。预编码序列 $\{d_k\}$ 的极性表示如表 4.1 第 3 行所示。最后, 由式(4.74)可得, 双二进制编码器输出的幅度电平如表 4.1 第 4 行所示。

为了检测原始二进制序列, 应用式(4.76)中的判决准则, 得到如表 4.1 第 5 行所示的二进制序列。这个结果表明, 当不存在噪声时, 原始二进制序列得到了正确检测。

表 4.1 例 4.3 中关于双二进制编码的举例

二进制序列 $\{b_k\}$	0	0	1	0	1	1	0
预编码序列 $\{d_k\}$	1	1	1	0	0	1	0
2 电平序列 $\{a_k\}$	+1	+1	+1	-1	-1	+1	-1
双二进制编码输出 $\{c_k\}$	+2	+2	0	-2	0	0	-2
应用式(4.76)中的判决准则得到的二进制序列	0	0	1	0	1	1	0

4.6.2 改进的双二进制信号

双二进制信号的频率响应 $H(f)$, 以及发送脉冲的功率谱密度在原点处均非零。但在某些应用中, 双二进制信号的这个特性反而会成为缺陷, 因为很多通信信道不能传输直流分量。我们可以通过使用第 IV 类部分响应或改进的双二进制技术来弥补这个缺陷。该技术涉及到两个二进制数的相关距离。这种特殊的相关形式通过减去相距 $2T_b$ 秒的幅度调制脉冲来实现, 如图 4.16 的框图所示。预编码器包括一个 $2T_b$ 秒的时延。改进的双二进制变换滤波器的输出与脉冲幅度调制器输出端的 2 电平输入序列 $\{a_k\}$ 有如下关系:

$$c_k = a_k - a_{k-2} \quad (4.77)$$

这里也产生了一个 3 电平信号。当 $a_k = \pm 1$ 时, c_k 取值为 +2, 0 或 -2。

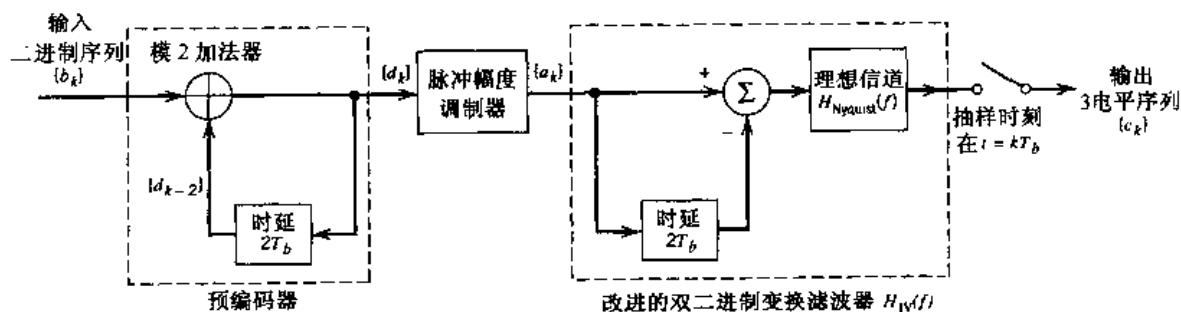


图 4.16 改进的双二进制信号方案

这种延迟滤波器和一个理想奈奎斯特信道级联(如图 4.16 所示)后的整体频率响应为

$$\begin{aligned} H_{IV}(f) &= H_{Nyquist}(f)[1 - \exp(-j4\pi f T_b)] \\ &= 2jH_{Nyquist}(f)\sin(2\pi f T_b) \exp(-j2\pi f T_b) \end{aligned} \quad (4.78)$$

其中, $H_{IV}(f)$ 的下标 IV 表示部分响应的相关类型, $H_{Nyquist}(f)$ 由式(4.68)定义。因此得到半周期正弦函数形式的整体频率响应, 即

$$H_{IV}(f) = \begin{cases} 2j \sin(2\pi f T_b) \exp(-j2\pi f T_b), & |f| \leq 1/2T_b \\ 0, & \text{其他} \end{cases} \quad (4.79)$$

相应的改进双二进制编码器的幅度响应与相位响应分别如图 4.17(a) 和图 4.17(b) 所示。改进的双二进制编码器的一个有用特性是其输出中没有直流分量。值得注意的是, 这种相关电平编码形式在频带边缘上具有与双二进制信号相同的连续性。

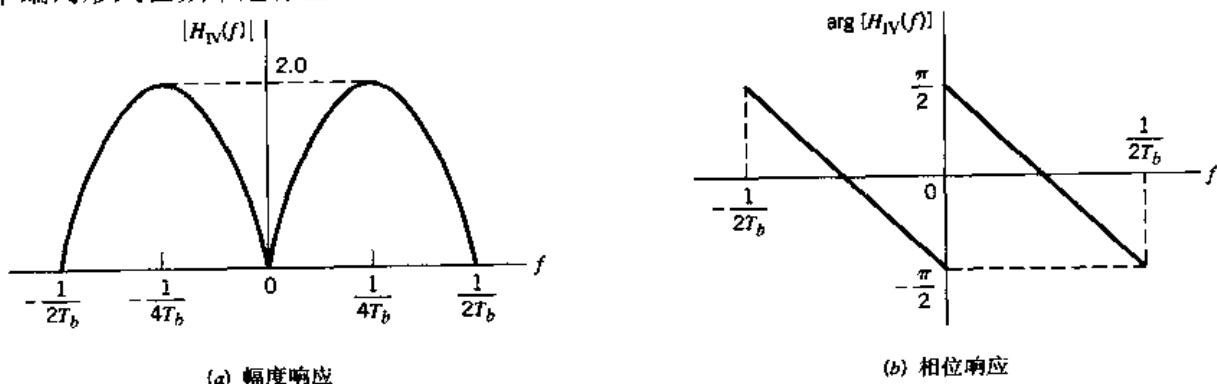


图 4.17 改进的双二进制变换编码器的频率响应

由式(4.78)的第一行和式(4.68)中 $H_{\text{Nyquist}}(f)$ 的定义, 不难发现改进的双二进制编码器的冲激响应由两个相互间隔 $2T_b$ 秒的 sinc(奈奎斯特)脉冲组成, 如下所示(这里去掉了比例因子):

$$\begin{aligned} h_{\text{IV}}(t) &= \frac{\sin(\pi t/T_b)}{\pi t/T_b} - \frac{\sin[\pi(t-2T_b)/T_b]}{\pi(t-2T_b)/T_b} \\ &= \frac{\sin(\pi t/T_b)}{\pi t/T_b} - \frac{\sin(\pi t/T_b)}{\pi(t-2T_b)/T_b} \\ &= \frac{2T_b^2 \sin(\pi t/T_b)}{\pi t(2T_b-t)} \end{aligned} \quad (4.80)$$

冲激响应如图 4.18 所示, 该图表明: 在抽样瞬间, 冲激响应只有三个可分辨的电平。值得注意的是, 与双二进制信号一样, 改进的双二进制信号的 $h_{\text{IV}}(t)$ 的尾部也以速率 $1/|t|^2$ 衰减。

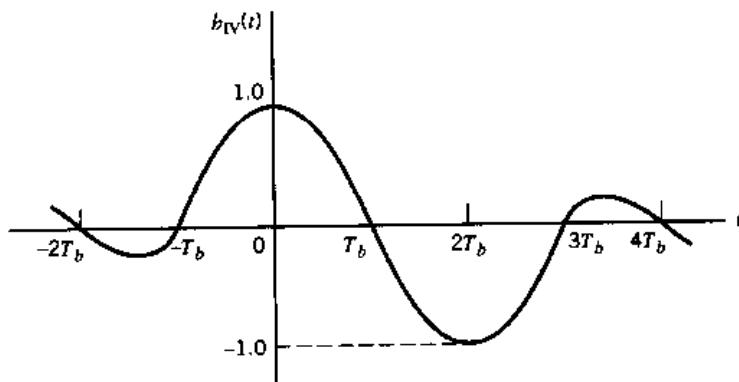


图 4.18 改进的双二进制变换滤波器的冲激响应

为了消除改进的双二进制系统中错误传播的可能性, 我们采用了一个与双二进制情况类似的预编码过程。具体做法是, 在产生改进的双二进制信号之前, 对相隔 $2T_b$ 秒的信号进行模 2 加, 即(见图 4.16 的前端):

$$\begin{aligned} d_k &= b_k \oplus d_{k-2} \\ &= \begin{cases} \text{符号 1} & \text{当 } b_k \text{ 与 } d_{k-2} \text{ 中有且仅有一个为 1 时} \\ \text{符号 0} & \text{其他} \end{cases} \end{aligned} \quad (4.81)$$

其中, $\{b_k\}$ 为输入的二进制数据序列, $\{d_k\}$ 为预编码器的输出序列。将由此生成的预编码序列 $\{d_k\}$ 送入脉冲幅度调制器, 然后送入改进的双二进制变换滤波器中。

在图 4.16 中, 假设脉冲幅度调制器对预编码序列 $\{d_k\}$ 使用极性表示, 则输出数字 c_k 等于 $-2, 0$ 或 $+2$ 。同时, 通过忽略 c_k 的极性, 可以在接收机输出端从 c_k 中提取检测数字 \hat{b}_k 。由此可得如下判决准则:

$$\begin{aligned} \text{如果 } |c_k| > 1, &\quad \text{则 } \hat{b}_k = 1 \\ \text{如果 } |c_k| < 1, &\quad \text{则 } \hat{b}_k = 0 \end{aligned} \quad (4.82)$$

当 $|c_k| = 1$ 时, 接收机在符号 1 和 0 之间进行一个随机猜测。与双二进制信号一样, 我们可以得到如下结论:

- 当不存在信道噪声时, 检测所得的二进制序列 $\{\hat{b}_k\}$ 恰好与发射机输入端的原始二进制

序列 $\{b_k\}$ 相同。

- 式(4.81)要求在预编码序列 $\{a_k\}$ 中增加两个额外比特。利用式(4.82)得到的译码序列 $\{\hat{b}_k\}$ 的成分不随这两个比特的选取而变化。

4.6.3 相关电平编码的一般形式(部分响应信号)

双二进制和改进的双二进制技术分别具有一个二进制数字和两个二进制数字的相关距离。由这两种技术,还可推广出统称为相关电平编码或部分响应信号的其他方法,如图 4.19 所示,其中 $H_{Nyquist}(f)$ 由式(4.68)定义。图中涉及到使用一个具有抽头权重 w_0, w_1, \dots, w_{N-1} 的抽头延迟线滤波器。具体地说,不同类型的部分响应信号方案都可以通过采用 N 个 sinc(理想奈奎斯特)脉冲的加权线性组合来实现,即

$$b(t) = \sum_{n=0}^{N-1} w_n \operatorname{sinc}\left(\frac{t}{T_b} - n\right) \quad (4.83)$$

适当选择式(4.83)中的抽头权重,就可以产生适合于不同应用的各种频谱形式。表 4.2 给出了五种不同类型的部分响应信号方案的具体细节。例如,在双二进制情况(第 I 类部分响应)中,有

$$w_0 = +1$$

$$w_1 = +1$$

且对于 $n \geq 2$ 有 $w_n = 0$ 。在改进的双二进制(第 IV 类部分响应)情况中,有

$$w_0 = +1$$

$$w_1 = 0$$

$$w_2 = -1$$

且对于 $n \geq 3$ 有 $w_n = 0$ 。

表 4.2 对应于图 4.19 的部分响应信号方案的不同分类

分 类	N	w_0	w_1	w_2	w_3	w_4	注 释
I	2	1	1				双二进制编码
II	3	1	2	1			
III	3	2	1	-1			
IV	3	1	0	-1			改进的双二进制编码
V	5	-1	0	2	0	-1	

部分响应信号方案的有用特性可以总结如下:

- 物理基带信道上的二进制数据传输,可以通过采用具有逐渐截止特性的滤波器来实现,其传输速率可以接近于奈奎斯特速率。
- 可以产生不同的频谱形式,以适应各种实际应用的需要。

利用这些特性所要付出的代价是,由于所使用的信号电平数的增加,需要一个更大的信噪比,因此当存在噪声时,系统将产生与相应的二进制 PAM 系统相同的平均符号差错概率。

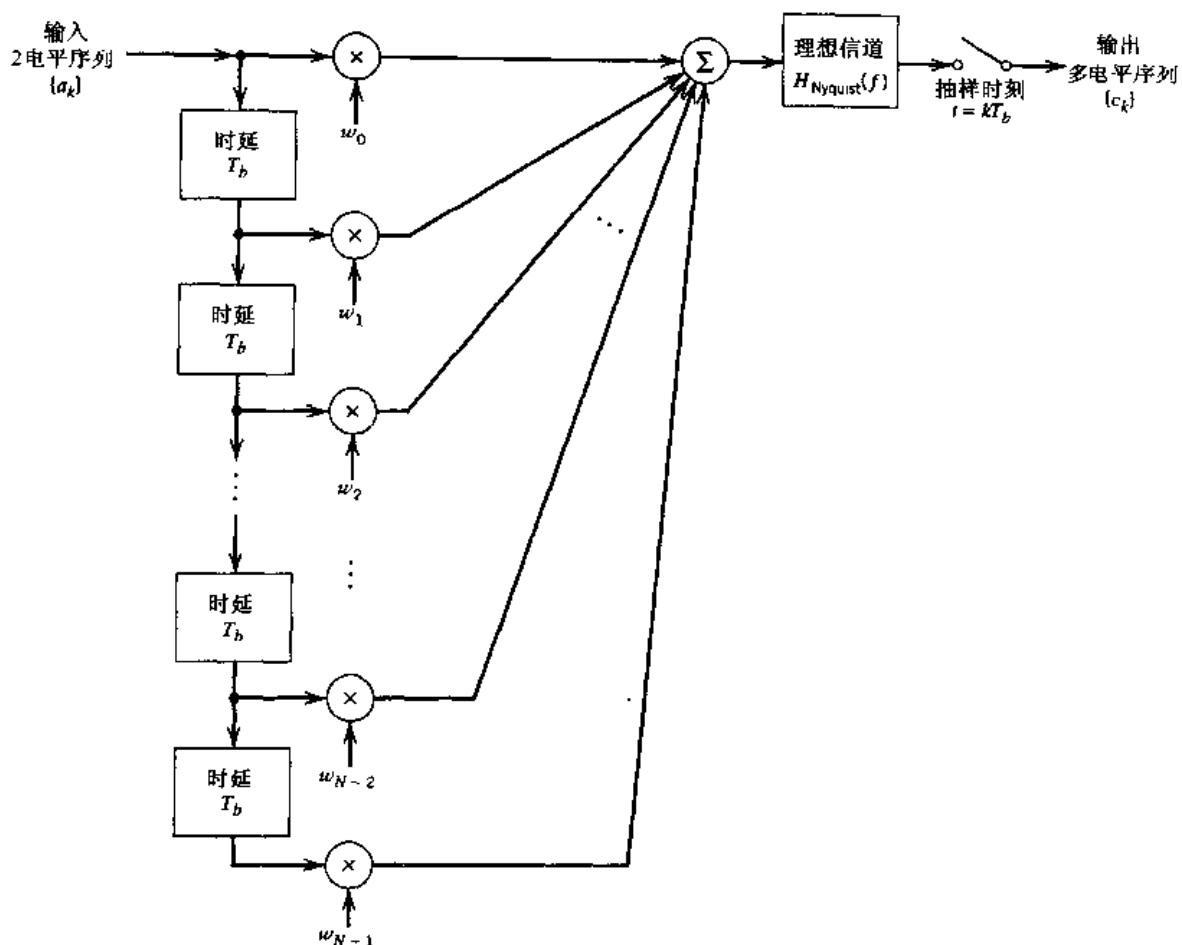


图 4.19 通用的相关编码方案

4.7 基带 M 进制 PAM 传输

在图 4.7 所示的基带二进制 PAM 系统中，脉冲幅度调制器产生二进制脉冲，即脉冲的电平为两个可能的幅度电平之一。而在一个基带 M 进制 PAM 系统中，脉冲幅度调制器产生的电平为 M 个可能的幅度电平之一，其中 $M > 2$ 。图 4.20(a) 为四进制 ($M = 4$) 系统中，二进制数据序列 0010110111 脉冲调制的形式。图 4.20(a) 所示的波形是基于图 4.20(b) 给出的 4 个可能二位组 (一对比特) 中每一个的电平表示法而绘制的。值得注意的是，这种表示方法是基于格雷编码的，这意味着四进制字符集中，任何一个二位组与相邻的二位组之间只有一个比特不同。

在一个 M 进制系统中，信源从一个由 M 个符号组成的字符集中选择一个符号序列进行发送。脉冲幅度调制器输出端的每个幅度电平对应一个不同的符号，因此，在发射端有 M 个不同的幅度电平被发送。接下来考虑一个 M 进制的 PAM 系统。该系统采用一个包含有 M 个等概率且统计独立的符号的字符，符号持续时间记为 T 秒。我们称 $1/T$ 为系统的信号速率，其单位为符号每秒或波特 (bauds)。将该系统的信号速率与一个等价的二进制 PAM 系统的信号速率联系在一起进行分析，是很有意义的。此时 M 的值为 2，且相继的二进制符号 1 和 0 是等概率且统计独立的，每个符号的持续时间记为 T_b 秒。在上述条件下，二进制 PAM 系统以 $1/T_b$ bps 的速率产生信息。在四进制 PAM 系统中，四种可能的符号可以由二位组 00, 01, 10 和

11进行标识。因此,每个符号代表2个比特的信息,且1波特等于2 bps。此结果可概括为,在一个 M 进制PAM系统中,1波特等于 $\log_2 M$ bps,且 M 进制PAM系统的符号持续时间 T 与等价的二进制PAM系统的比特持续时间 T_b 有如下关系:

$$T = T_b \log_2 M \quad (4.84)$$

因此,在一个给定的信道带宽下,采用 M 进制的PAM系统,能够以比相应的二进制PAM系统快 $\log_2 M$ 倍的速率传输信息。但在相同的平均符号差错概率下, M 进制PAM系统需要更大的发射功率。具体地说,当 M 远大于2且平均符号差错概率小于1时, M 进制PAM系统的发射功率等于在二进制PAM系统的发射功率上乘以因子 $M^2/\log_2 M$ 。

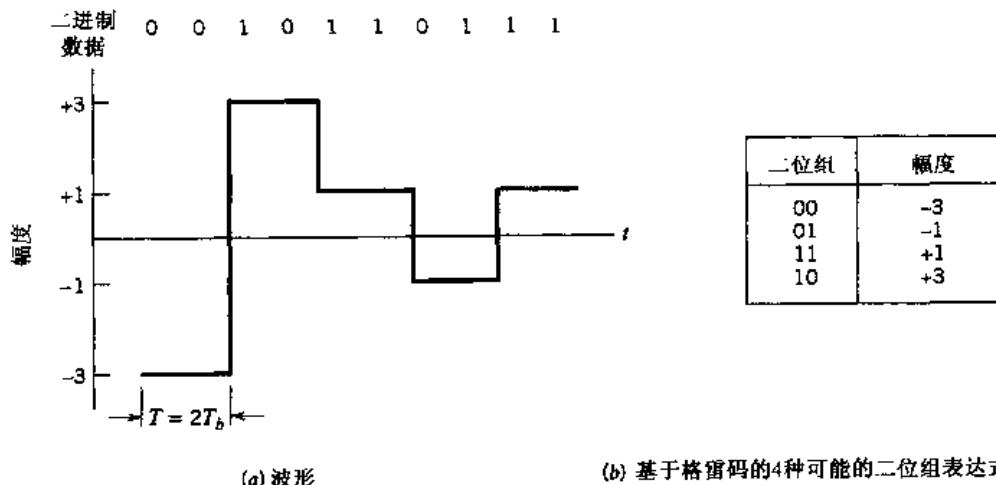


图4.20 四进制系统的输出

在一个基带 M 进制系统中,信源发出的符号序列首先由一个处于发射机输入端的脉冲幅度调制器转变为一个 M 电平PAM脉冲串。接着,和二进制PAM系统一样,该脉冲串由发射滤波器进行整形,并通过通信信道传输,信道中的噪声和失真将破坏信号波形。接收信号通过接收滤波器,然后以与发射机同步的速率进行抽样。将每个抽样值与预设的门限值(也称为切割电平)进行比较,以判决发送的是哪个符号。因此, M 进制PAM脉冲幅度调制器及其判决设备的设计,远比二进制PAM系统更为复杂。ISI、噪声和不完善的同步系统都会引发接收机输出端的错误。发射和接收滤波器就是为使这些差错最小化而设计的。这些滤波器的设计过程与4.5节和4.6节中讨论的基带二进制PAM系统非常相似。

4.8 数字用户线

对基带数据传输的研究进行到这里,我们有必要向读者介绍一种快速成长的应用:数字用户线(DSL)。^[6]DSL工作于一个提供了用户终端(例如,计算机)和电话公司的中心局(CO)间直接连接的本地环路(小于1.5 km)。如图4.21所示。通过CO,DSL用户被连接到一个基于ATM和IP技术的宽带骨干数据网。这些技术和相关的网络资源(即光纤或同步光网络)在“背景与预览”一节中曾经进行过讨论。因此,承载信息的信号以数字方式从用户终端传输到因特网服务提供商。在其通过数据网传输期间,信号以一定的规则进行交换或路由选择。

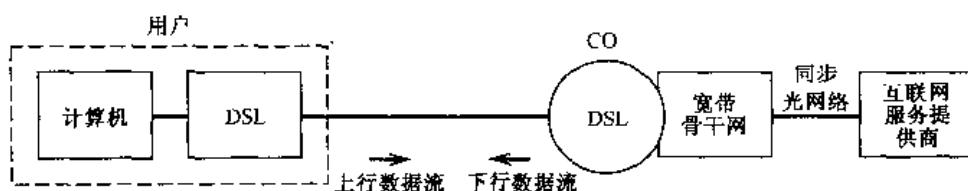
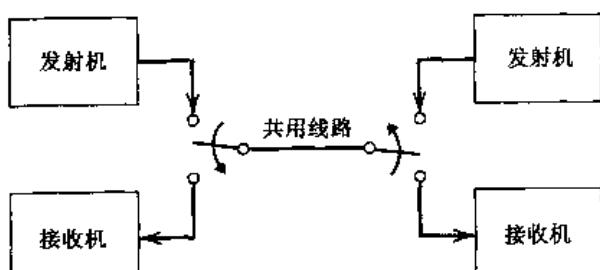


图 4.21 描述 DSL 工作环境的框图

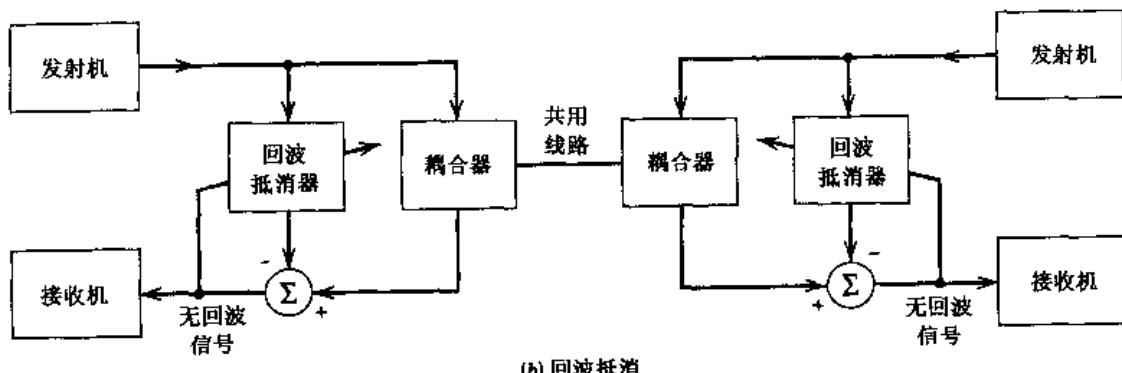
在廉价的应用中,DSL 使用的双绞线能够提供高数据速率,全双工的数字传输能力。(正如在“背景与预览”一章中所提到的那样,双绞线也在普通电话通信中使用。)为了实现全双工的两线传输,可采用下面两种操作模式中的一种:

1. **时间压缩(TC)复用。**其中,在同一线路上的两个不同方向上的数据传输在时间上是分开的。具体地说,数据块是以交替传输的原则在每个方向上进行传输的,如图 4.22(a)所示。

各个数据脉冲之间将插入保护时间。因此,线路速率比数据速率的两倍稍高。



(a) 时间压缩复用



(b) 回波抵消

图 4.22 全双工操作

2. **回波抵消模式。**这种模式支持同一条线路两个方向上数据的同时流动。为了实现这种传输模式,每个收发信机(发射机/接收机)包含了一个耦合器。耦合器的作用是将发射信号与接收信号相分离以及进行 2-4 线转换,如图 4.22(b)所示。耦合器,或更准确地说是耦合变换器,主要由一个有三端口(终端对)的电桥构成,如图 4.23 所示。如果该电桥没有很好的平衡,耦合器的发射端就会与接收端相耦合,由于近端(本地)发射信号与近端(本地)接收机之间产生的泄露,导致了回波的产生。为了消除这种无用的回波,每个收发信机必须包含一个回波抵消器,如图 4.22(b)所示。在这种操作模式中,由于数据可以同时在两个方向上通过线路流动,因此线路速率等于数据速率。

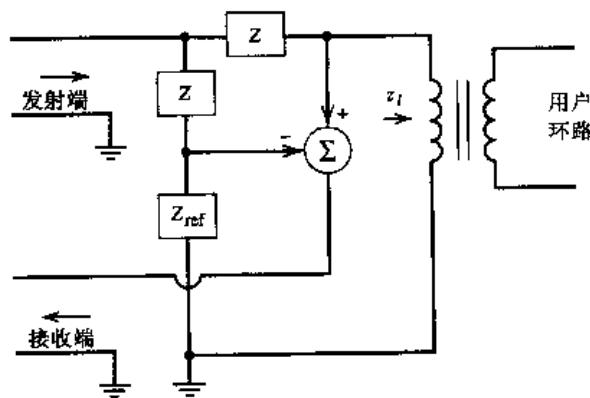


图 4.23 耦合变换器的简化电路。为了保持电桥平衡,参考阻抗 Z_{ref} 应等于线路阻抗 z_t

综上所述,回波抵消模式提供了比时间压缩复用模式好得多的数据传输性能,但这是以增加系统复杂度为代价得到的。将整个收发器集成在单个 VLSI 芯片中,尽管系统复杂度仍有所增加,但费用还是可接受的。在北美,回波抵消方法已被用来进行收发器的设计。关于自适应回波抵消器的讨论见习题 4.31。

除回波之外,还有其他需要考虑的传输媒质损耗。其中两个主要的损耗就是 ISI 和串话(crosstalk),下面依次进行讨论。

对于一阶近似,双绞线的平方幅度响应为

$$|H(f)|^2 = \exp(-\alpha\sqrt{f}) \quad (4.85)$$

其中

$$\alpha = k \frac{l}{l_0} \quad (4.86)$$

在式(4.85)中,频率 f 的单位是 kHz, k 为双绞线的物理常数, l_0 为参考长度(例如 km), l 为双绞线的实际长度。式(4.85)指出在基带数据传输中使用双绞线所带来的一个主要损耗是频率响应的逐渐下降,这将导致 ISI 的增大。

串话出现的主要原因是在一条电缆的相邻双绞线中存在的电容耦合。典型地,绝大部分串话产生在电缆中最靠近的 5 到 7 个双绞线之间。在任何情况下,以下两种串话都可以在接收机中观察到:

1. 近端串话(NEXT),由与接收机处于电缆同一端的发射机产生,如图 4.24(a)所示。
2. 远端串话(FEXT),由与接收机距离较远的发射机产生,如图 4.24(b)所示。

自然地,远端串话与信号一样要遭受线路损失,而近端串话就不会。因此,在如图 4.22(b)所示的回波抵消方法里,信号在电缆中沿两个方向传输,近端串话要比远端串话强得多。因此,在本书以后的内容里,将忽略远端串话的影响。

实际上,近端串话和 ISI 是决定 DSL 环路性能的两个最重要的因素。图 4.25 为受这两种损耗影响的双绞线信道模型。由于双绞线通常都传输相似的信号,通过有如下近似式的串话频率响应 $H_{NEXT}(f)$,可将近端串话建模为一个与发射信号具有相同功率谱密度的信号。即 $H_{NEXT}(f)$ 近似为

$$H_{NEXT}(f) = \beta f^{3/2} \quad (4.87)$$

其中, β 为电缆常量。在图 4.25 中需要注意的重点是, 发射信号与干扰信号具有相同的功率谱密度, 二者仅在各自的频率响应上有所不同, 见式(4.85)和式(4.87)。对该模型进行仿真研究时, 发射信号由一个随机数据序列表示, 而干扰则由一个高斯噪声序列表示。

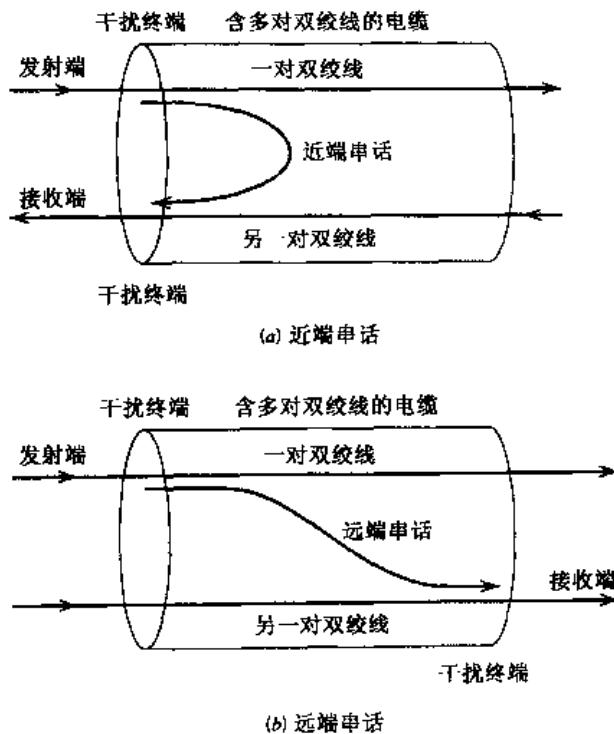


图 4.24 两种串话现象

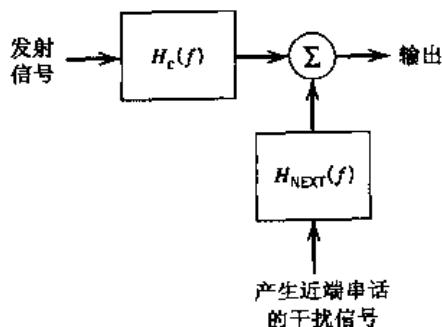


图 4.25 双绞线信道模型

4.8.1 用于 DSL 的线路码

现在, 我们已经认识了主要的传输损耗, 下面描述发射信号频谱所应当具备的特性:

1. 发射信号的功率谱密度在零频率处应为零, 其原因是耦合变换器不能通过直流分量。
2. 发射信号的功率谱密度在高频处应较低, 其原因如下:
 - 双绞线中的传输衰减在高频处最为严重。
 - 由于高频处电容耦合性的增强, 相邻双绞线的串话将急剧增加。这让我们联想到电容的阻抗与频率值成反比的性质。

为了获得上述特性,我们选择线路码,将输入数据比特流映射为电脉冲,以便在线路上进行传输。每种线路码都有各自的优缺点。下面列出了一些可用的线路码:

- **曼彻斯特码。**这种线路码简单且无直流分量。其缺点是占用了很大的频带,因此对近端串话和 ISI 非常敏感。(有关曼彻斯特码的讨论见 3.7 节。)
- **改进的双二进制码。**这种线路码无直流分量,频谱效率适中,且具有最小的 ISI。但是,对改进的双二进制码串话性能的仿真研究表明,在情况最差的用户线路上,这种码对近端串话和 ISI 的抵抗能力要比分组码低 2~3 dB。(有关改进的双二进制码的讨论见 4.6 节。)
- **双极性码。**这种线路码将连续的 1 用等值的正负电平交替表示,将符号 0 用零电平值表示。双极性码信号无直流分量。计算机仿真表明在所有数字用户环路上,这种码的近端串话和 ISI 性能只稍低于改进的双二进制码。有关双极性码即交替信号反转(AMI)码的讨论见 3.7 节。
- **2B1Q 码。**2B1Q 代表被编码为一个四进制符号的两个二进制数。这种码是一种代表 4 电平 PAM 信号的分组码,如图 4.20 所示。假设符号 1 和 0 等概,2B1Q 码平均来说无直流分量。这里讨论的所有线路码中,2B1Q 码波特率的减少是最多的,其抵抗近端串话和 ISI 的性能也是最好的。

与曼彻斯特码、改进的双二进制码、双极性码以及其他这里没有提到的线路码^[7]相比,2B1Q 码具有更好的通信能力,因此 2B1Q 码被采用作为 DSL 的北美标准。

采用 2B1Q 码作为线路码,并利用 VLSI 技术集成了自适应均衡器和回波抵消器的收发器,在绝大多数双绞线用户线路上,能够以 160 Kbps 的速率在全双工工作模式下实现 10^{-7} 的误比特率。当出现最差情况即近端串话率达到 1% 时,2B1Q 码仍具有 12 dB 的噪声容限和 10^{-7} 的误比特率,这对于 DSL 来说,是一个可以接受的性能标准。噪声容限即误码率为不超过 10^{-7} 时可以容忍的接收机噪声值(包括未抵消的回波)。

4.8.2 非对称数字用户线

DSL 的另一个重要类型是非对称数字用户线(ADSL)。ADSL 是用来在单个双绞线上同时支持三项服务的本地传输系统:

1. 可达 9 Mbps 比特率的数据传输下行数据流(传向用户)。
2. 可达 1 Mbps 比特率的数据传输上行数据流(离开用户)。
3. 传统电话服务(POTS)。

下行数据流和上行数据流的比特率取决于用于传输的双绞线长度。ADSL 被称为“非对称”是因为下行数据流比特率要远大于上行数据流比特率。模拟语音在基带频率上进行传输,并使用频分复用与下行数据流和上行数据流的通带传输相结合。如图 4.26(a)所示,上行数据流的传输在一个不同于下行数据流传输的频带上进行以避免串话。同时,一个保护频带被插入到 POTS 频带与上行数据流传输频带之间。通过使用一对分离器,使 ADSL 与 POTS 信号在本地环路的共存成为可能。其中一个分离器位于本地环路的 CO 终端,另一个则位于用户终端。从功能上说,一个分离器可分为两个双向滤波器,如图 4.26(b)所示:

- 一个用于基带传输或提取语音信号的低通滤波器。
- 一个用于通常传输或提取 ADSL 数据的高通滤波器。

分离器有效地起到了频分复接器或分接器的作用,具体是复接器还是分接器,取决于信号传输的方向。

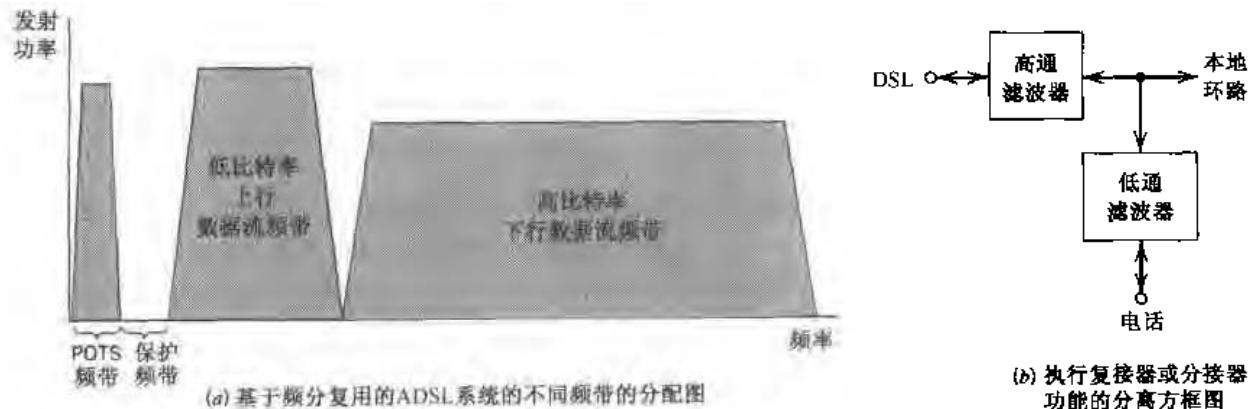


图 4.26 ADSL 系统(注意,分离器的两个滤波器都是双向滤波器)

让 DSL 非对称的目的是为了提供“视频点播”。在这样的应用中,一个用户需要一个高吞吐量的信道从 CO 按需下载高带宽的视频数据。在相反方向上,一个吞吐量低得多的信道就足以用来发送定制信息及实时的控制命令。例如,一个面向因特网提供 1.544 Mbps 的 DS1 速率的下行数据流传输和大约 160 Kbps 的上行数据流传输的 ADSL,就能够满足这样的应用要求。这样的一个系统实现了约为 10:1 的非对称比,避免了 IP 协议中确认分组(包)流量的问题,从而获得了更高的数据传输速率。

要以 DS1 或更高的速率在双绞线上传输数据是很困难的,为此必须采用复杂的调制技术。对该问题的讨论参见第 6 章。

4.9 最佳线性接收机

下面继续对图 4.7 所示的基带数据传输系统的研究。我们已分别讨论了以下两种信道条件:

- 信道噪声单独作用的信道,并引出了匹配滤波器的公式。
- ISI 单独作用的信道,并引出了用于实现奈奎斯特信道的脉冲整形发射滤波器的公式。

然而,在实际情况中,信道噪声和 ISI 往往是同时作用的。在本节中,我们将阐述如何在既具有弥散特性又受到噪声影响的线性信道通用情况下进行线性接收机的优化设计。

在设计线性接收机时,往往将接收机看成是后面带有判决设备的迫零均衡器(zero-forcing equalizer)。均衡器采用这种形式,其目的就是为了使得在除 $k = 0$ 之外的所有抽样时刻 $t = kT$,信道输出的 ISI“被迫为零”,其中假设 $k = 0$ 时产生的是有用符号。在这种条件下,如果信道噪声 $w(t)$ 为零,那么,按照奈奎斯特准则,逐符号的检测可确保是最优的。

迫零均衡器相对来说易于实现,因为它忽略了信道噪声 $w(t)$ 的影响。这种忽略带来的严重后果是噪声加强时系统整体性能的下降。这是迫零均衡器的一个固有现象(见习题 4.32)。

一个更加巧妙的接收机设计方法就是使用均方误差准则,这种方法在减小信道噪声和减小 ISI 之间进行了折中处理。实际上,对于给定的计算复杂度来说,按照均方误差准则设计的均衡器总是具有比相应迫零均衡器更好或与之相同的性能。在本书后面的内容中,我们就着重考虑均方误差准则。

图 4.7 所示的基带二进制数据传输系统中,冲激响应为 $c(t)$ 的接收滤波器对信道输出 $x(t)$ 产生如下响应:

$$y(t) = \int_{-\infty}^{\infty} c(\tau)x(t - \tau) d\tau \quad (4.88)$$

信道输出 $x(t)$ 定义为

$$x(t) = \sum_k a_k q(t - kT_b) + w(t) \quad (4.89)$$

其中, a_k 为时刻 $t = kT_b$ 的发送符号, $w(t)$ 为信道噪声。时间函数 $q(t)$ 为两个冲激响应的卷积:与脉冲整形发射滤波器相关的 $g(t)$ 和与信道相关的 $b(t)$ 。将式(4.89)代入式(4.88)并在时刻 $t = iT_b$ 对输出 $y(t)$ 进行抽样,得

$$y(iT_b) = \xi_i + n_i \quad (4.90)$$

其中, ξ_i 为定义如下的信号分量:

$$\xi_i = \sum_k a_k \int_{-\infty}^{\infty} c(\tau)q(iT_b - kT_b - \tau) d\tau \quad (4.91)$$

n_i 为定义如下的噪声分量:

$$n_i = \int_{-\infty}^{\infty} c(\tau)w(iT_b - \tau) d\tau \quad (4.92)$$

接收机的理想工作条件就是 $y(iT_b) = a_i$, 其中 a_i 为发送符号。偏离该条件会产生误差信号

$$\begin{aligned} e_i &= y(iT_b) - a_i \\ &= \xi_i + n_i - a_i \end{aligned} \quad (4.93)$$

因此,可以正式定义均方误差为

$$J = \frac{1}{2} E[e_i^2] \quad (4.94)$$

其中, E 为期望运算符,引入因子 $1/2$ 是为了表达的方便。将式(4.93)代入式(4.94)并展开各项,有

$$J = \frac{1}{2} E[\xi_i^2] + \frac{1}{2} E[n_i^2] + \frac{1}{2} E[a_i^2] + E[\xi_i n_i] - E[n_i a_i] - E[\xi_i a_i] \quad (4.95)$$

现在我们按照这 6 项期望值在式中出现的顺序分别进行求值:

1. 在稳态环境中,均方项 $E[\xi_i^2]$ 独立于接收滤波器输出的抽样时刻 $t = iT_b$ 。因此,可将该项简化如下:

$$E[\xi_i^2] = \sum_l \sum_k E[a_l a_k] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\tau_1)c(\tau_2)q(lT_b - \tau_1)q(kT_b - \tau_2) d\tau_1 d\tau_2$$

假设二进制符号 $a_k = \pm 1$ 如式(4.42)所示,且发送符号统计独立,即

$$E[a_i a_k] = \begin{cases} 1 & l = k \\ 0 & \text{其他} \end{cases} \quad (4.96)$$

可将均方项 $E[\xi_i^2]$ 进一步简化为

$$E[\xi_i^2] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_q(\tau_1, \tau_2) c(\tau_1) c(\tau_2) d\tau_1 d\tau_2 \quad (4.97)$$

其中

$$R_q(\tau_1, \tau_2) = \sum_k q(kT_b - \tau_1) q(kT_b - \tau_2) \quad (4.98)$$

这里因子 $R_q(\tau_1, \tau_2)$ 为序列 $\{q(kT_b)\}$ 的时间自相关函数。该序列的平稳性意味着(见 1.5 节):

$$R_q(\tau_1, \tau_2) = R_q(\tau_2 - \tau_1) = R_q(\tau_1 - \tau_2)$$

2. 利用式(4.92),得到由信道噪声引起的均方项 $E[n_i^2]$:

$$\begin{aligned} E[n_i^2] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\tau_1) c(\tau_2) E[w(iT_b - \tau_1) w(iT_b - \tau_2)] d\tau_1 d\tau_2 \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\tau_1) c(\tau_2) R_w(\tau_2 - \tau_1) d\tau_1 d\tau_2 \end{aligned} \quad (4.99)$$

其中, $R_w(\tau_2 - \tau_1)$ 为信道噪声, $w(t)$ 为整体平均自相关函数。假设 $w(t)$ 为白噪声且功率谱密度为 $N_0/2$,有

$$R_w(\tau_2 - \tau_1) = \frac{N_0}{2} \delta(\tau_2 - \tau_1) \quad (4.100)$$

因此, $E[n_i^2]$ 的表达式简化为

$$E[n_i^2] = \frac{N_0}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\tau_1) c(\tau_2) \delta(\tau_2 - \tau_1) d\tau_1 d\tau_2 \quad (4.101)$$

3. 根据式(4.96),由发送符号 a_i 引起的均方项 $E[a_i^2]$ 为 1,即

$$E[a_i^2] = 1 \quad \text{对于所有 } i \quad (4.102)$$

4. 交叉乘积项 $\zeta_i n_i$ 的期望值为零,其原因有两个:第一, ζ_i 与 n_i 相互独立;第二,信道噪声 $w(t)$ 与 n_i 都具有零均值。即

$$E[\zeta_i n_i] = 0 \quad \text{对于所有 } i \quad (4.103)$$

5. 由于相似的原因,交叉乘积项 $n_i a_i$ 的期望值也为零,即

$$E[n_i a_i] = 0 \quad \text{对于所有 } i \quad (4.104)$$

6. 最后,由式(4.91)可得交叉乘积项 $\zeta_i a_i$ 的期望值如下:

$$E[\zeta_i a_i] = \sum_k E[a_k a_i] \int_{-\infty}^{\infty} c(\tau) q(iT_b - kT_b - \tau) d\tau \quad (4.105)$$

利用式(4.96)描述的发送符号的统计独立性,可将该期望简化为

$$E[\xi; a_i] = \int_{-\infty}^{\infty} c(\tau) q(-\tau) d\tau \quad (4.106)$$

将式(4.97), 式(4.101)~式(4.104)和式(4.106)代入式(4.95), 可将图 4.7 所示的二进制数据传输系统的均方误差 J 表示为

$$J = \frac{1}{2} + \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(R_q(t - \tau) + \frac{N_0}{2} \delta(t - \tau) \right) c(t) c(\tau) dt d\tau - \int_{-\infty}^{\infty} c(t) q(-t) dt \quad (4.107)$$

为了表达的方便, 式中对变量进行了如下替换: 第一个积分中的 τ_1 和 τ_2 分别由 t 和 τ 代替, 第二个积分中的 τ 由 t 代替。同时, 根据式(4.96)的假设, 均方误差 J 的上述表达式实际上已经对发送符号 a_k 的方差进行了归一化。

利用式(4.107)中关于均方误差 J 的公式, 可对图 4.7 所示的接收滤波器进行具体设计。将式(4.107)对接收滤波器的冲激响应 $c(t)$ 求导, 并令结果为零, 得

$$\int_{-\infty}^{\infty} \left(R_q(t - \tau) + \frac{N_0}{2} \delta(t - \tau) \right) c(\tau) d\tau = q(-t) \quad (4.108)$$

式(4.108)是(用均方误差方法得到的)最佳均衡器的冲激响应 $c(t)$ 的公式。按此设计的均衡器称为最小均方误差(MMSE)均衡器。

对式(4.108)中两边取傅里叶变换, 得到

$$\left(S_q(f) + \frac{N_0}{2} \right) C(f) = Q^*(f) \quad (4.109)$$

其中, $c(t) \Leftrightarrow C(f)$, $q(t) \Leftrightarrow Q(f)$, 以及 $R_q \Leftrightarrow S_q(f)$ 。由式(4.109)求解 $C(f)$, 得

$$C(f) = \frac{Q^*(f)}{S_q(f) + \frac{N_0}{2}} \quad (4.110)$$

由习题 4.33 可知, 序列 $\{q(kT_b)\}$ 的功率谱密度可表示为

$$S_q(f) = \frac{1}{T_b} \sum_k \left| Q\left(f + \frac{k}{T_b}\right) \right|^2 \quad (4.111)$$

这意味着最佳线性接收机的频率响应 $C(f)$ 是一个周期为 $1/T_b$ 的函数。式(4.110)将最佳线性接收机解释为两个基本分量^[8]的级联:

- 冲激响应为 $q(-t)$ 的匹配滤波器, 其中 $q(t) = g(t) \star b(t)$ 。
- 频率响应为周期函数 $S_q(f) + (N_0/2)$ 的倒数的横向(抽头延迟线)均衡器。

为了精确实现式(4.110), 需要一个无限长度的均衡器。实际上, 只要 N 足够大, 就可以通过采用一个系数为 $\{c_k\}_{k=-N}^N$ 的有限集合的均衡器来近似地作为最佳解。相应地, 接收机就有了图 4.27 所示的形式。注意图 4.27 中标有 z^{-1} 字样的部分为一个等于 T_b 的时延, 这意味着均衡器的抽头间隔恰好等于比特持续时间 T_b 。这样的均衡器称为是与发射机相同步的。

4.9.1 实际考虑

图 4.27 的最小均方误差接收机在实验室条件下工作得很好，并且达到了系统的均衡。可通过一组系数 $\{c_k\}_{k=-N}^N$ 来确定横向均衡器的特性。该均衡器与式(4.110)中的频率响应 $C(f)$ 已经相当接近。但在实际的通信环境中，信道通常是时变的，例如，在 PSTN 中，有两个因素将影响不同链路上脉冲失真的分布：

- 进行交换的各个链路的传输特性的不同。
- 一个连接中链路数的不同。

其结果就是：电话信道是随机的。因此，采用基于平均信道特性而设计的匹配滤波器和均衡器的固定搭配，无法充分减小 ISI 和信道噪声的影响。为了充分利用电话信道的传输能力，需要采用一种自适应的接收机^[9]。这种接收机能以组合的方式使匹配滤波器和均衡器自适应运行。由于均衡器系数可按照一个内置算法而自动调整，所以这种接收机是自适应的。

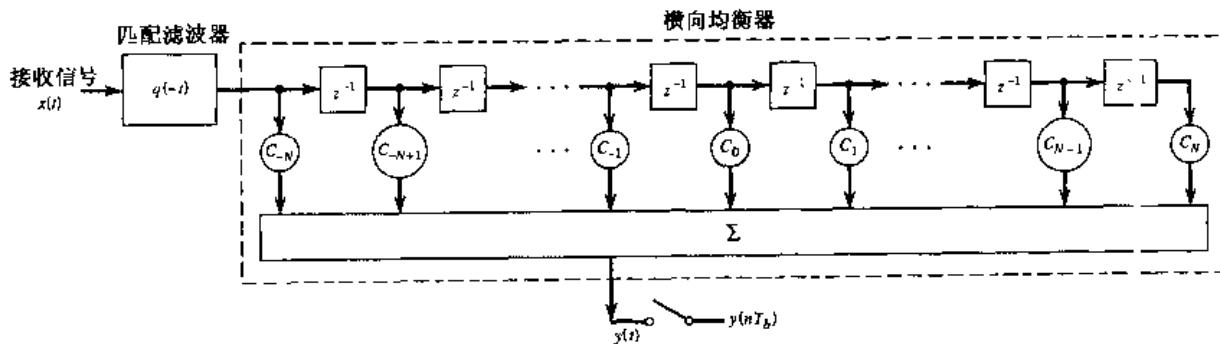


图 4.27 由匹配滤波器和横向均衡器级联而成的最优线性接收机

此外，我们还希望均衡器的抽头间隔小于码元周期。通常的做法是将相邻抽头之间的间隔设为 $T/2$ 。这样得到的最终结构称为小间隔均衡器(FSE)。与传统的同步均衡器相比，FSE 具有更有效的时延失真补偿能力。FSE 的另一个优点是数据传输可以在任意抽样状态下开始。但对 FSE 的数学分析要比同步均衡器的更为复杂，因此这里不再赘述^[10]。

4.10 自适应均衡

本节将提出一种简单而有效的未知特性线性信道的自适应均衡算法。图 4.28 为自适应同步均衡器的结构。该均衡器同时也具有匹配滤波器的功能。为什么需要自适应均衡呢？为了回答这个问题，首先要了解一点，那就是在平均数据呼叫期间，典型的电话信道几乎不会发生变化。因此，在数据传输之前，均衡器就在一个通过信道发送的训练序列的指导下进行了调整。该训练序列的一个同步版本在接收机中产生，并经过一个等于信道传输时延的时移之后，作为期望响应而送入均衡器。实际应用中普遍采用的训练序列是伪噪声(PN)序列。该序列由一个具有与噪声类似特性的确定周期序列组成。一般是采用两个同样的 PN 序列发生器，一个位于发射端，另一个位于接收端。当训练过程结束后，PN 序列发生器关闭，自适应均衡器就可以进行正常的数据传输了。对 PN 序列发生器的详细描述见第 7 章。

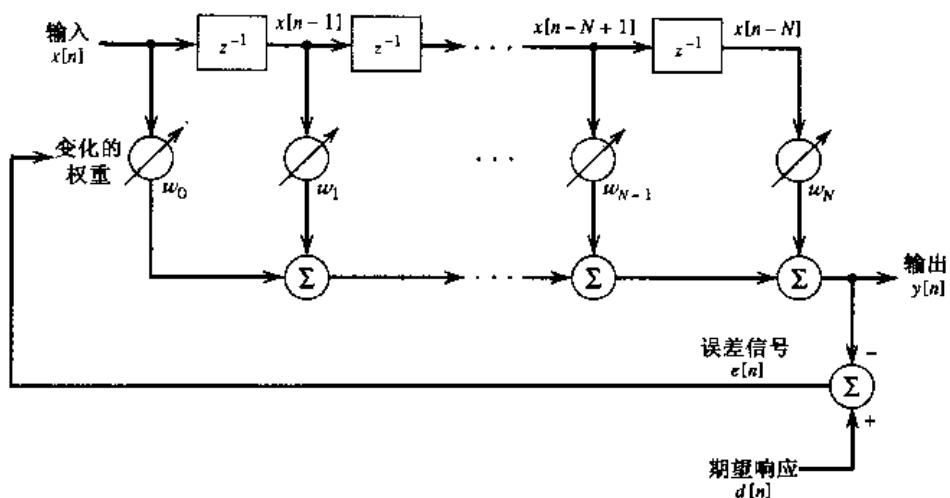


图 4.28 自适应均衡器的框图

4.10.1 LMS 算法(回顾)

为了简化符号, 设

$$\begin{aligned}x[n] &= x(nT) \\y[n] &= y(nT)\end{aligned}$$

则抽头延迟线均衡器对输入序列 $\{x[n]\}$ 的响应输出 $y[n]$ 就由离散卷积和(如图 4.28)定义为

$$y[n] = \sum_{k=0}^{N-1} w_k x[n-k] \quad (4.112)$$

其中, w_k 为第 k 个抽头的权重, $N+1$ 为抽头总数。抽头权重就是自适应滤波器的系数。这里假设输入序列 $\{x[n]\}$ 具有有限能量。图 4.28 中采用了与图 4.27 中不同的符号来表示均衡器的权重, 是为了强调图 4.28 所示的均衡器也具有匹配滤波器的功能。

通过在抽样时刻观察期望脉冲波形与滤波器输出端实际脉冲波形之间的误差, 然后用该误差来估计滤波器权重为接近最佳值而应进行的改变, 就可以实现自适应。对于自适应, 我们可以使用一个基于最小化峰值失真的准则。峰值失真的定义为均衡器输出端在最差情况下的 ISI。基于该准则的自适应均衡器是建立在 4.9 节所述的迫零概念之上的。只有在输入峰值失真小于 100% (即 ISI 不太严重) 时, 该均衡器才是最佳的。更好的方法是采用应用更为普遍的均方误差准则。与基于峰值失真准则的自适应均衡器相比, 基于均方误差准则的自适应均衡器对时间扰动不太敏感。接下来就采用均方误差准则来推导自适应均衡器的算法。

用 $a[n]$ 表示期望响应, 其定义为第 n 个发送二进制符号的极性表示。用 $e[n]$ 表示误差信号, 其定义为期望响应 $a[n]$ 与均衡器实际响应 $y[n]$ 之差, 即

$$e[n] = a[n] - y[n] \quad (4.113)$$

在用于自适应均衡器的 LMS 算法^[1] 中, 随着算法的迭代进行, 由于误差信号 $e[n]$ 的存在, 必须对均衡器的各个抽头权重不断进行调整。用于自适应预测的 LMS 算法的推导见 3.13 节。将式(3.72)改写为最一般的形式, 我们可以用文字将 LMS 算法表达如下:

$$\begin{pmatrix} \text{第 } k \text{ 个权重的更新值} \\ \end{pmatrix} = \begin{pmatrix} \text{第 } k \text{ 个权重量} \\ \end{pmatrix} + \begin{pmatrix} \text{步长参数} \\ \end{pmatrix} \times \begin{pmatrix} \text{应用到第 } k \text{ 个权重的输入信号} \\ \end{pmatrix} \begin{pmatrix} \text{误差信号} \\ \end{pmatrix} \quad (4.114)$$

用 μ 表示步长参数。由图 4.28 可见, 在时间步骤 n 上, 应用到第 k 个抽头权重的输入信号为 $x[n-k]$ 。因此 $\hat{w}_k(n)$ 表示时间步骤 n 上第 k 个抽头权重的旧值, 则由式(4.114)可知, 时间步骤 $n+1$ 上第 k 个抽头权重的更新值为

$$\hat{w}_k[n+1] = \hat{w}_k[n] + \mu x[n-k]e[n], \quad k = 0, 1, \dots, N \quad (4.115)$$

其中

$$e[n] = a[n] - \sum_{k=0}^N \hat{w}_k[n]x[n-k] \quad (4.116)$$

这两个公式组成了用于自适应均衡器的 LMS 算法。注意不要将图 4.28 中自适应均衡器的长度与图 4.27 中的均衡器长度相混淆。

可使用矩阵符号来简化 LMS 算法的公式。用 $(N+1)$ 行 1 列的矢量 $\mathbf{x}[n]$ 表示均衡器的抽头输入

$$\mathbf{x}[n] = [x[n], \dots, x[n-N+1], x[n-N]]^T \quad (4.117)$$

其中, 上标 T 表示矩阵转置。相应地, 用 $(N+1)$ 行 1 列的矢量 $\hat{\mathbf{w}}[n]$ 表示均衡器的抽头权重:

$$\hat{\mathbf{w}}[n] = [\hat{w}_0[n], \hat{w}_1[n], \dots, \hat{w}_N[n]]^T \quad (4.118)$$

然后, 可用矩阵符号将式(4.112)中的卷积和改写为紧凑形式:

$$y[n] = \mathbf{x}^T[n]\hat{\mathbf{w}}[n] \quad (4.119)$$

其中, $\mathbf{x}^T[n]\hat{\mathbf{w}}[n]$ 为矢量 $\mathbf{x}[n]$ 和 $\hat{\mathbf{w}}[n]$ 的内积。现在, 我们可将自适应均衡器的 LMS 算法总结如下:

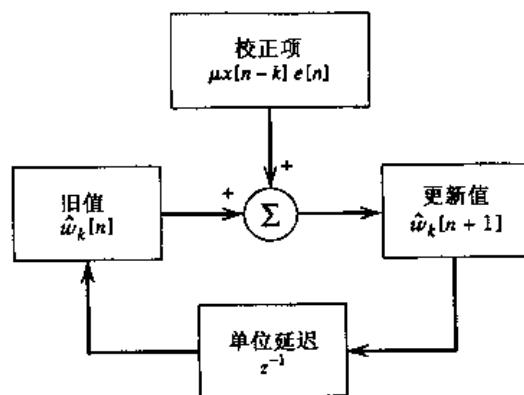
1. 通过令 $\hat{\mathbf{w}}[1] = 0$ (即在时刻 $t = T$ 且 $n = 1$ 时, 设均衡器的所有抽头权重为零) 来初始化该算法。
2. 对于 $n = 1, 2, \dots$, 计算

$$\begin{aligned} y[n] &= \mathbf{x}^T[n]\hat{\mathbf{w}}[n] \\ e[n] &= a[n] - y[n] \\ \hat{\mathbf{w}}[n+1] &= \hat{\mathbf{w}}[n] + \mu e[n]\mathbf{x}[n] \end{aligned}$$

其中, μ 为步长参数。

3. 继续迭代计算直到均衡器达到“稳定状态”, 这意味着均衡器的实际均方误差实质上达到了一个常数值。

LMS 算法是一个反馈系统的例子, 如图 4.29 的框图所示。该算法与第 k 个滤波器系数有关。因此, 该算法有可能是发散的(即自适应均衡器可能是不稳定的)。遗憾的是, 很难进行 LMS 算法的收敛性分析。但如果将步长参数 μ 设为一个较小的值, 在大量的迭代之后, LMS 算法的表现就粗略地近似于最佳梯度法(steepest-descent algorithm)。最佳梯度法是使用实际斜率而不是噪声估计来计算抽头权重的。(关于最佳梯度法的讨论见 3.13 节。)

图 4.29 涉及第 k 步抽头权重的 LMS 算法的信号流程图

4.10.2 均衡器的工作

自适应均衡器存在两种工作模式, 即训练模式和判决模式, 如图 4.30 所示。如前所述, 在训练模式中, 一个已知的 PN 序列被发送并有一个它的同步版本在接收机中产生, 经过等于传输时延的时移之后, 该同步版本作为期望响应被送入自适应均衡器。均衡器的抽头权重按照 LMS 算法进行调整。

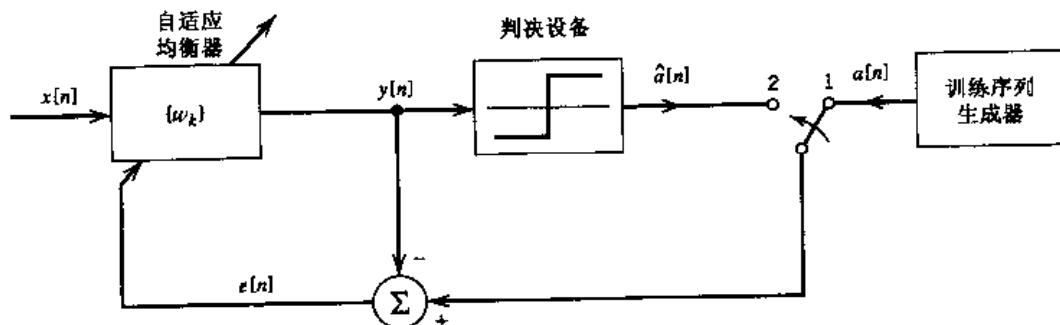


图 4.30 自适应均衡器的两种工作模式。对于训练模式, 开关在位置 1; 对于跟踪模式, 开关在位置 2

当训练过程完成后, 自适应均衡器就切换到其第二种工作模式: 判决模式。在该工作模式中, 误差信号定义为

$$e[n] = \hat{a}[n] - y[n] \quad (4.120)$$

其中, $y[n]$ 为均衡器在时刻 $t = nT$ 的输出, $\hat{a}[n]$ 为发送符号 $a[n]$ 最后(但不一定必要)的正确估计。此时, 在正常工作时, 接收机所进行的判决具有很高的正确概率。这意味着误差估计在绝大多数时间里是正确的。此外, 工作于判决模式的自适应均衡器能够跟踪信道特性相对缓慢的变化。

可以证明, 步长参数越大, 自适应均衡器的跟踪能力就越强。但是, 过大的步长会导致不可接受的很高的超均方误差。超均方误差的定义为误差信号超过最小可获得值 J_{\min} (当抽头的加权值为最优值时)的均方的部分。因此, 选择适当的步长参数 μ 实际上就是要兼顾快速跟踪和超均方误差两个方面。

4.10.3 判决反馈均衡器

为了进一步研究自适应均衡器,考虑一个冲激响应抽样形式为序列 $\{b[n]\}$ ($b[n] = b(nT)$)的基带信道。当不存在噪声时,该信道对输入序列 $\{x[n]\}$ 的响应由离散卷积和表示:

$$\begin{aligned} y[n] &= \sum_k b[k]x[n - k] \\ &= b[0]x[n] + \sum_{k<0} b[k]x[n - k] + \sum_{k>0} b[k]x[n - k] \end{aligned} \quad (4.121)$$

上式的第一项代表了希望得到的数据码元。第二项是由信道冲激响应的先前值(出现在主样值 $b[0]$ 之前)所引起的。第三项是由信道冲激响应的后继值(出现在主样值 $b[0]$ 之后)所引起的。信道冲激响应的先前值和后继值如图 4.31 所示。判决反馈均衡器^[12]的思想就是采用(基于信道冲激响应先前值的)数据判决来处理后继值。显然,先前值的判决必须是正确的。如果满足该条件,那么判决反馈均衡器就能够提供优于抽头延迟线均衡器的均衡效果。

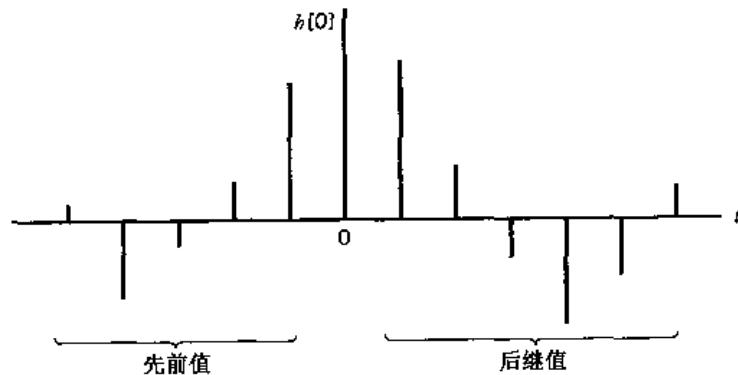


图 4.31 一个离散时间信道的冲激响应。先前值和后继值

判决反馈均衡器由一个前馈部分、一个反馈部分和一个判决设备组成,如图 4.32 所示。前馈部分由一个抽头间隔等于信号速率倒数的抽头延迟线滤波器组成。需要进行均衡的数据序列被送至该部分。反馈部分由另一个抽头间隔等于信号速率倒数的抽头延迟线滤波器组成。送至反馈部分的输入即判决设备的输出。反馈部分的作用就是从未来抽样的估计值中减去由判决设备输出的先前检测符号引入的 ISI 部分。

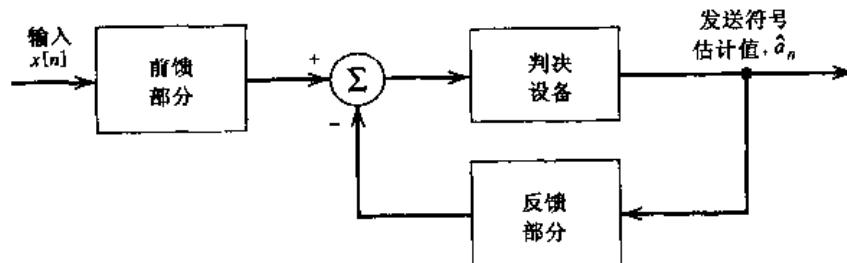


图 4.32 判决反馈均衡器的框图

值得注意的是,反馈环路中判决设备的引入使得均衡器本质上是非线性的。因此这种均衡器与一个普通的抽头延迟线均衡器相比更难进行分析。但是可采用均方误差准则来获得判决反馈均衡器的数学上易于处理的最佳形式。实际上,LMS 算法可以基于共同误差信号连带

地调节前向抽头权重和反馈抽头权重(参见习题 4.37)。

结合文献中对线性均衡器和判决反馈均衡器的比较性评价^[13] 可得,当线性信道的频率响应以严重的幅度失真或急剧的幅度截止作为其特性时,判决反馈均衡器就比等抽头数的线性均衡器性能好得多。这里假定判决反馈均衡器中的反馈判决都是正确的。以急剧的幅度截止为例(见“背景与预览”一章中图 8 所示的电话信道频率响应)。

和线性均衡器不同,判决反馈均衡器会产生错误传播。然而,尽管判决反馈均衡器是一个反馈系统,但错误传播不会无限持续下去。更准确地说,判决错误倾向于突发地出现。为了证明这种特性,我们提供了下述直观推理^[14]:

- 用 L 表示判决反馈均衡器中反馈部分的抽头个数。在连续 L 步正确判决之后,反馈部分中的所有判决错误将被冲掉。这里的错误传播的持续时间是有限的。
- 当一个判决错误产生时,下一个判决也错误的概率显然不超过 $1/2$ 。
- 用 K 表示错误传播的持续时间,即连续 L 步正确判决需要的符号数。则平均差错概率为 $(K/2)P_0$,其中, $K/2$ 为由单个判决错误引起的平均差错数, P_0 为前 L 个判决都正确的条件下的差错概率。
- 在一个公平的抛硬币实验中,为得到连续 L 个人头(代表无错误),需要的抛硬币的平均次数 K 为 $2(2^L - 1)$ 。

因此,判决反馈均衡器中错误传播的影响,就是使平均差错率大约等于第一次判决出错的概率乘以因子 2^L 。例如,对 $L = 3$,由于错误传播造成的平均差错率的增长小于一个数量级。

4.11 计算机实验:眼图

在本章的前几节中,我们已经讨论了为处理信道噪声和 ISI 对基带脉冲传输系统性能影响而采用的各种技术。最后也是最重要的,是怎样评价在实际通信环境中二者对整体系统性能的综合影响。一种评价工具就是所谓的眼图。眼图的定义为在一个特定的信号间隔内观察到的所关心信号(例如接收信号,接收机输出)的所有可能值的同步重叠。当采用二进制波形时,眼图的形状类似于人眼,因此而得名。眼图的内部区域称为眼睛开口(eye opening)。

眼图提供了与数据传输系统性能有关的大量有用信息,如图 4.33 所示。具体地说,可有如下结论:

- 眼睛开口的宽度定义了使得接收信号能够不受 ISI 引起的误差影响的抽样时间间隔。很明显,抽样的首选时刻是眼睛张开最宽的时刻。
- 系统对定时误差的敏感性定义为当抽样时刻变化时,眼睛闭合的速率。
- 在规定的抽样时刻,眼睛开口的高度定义了系统的噪声容限。

当 ISI 的影响非常严重时,来自眼图上部的迹线会与来自下部的迹线相交叉,结果眼睛就完全闭合了。此时系统中就不可避免地存在着由于 ISI 和噪声的共同影响而产生的误差。

在 M 进制系统的例子中,眼图包含了 $(M - 1)$ 个互相垂直堆叠在一起的眼睛开口,其中 M 是用于组成发射信号的离散幅度电平数。在带有真正的随机数据的严格线性系统中,所有这些眼睛开口都将是相同的。

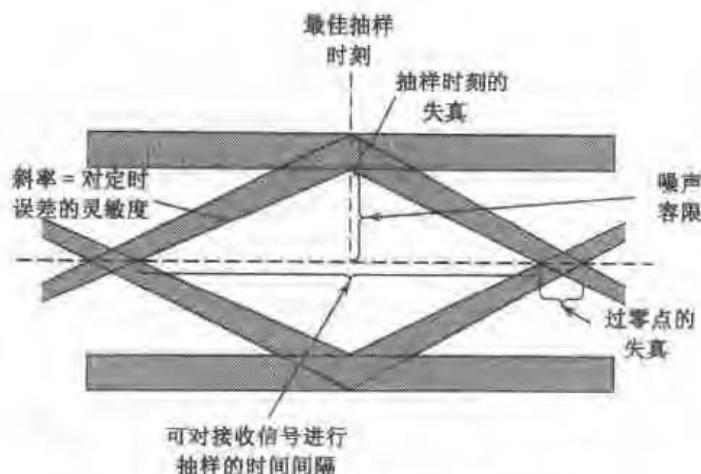


图 4.33 眼图的示意图

在接下来的两个实验中,我们利用计算机仿真来分别研究在无噪声、有噪声和带限条件下的四进制($M=4$)基带 PAM 传输系统的眼图。信道的非线性对眼图的影响将在习题 4.38 中进行讨论。

4.11.1 实验 1:信道噪声的影响

图 4.34(a)为理想条件下系统的眼图:没有信道噪声,也没有带宽限制。使用的信源符号由计算机随机产生并经过升余弦脉冲整形。用于产生眼图的系统参数如下:奈奎斯特带宽 $W=0.5$ Hz, 滚降因子 $\alpha=0.5$, 符号持续时间 $T=T_b \log^2 M = 2T_b$ 。图 4.34 的开口是理想的,这意味着系统性能是可靠的。注意该图有 $M-1=3$ 个开口。

图 4.34(b)和图 4.34(c)分别为存在信道噪声时系统的眼图。这两张图分别在信噪比 $SNR=20$ dB 和 10 dB 时进行仿真得到,其中信噪比在信道输出端测量得到。如图 4.34(b)所示,当 $SNR=20$ dB 时,信道噪声的影响几乎觉察不到。但当 $SNR=10$ dB 时,图 4.34(c)中的眼睛开口就几乎看不清了。

4.11.2 实验 2:带宽限制的影响

图 4.35(a)和图 4.35(b)为与实验 1 中参数相同的四进制系统的眼图,但本实验是在带限条件下的无噪声信道中进行的。具体地说,此时的信道模型就是一个平方幅度响应如下的低通巴特沃兹(Butterworth)滤波器:

$$|H(f)|^2 = \frac{1}{1 + (f/f_0)^{2N}}$$

其中, N 为滤波器级数, f_0 为截止频率, 其值为 3 dB。在图 4.35(a)所对应的计算机实验中,有:

$$N = 25 \text{ 和 } f_0 = 0.975 \text{ Hz}$$

PAM 传输系统所要求的带宽为

$$B_T = W(1 + \alpha) = 0.75 \text{ Hz}$$

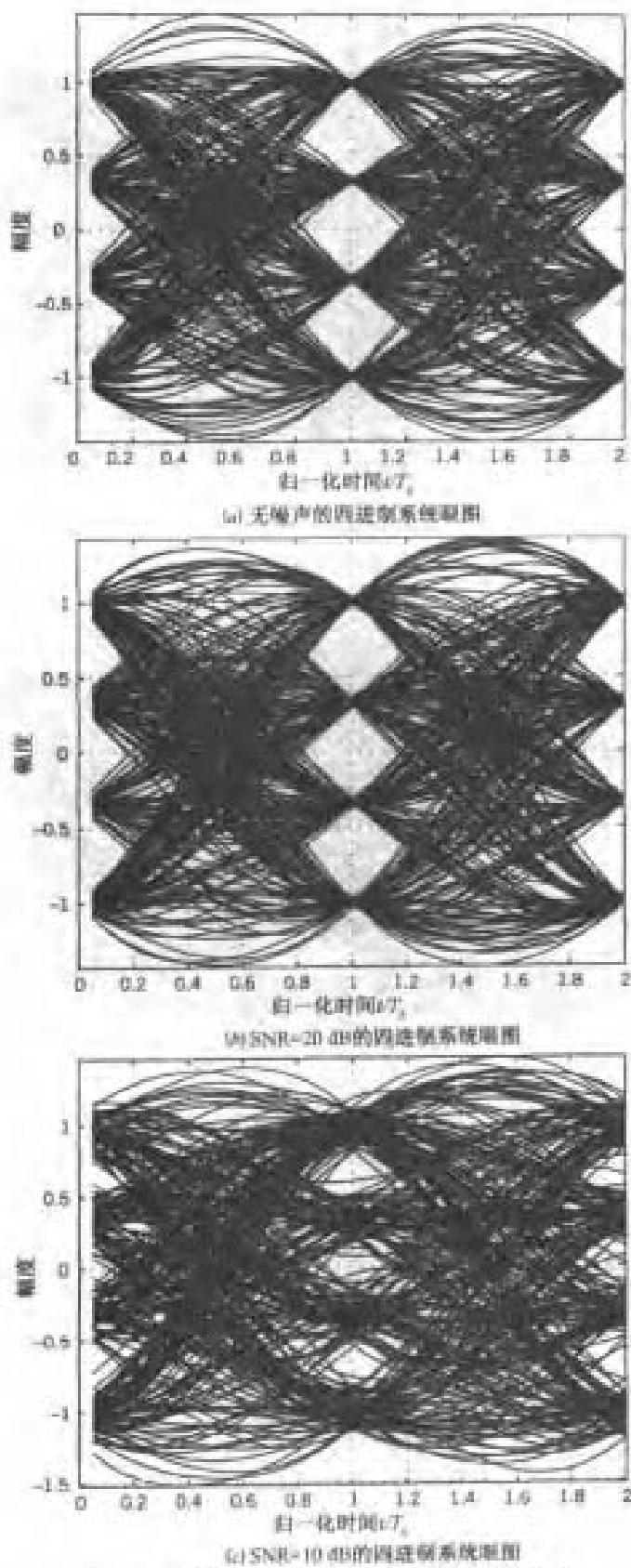


图 4.34 无噪声与有噪声时的四进制系统眼图

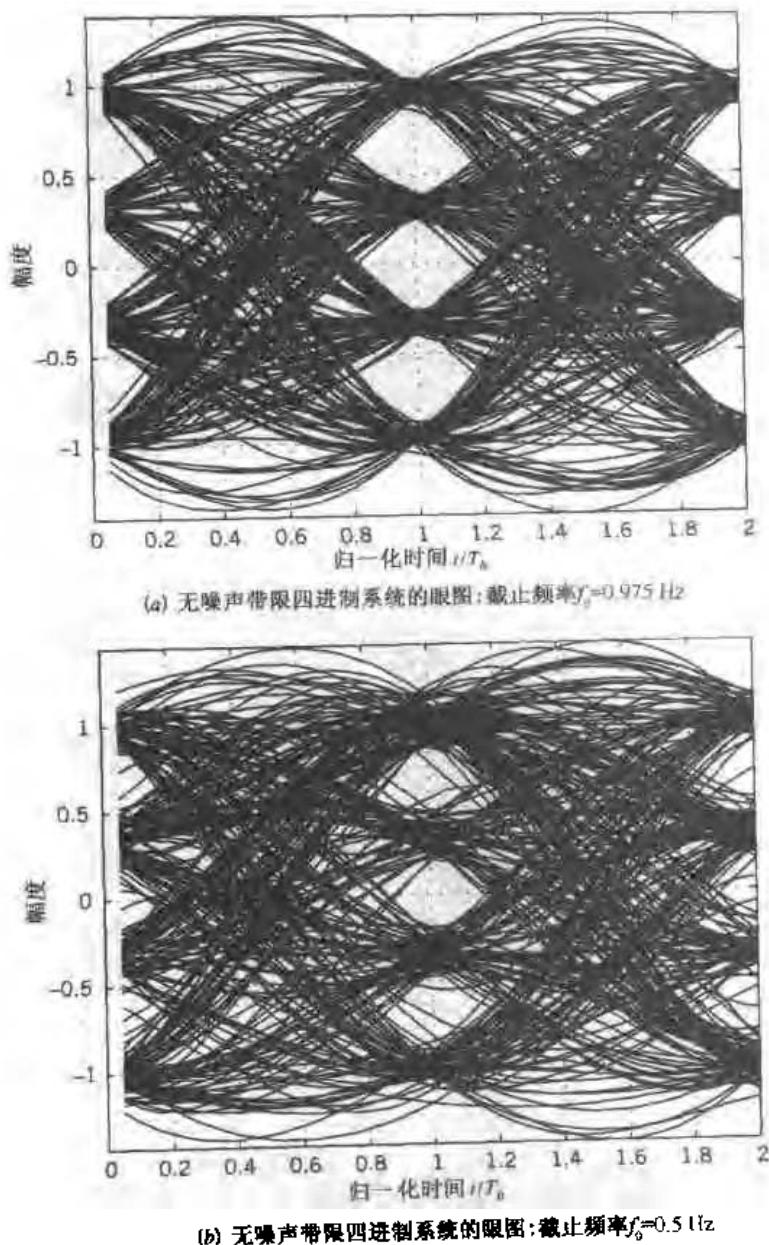


图 4.35 无噪声带限四进制系统的眼图

尽管信道带宽(即截止频率)大于需要值,但与图 4.34(a)中相比,信道带宽对通带的影响仍表现为眼睛开口尺寸的减少。如图 4.34(a)所示,在时刻 $t = 1\text{ s}$ 时眼图不再是可区分的值,而是一个模糊区域。

通过将信道建模为一个 $N = 25$ 且 $f_0 = 0.5\text{ Hz}$ 的低通巴特沃兹滤波器,图 4.35(b)中的信道带宽得以进一步减小。信道带宽的减小将进一步地减小眼睛张开的程度。

4.12 总结与讨论

在本章中,我们研究了信道噪声和 ISI 对基带脉冲传输系统性能的影响。ISI 不同于噪声,因为它是由于信道的频率响应偏离理想低通滤波器(奈奎斯特信道)而产生的依赖于信号的干

扰形式。当没有发射信号时, ISI 也就消失了。这种偏离的结果是导致接收脉冲要受到先前脉冲尾部和后继脉冲前端的影响。

根据接收信噪比,对于具有固定特性的信道,我们可以区分在基带脉冲传输系统中所出现的三种不同情况:

1. 相对于信道噪声而言, ISI 的影响可以忽略。

在这种情况下,合适的方法就是采用匹配滤波器。匹配滤波器是最大化峰值脉冲信噪比下的最佳线性时不变滤波器。

2. 接收信噪比足够高,以至于可以忽略信道噪声的影响。

在这种情况下,我们需要警惕 ISI 对接收机中信号重构的影响。特别地,必须对接收脉冲波形加以控制。这可以用两种不同方法之一来实现:

- 采用升余弦频谱作为基带脉冲传输系统的整体频率响应。

- 采用相关电平编码或部分响应信号,以可控方式将 ISI 引入到发射信号中。

3. ISI 和信道噪声都不能忽略。

对于这种更复杂情况,为了得到易于进行数学处理的解,可采用均方误差准则。所得的最佳线性接收机称为最小均方误差接收机。这种接收机由匹配滤波器和线性横向(抽头延迟线)均衡器级联而成。

信道是随机的,如在电信环境中就经常出现这种情况。此时,基于平均信道特性而设计的匹配滤波器就不能满足要求了。在这种情况下,较好的办法就是采用自适应均衡器,其目的就是对数据传输期间信道频率响应的变化进行自动补偿。抽头延迟线滤波器和调整抽头权重的 LMS 算法的结合使用,为实现自适应均衡器的简单但非常有效的方法提供了基础。自适应均衡器能够处理 ISI 和在非稳态环境中的接收机噪声的混合影响。它的实际用途在于,如今商用的几乎每一只用在语音级电话信道上进行数据传输的调制解调器(调制器-解调器)都将自适应均衡器作为不可缺少的部分。

自适应滤波器的另一个重要应用是在 DSL 收发器关键部分(即回波抵消器)的设计中。DSL 通常使用双绞线作为传输媒质,这和普通电话信道一样。和电话信道不同的是,DSL 是用来在数字网络与用户终端之间提供 64 Kbps 或更高速率的高速数据传输的。

注释与参考

- [1] 关于基带脉冲传输的经典书籍有 Lucky, Salz 和 Weldon(1968)以及 Sunde(1969)。要得到相关的其他方面的详细资料,见 Gitlin, Hayes 和 Weinstein(1992), Proakis(1995)和 Benedetto 以及 Biglieri 和 Castellani(1987)。
- [2] 对匹配滤波器特性的描述首次出现在 North 的一篇分类报告中(RCA 实验室报告 PTR-6C, 1943 年 6 月),该报告在 20 年后发表,见 North(1963)。Van Vleck 和 Middleton(1946)也独立获得了相似的结果。要参阅有关匹配滤波器及其特性的材料,见 Turin(1960, 1976)。
- [3] 误差函数,记为 $\text{erf}(u)$,在该文献中由一些不同方法所定义。我们采用的是下面的定义:

$$\operatorname{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u \exp(-z^2) dz$$

误差函数有下面两个有用特性：

(i) $\operatorname{erf}(-u) = -\operatorname{erf}(u)$

这个特性称为对称相关。

(ii) 当 u 接近于无穷大时, $\operatorname{erf}(u)$ 接近于 1, 即

$$\frac{2}{\sqrt{\pi}} \int_0^\infty \exp(-z^2) dz = 1$$

互补误差函数定义为

$$\operatorname{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^\infty \exp(-z^2) dz$$

它和误差函数的关系如下：

$$\operatorname{erfc}(u) = 1 - \operatorname{erf}(u)$$

表 A6.6 给出了 u 在 0 到 3.3 之间时的误差函数 $\operatorname{erf}(u)$ 的值。

对一个较大的正数 u 来说, $\operatorname{erfc}(u)$ 有两个简单界限, 一个下界和一个上界, 即

$$\frac{\exp(-u^2)}{\sqrt{\pi u}} \left(1 - \frac{1}{2u^2} \right) < \operatorname{erfc}(u) < \frac{\exp(-u^2)}{\sqrt{\pi u}}$$

互补误差函数为 4.3 节中的符号差错概率的紧凑形式提供了基础。在该文献中经常使用的另一个函数为 Q 函数。考虑一个标准化的具有零均值和单位方差的高斯随机变量 X 。随机变量 X 的一个观察值大于 v 的概率由 Q 函数给出：

$$Q(v) = \frac{1}{\sqrt{2\pi}} \int_v^\infty \exp\left(-\frac{x^2}{2}\right) dx$$

Q 函数定义了标准化的高斯序列下的面积。 Q 函数与互补误差函数的关系如下：

$$Q(v) = \frac{1}{2} \operatorname{erfc}\left(\frac{v}{\sqrt{2}}\right)$$

反过来, 设 $u = v/\sqrt{2}$, 有

$$\operatorname{erfc}(u) = 2Q(\sqrt{2}u)$$

- [4] 式(4.49)或式(4.53)中所描述的标准首先由奈奎斯特在研究电报传输理论时明确地表达出来。奈奎斯特在 1928 年的论文就是一个经典。在该论文中, 该标准称为奈奎斯特第一准则。在 1928 年的论文中, 奈奎斯特描述了另一种方法, 称为奈奎斯特第二准则。这种方法与其说利用了接收信号中的中央样值, 倒不如说是利用了不同符号之间的过渡瞬间。对奈奎斯特第一和第二准则的讨论见 Bennett (1970, pp. 78–92) 和 Gibby 与 Smith (1965)。而奈奎斯特第三准则在 Sunde (1969) 中进行了讨论, 也可以参考 Pasupathy (1974) 和 Sayar 与 Pasupathy (1987)。
- [5] 相关电平编码和部分响应信号是同义的, 在该文献中这两个术语都有使用。相关电平编码的思想首先见于 Lender (1963)。Lender 的工作由 Kretzmer (1966) 归纳为二进制数据传输。有关相关电平编码技术的进一步细节, 见 Gitlin, Hayes 和 Weinstein (1992), 也可参见 Pasupathy (1977), Kabal, Pasupathy (1975), 以及 Sousa 和 Pasupathy

(1983)。

- [6] 在4.8节中提到的DSL的资料基于Lin和Tzeng(1988), Lechleider(1989)的两篇论文以及Starr,Cioffi,以及Silverman(1999)和Chen(1998)。
- [7] 有关用于数字用户环路的线路码的讨论,参见Gitlin等人(1992)的论文。
- [8] 在Ericson(1971)中,对于每个“合理”的性能准则,最佳接收机可以由一个匹配滤波器接一个抽头延迟线均衡器来实现,如图4.27所示。作为对4.9节中考虑的均方误差准则的补充,我们关心的合理的性能准则包括:
 - (i) 符号差错概率的最小化。
 - (ii) 迫零均衡(将ISI减小到零),以及该约束条件下的最小符号差错概率。
 - (iii) 在抽样瞬间信噪比的最小化。
- 标准(i)是线性接收机最优化的最本质的方法。Aaron和Tufts(1966)曾对这种方法进行了后续研究,但遗憾的是,由于太复杂而无法给出成形的结论。由Lucky等人继续研究的准则(ii)就是一个简化了许多的方法。准则(iii)见George(1965)。
- [9] 电话信道的自适应均衡首先由Lucky(1965,1966)提出。从那时开始,一大批自适应均衡的方案陆续发表,这为具体的同步数据传输系统提供了均衡方案。有关自适应均衡的内容,可参阅Proakis(1975)和Qureshi(1982,1985)。自适应均衡的细节在Gitlin,Hayes和Weinstein(1992,Chapter 8)及Proakis(1995,Chapter 6)中也有讨论。
- [10] 有关分段间隔的均衡器的早期研究工作开始于Brady(1970)。Ungerboeck(1976)和Gitlin及Weinstein(1981)进行了后续工作。有关分段间隔均衡器的详细讨论见Gitlin等人(1992)的论文。
- [11] LMS算法由Widrow和Hoff,Jr.(1960)首先提出。有关LMS算法的详细集中分析,见Haykin(1996,Chapter 9),以及Widrow和Stearns(1985,Chapter 6)。
- [12] 判决反馈均衡器首先由Austin(1967)进行了讨论。用于最小均方误差的最优化判决反馈均衡器首先由Monsen(1971)完成。判决反馈均衡器的简易说明见Gitlin,Hayes和Weinstein(1992,pp.500–510)。Tomlinson(1971),Harashima以及Miyakawa(1972)提出了一个用于消除判决反馈均衡器中误差传播的设备。该设备称为Tomlinson-Harashima预编码器,在发射中作为调制器的预处理器而存在。这种预编码器的基本思想是把判决反馈均衡器中的反馈部分搬到发射机中,这样就不可能产生判决误差了。但这种改进可能会导致发射功率大大增加。绝大多数情况下,采用模算法来克服这种功率增加。
- [13] 关于线性均衡器与判决反馈均衡器性能的比较,参见Gitlin等人(1992)的著作以及Proakis(1995)。
- [14] 对章节4.10提出的判决反馈均衡器中的误差传播的直观讨论参见Gitlin等人(1992)的著作。
- [15] 对存在误差传播的判决反馈均衡器的符号差错概率的严格论述,见Duttweiler等人(1974)。该论文表明,在最差的ISI情况下, P_e 需乘以因子 2^L (对应于使得在高信噪比下不会发生判决错误的差错概率),其中 L 为反馈部分的抽头数。Duttweiler等人的研究结果为4.10节的直观参数提供了理论性的佐证。

习题

匹配滤波器

4.1 考虑图 P4.1 所示的信号 $s(t)$ 。

- (a) 确定与该信号匹配的滤波器的冲激响应并绘制其时间函数草图。
- (b) 绘制该匹配滤波器输出的时间函数图。
- (c) 该输出的峰值是多少?

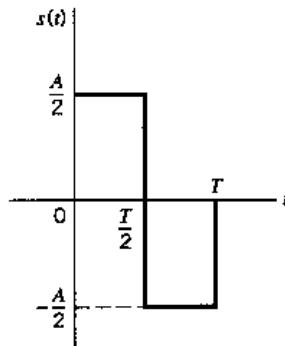


图 P4.1

4.2 图 P4.2(a)显示了一对在间隔 $[0, T]$ 上相互正交的脉冲。在本题中, 我们通过考察这种脉冲对的使用来研究二维匹配滤波器。

- (a) 分别确定脉冲 $s_1(t)$ 和 $s_2(t)$ 各自的匹配滤波器, 其中 $s_1(t)$ 的匹配滤波器与习题 4.1 中相同。
- (b) 通过并行连接(a)部分中的两个匹配滤波器, 组成一个二维匹配滤波器, 如图 P4.2(b)所示。试证明下列结论:
 - (i) 当脉冲 $s_1(t)$ 被送入该二维滤波器时, 下方匹配滤波器的响应为零。
 - (ii) 当脉冲 $s_2(t)$ 被送入该二维滤波器时, 上方匹配滤波器的响应为零。归纳你考察的结论。

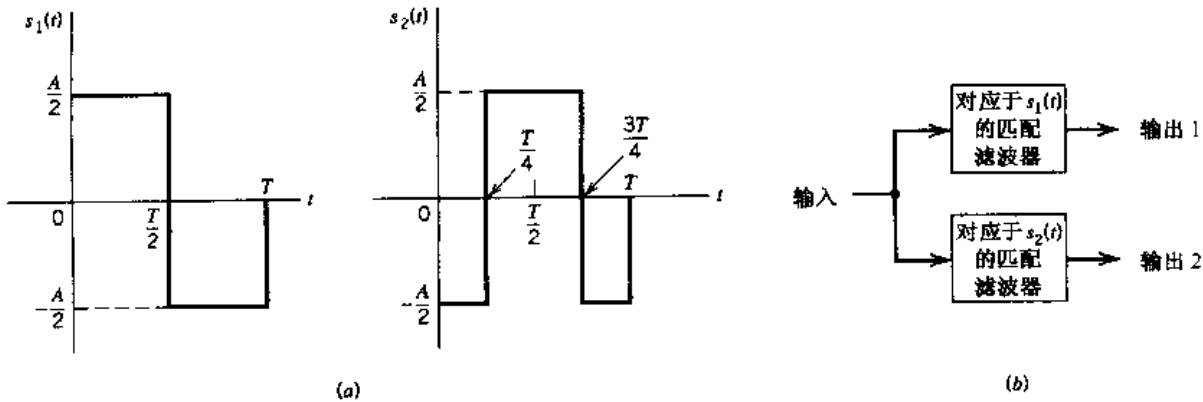


图 P4.2

4.3 设有一个定义如下的矩形脉冲：

$$g(t) = \begin{cases} A, & 0 \leq t \leq T \\ 0, & \text{其他} \end{cases}$$

假设用一个带宽为 B 的理想低通滤波器来近似 $g(t)$ 对应的匹配滤波器。本题的主要目标是求峰值脉冲信噪比的最大值。

- (a) 确定 B 的最优值,使理想低通滤波器提供对该匹配滤波器的最佳近似。
- (b) 理想低通滤波器与该匹配滤波器相差多少 dB?
- 4.4 本题探索另一种匹配滤波器的近似实现方法。采用一个简单的电阻 - 电容(RC)低通滤波器,如图 P4.4 所示。该滤波器的频率响应为

$$H(f) = \frac{1}{1 + if/f_0}$$

其中, $f_0 = 1/2\pi RC$ 。输入信号 $g(t)$ 是一个幅度为 A 且持续时间为 T 的矩形脉冲。要求对该滤波器的 3 dB 截止频率 f_0 的选择进行最优化,使得滤波器输出端的峰值脉冲信噪比为最大。并证明 f_0 的最佳值为 $0.2/T$,此时匹配滤波器损失的信噪比大约为 1 dB。

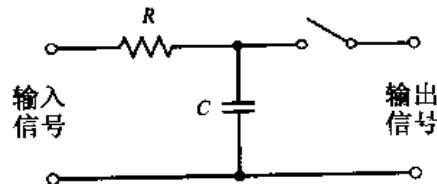


图 P4.4

差错概率的计算

- 4.5 图 4.4 所示接收机的最佳门限一般由式(4.37)给出。在下列两种情况下,根据图 4.4 分析最佳门限是怎样影响式(4.35)中两个平均符号差错概率 P_e 分量的:

- (a) $p_0 > p_1$
- (b) $p_0 < p_1$

其中, p_0 和 p_1 分别是符号 0 和 1 的先验概率。

- 4.6 在一个二进制 PCM 系统中,符号 0 和 1 分别具有先验概率 p_0 和 p_1 。随机变量 Y (其抽样值为 y) 的条件概率密度函数是在一个信号间隔的结束时刻,通过对图 4.4 接收机中匹配滤波器的输出进行抽样而获得。若发送的是符号 0,将该密度函数记为 $f_Y(y|0)$ 。类似地,若发送的是符号 1,则将该密度函数记为 $f_Y(y|1)$ 。用 λ 表示接收机使用的门限值。如果抽样值 $y > \lambda$,接收机就会判决为符号 1。否则判决为符号 0。证明使平均差错概率最小的最佳门限值 λ_{opt} 由下式给出:

$$\frac{f_Y(\lambda_{opt}|1)}{f_Y(\lambda_{opt}|0)} = \frac{p_0}{p_1}$$

- 4.7 一个使用极性 NRZ 信号的二进制 PCM 系统以等于 10^{-6} 的平均差错概率工作于差

错门限之上。假设信号速率变为 2 倍。请找到新的平均差错概率。可利用表 A6.6 来计算互补误差函数。

- 4.8 一个连续时间信号被抽样并作为 PCM 信号发送。接收机中判决设备输入端的随机变量有 $0.01 \text{ V}^{[2]}$ 的方差。

(a) 假设使用极性 NRZ 信号, 确定要使平均差错率为在 10^8 比特中只有不超过 1 个出错, 发送的脉冲幅度为多少?

(b) 如果有附加的干扰出现, 导致差错率增加到在 10^6 比特中有 1 个出错, 那么该干扰的方差是多少?

- 4.9 一个二进制 PCM 系统使用非极性 NRZ 信号发送符号 1 和 0。符号 1 由一个幅度为 A 且持续时间为 T_b 的矩形脉冲表示。信道噪声模型为均值为零, 功率谱密度为 $N_0/2$ 的加性高斯白噪声。假设符号 1 和 0 等概率出现。采用 4.3 节中描述的匹配滤波器。试确定接收机输出端的平均差错率的表达式。

- 4.10 假设采用习题 4.9 中的单极性 RZ 信号。符号 1 由一个幅度为 A 且持续时间为 $T_b/2$ 的脉冲表示, 符号 0 由不发送脉冲表示。

试证明: 为了获得同样的平均符号差错概率, 这种形式的单极性信号需要两倍于单极性 NRZ(例如, 通 - 断)信号的平均功率。

- 4.11 采用图 4.4 所示的 PCM 接收机双极性 NRZ 信号。发射信号 $s(t)$ 为

二进制符号 1: $s(t) = \pm A, 0 < t \leq T$

二进制符号 0: $s(t) = 0, 0 < t \leq T$

假设二进制符号 0 和 1 是等概率的, 计算这种接收机的平均符号差错概率 P_s 。

升余弦谱

- 4.12 图 P4.12 中的 NRZ 脉冲可以看成是奈奎斯特脉冲的一个原型。比较这两种脉冲的频谱特性。

- 4.13 试求式(4.60)中定义的频率函数 $P(f)$ 的傅里叶反变换。

- 4.14 模拟信号经抽样、量化后, 编码为二进制 PCM 信号。该 PCM 系统的具体参数如下:

抽样率 = 8 kHz

量化级数 = 64

该 PCM 信号采用离散脉冲幅度调制并经基带信道传输。

如果每个脉冲允许采用的幅度电平数为 2, 4 或 8, 求传输该 PCM 信号所需要的最小带宽。

- 4.15 考虑一个被设计为具有升余弦谱 $P(f)$ 的基带二进制 PAM 系统。输出脉冲 $p(t)$ 由式(4.62)定义。要使该系统具有线性相位响应, 该脉冲应进行怎样的调整?

- 4.16 计算机以 56 Kbps 的速率输出二进制数据。计算机输出通过具有升余弦频谱的二进制 PAM 系统发送。求下列各滚降因子所要求的传输带宽: $\alpha = 0.25, 0.5, 0.75, 1.0$ 。

- 4.17 重复习题 4.16 的计算。其中条件变为已知在计算机输出中的每一组连续三个二进制数字被编码为 8 个可能幅度电平之一, 且生成信号采用一个具有升余弦谱的 8 电

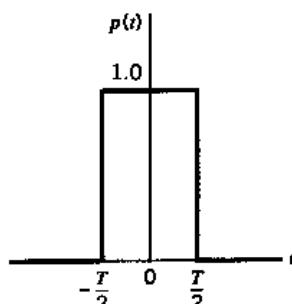


图 P4.12

平 PAM 系统发送。

- 4.18 模拟信号经抽样、量化后，编码为二进制 PCM 波形。使用的量化级数为 128。将一个同步脉冲加到每个代表模拟抽样的码字的末端。最终的 PCM 信号采用一个具有升余弦频谱的四进制 PAM 系统，经过带宽为 12 kHz 的信道发送。滚降因子为 1。
 (a) 求信息通过该信道传输的速率(bps)。
 (b) 求模拟信号的抽样率。对于该模拟信号的最高频率分量，抽样率的最大可能值是多少？
- 4.19 二进制 PAM 信号经一个绝对最大带宽为 75 kHz 的基带信道传输。信号比特持续时间为 $10 \mu\text{s}$ 。求一个满足上述要求的升余弦频谱。

相关电平编码

- 4.20 双二进制、三进制以及双极性信号技术有一个共同特征：具有三个幅度电平。试问：双二进制技术与其他两种技术有何差别？
- 4.21 二进制数据流 001101001 被送至一个双二进制系统的输入端。
 (a) 在不使用预编码器的情况下，求双二进制编码器的输出和相应接收机的输出。
 (b) 假设由于传输期间出现错误，在接收机输入端由第二个数字产生的电平变为零。求新的接收机输出。
- 4.22 重复习题 4.21 的计算，其中假设在发射机中采用了预编码器。
- 4.23 图 P4.23 所示方案可看成是一个差分编码器（由一个模 2 加法器和一个单位延迟单元组成）与一个特殊形式的相关编码器（由一个单位延迟单元和加法器组成）的级联。其中单位延迟单元如图 P4.23 所示。差分编码器和相关编码器中各有一个单位延迟单元。在该差分编码器中，符号 0 代表有变化，符号 1 代表无变化。

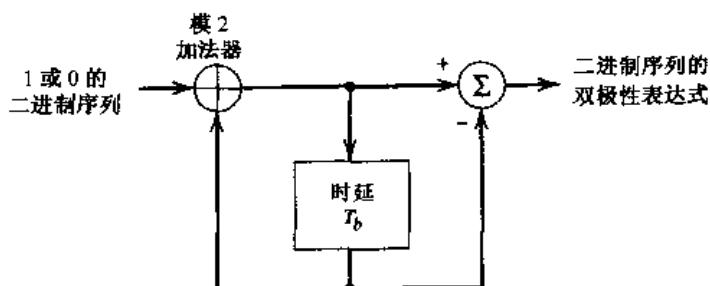


图 P4.23

- (a) 求图 P4.23 所示方案中相关编码器的频率响应以及冲激响应。
 (b) 证明该方案可以用来将一个二进制序列（送至输入端）的通 - 断信号，转变为输出端序列的双极性信号。可以序列 010001101 为例来演示这种转变。
 对二进制序列的通 - 断、双极性和差分编码的描述，见 3.7 节。
- 4.24 考虑一个随机二进制信号 $x(t)$ 。其中，1 和 0 等概率出现。相邻时间帧的符号统计独立。符号 1 由 A 伏表示，符号 0 由零伏表示。这种通 - 断二进制波形被送入图 P4.23 所示的电路中。
 (a) 利用习题 4.23 中的结论，证明在该电路输出端出现的双极性信号 $y(t)$ 的功率

谱密度等于

$$S_X(f) = T_b A^2 \sin^2(\pi f T_b) \operatorname{sinc}^2(f T_b)$$

(b) 分别绘制通 - 断信号和双极性二进制信号的功率谱密度图，并对它们进行比较。

4.25 二进制数据流 011100101 被送至一个改进的双二进制系统的输入端。

(a) 在不使用预编码器的情况下，求该改进的双二进制编码器的输出以及相应接收机的输出。

(b) 假设由于传输期间出现错误，由第三个数字产生的电平变为零。求新的接收机输出。

4.26 重复习题 4.25 的计算，其中假设在发射机中采用了预编码器。

M 进制 PAM 系统

4.27 考虑一个采用 M 个离散幅度电平的基带 M 进制系统。接收机模型如图 P4.27 所示。接收机在如下条件下工作：

(a) 接收波形中的信号分量为

$$m(t) = \sum_n a_n \operatorname{sinc}\left(\frac{t}{T} - n\right)$$

其中， $1/T$ 为信号的波特率。

(b) 如果 M 为偶数，则幅度电平为 $a_n = \pm A/2, \pm 3A/2, \dots, \pm (M-1)A/2$ 。如果 M 是奇数，则幅度电平为 $a_n = 0, \pm A, \dots, \pm (M-1)A/2$ 。

(c) M 个电平是等概的，且在相邻时隙中发送的符号统计独立。

(d) 信道噪声 $w(t)$ 是均值为零，功率谱密度为 $N_0/2$ 的高斯白噪声。

(e) 该低通滤波器是带宽 $B = 1/2T$ 的理想滤波器。

(f) 若 M 为偶数，则判决设备中使用的门限电平为 $0, \pm A, \dots, \pm (M-2)A/2$ 。若 M 为奇数，则门限电平为 $\pm A/2, \pm 3A/2, \dots, \pm (M-2)A/2$ 。

该系统中的平均符号差错概率定义为

$$P_e = \left(1 - \frac{1}{M}\right) \operatorname{erfc}\left(\frac{A}{2\sqrt{2}\sigma}\right)$$

其中， σ 为判决设备输入端噪声的标准方差。试通过确定 $M = 2, 3, 4$ 时的 P_e 来证明该一般公式的正确性。

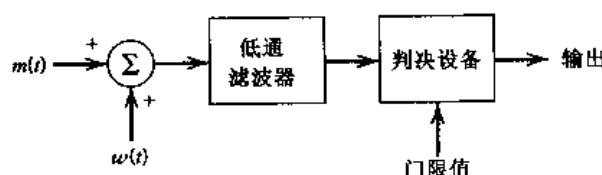


图 P4.27

4.28 假设在如习题 4.27 所述的具有 M 个等概率幅度电平的基带 M 进制 PAM 系统中，

平均符号差错概率 P_e 小于 10^{-6} , 因此译码时的错误可以忽略不计。证明在该系统中接收信噪比的最小值可以近似地表示为

$$(\text{SNR})_{\min} \approx 7.8(M^2 - 1)$$

DSL

4.29 DSL 中的串话幅度分布的数学模型可以是高斯型的。请证明这样一个模型的正确性。提示:一个电缆通常包括很多条双绞线。

4.30 (a) 推导使用 2B1Q 线路码的发射信号的功率谱密度公式。

(b) 绘出下列线路码的功率谱:

- 曼彻斯特码
- 改进的双二进制码
- 双极性归零码
- 2B1Q 码

并就这些线路码在数字用户环路中的适应性来比较其品质的相对优劣。

4.31 本题中采用 LMS 算法来精确表达用于 DSL 的自适应回波抵消器。自适应回波抵消器的基本原理是, 合成一个回波的复制品, 并以自适应的方式从返回信号中减掉它, 如图 P4.31 所示。合成的回波用 $\hat{r}[n]$ 表示。将发射信号通过一个与回波路径的传输函数理想匹配的自适应滤波器, 就能够产生合成回波。返回的接收信号为实际回波 $r[n]$ 和接收信号 $x[n]$ 之和。返回的接收信号可看成是自适应滤波过程的期望响应。

试利用 LMS 算法, 精确描述图 P4.31 所示的自适应回波抵消器的工作原理。

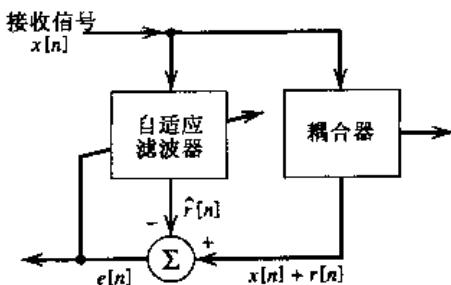


图 P4.31

均衡

4.32 图 P4.32 所示为一个线性信道和一个同步抽头延迟线均衡器的级联。该信道的冲激响应为 $c(t)$, 均衡器的冲激响应为 $b(t)$ 。 $b(t)$ 的定义为

$$b(t) = \sum_{k=-N}^N w_k \delta(t - kT)$$

其中, T 为均衡器中相邻抽头的间隔, w_k 为其抽头权重(系数)。图 P4.32 所示级联系统的冲激响应计为 $p(t)$ 。 $p(t)$ 以 $1/T$ 的速率进行均匀抽样。为消除 ISI, 就必须满足奈奎斯特准则, 即

$$p(nT) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

(a) 假定系统满足奈奎斯特准则, 试证明最终的迫零均衡器的 $(2N + 1)$ 个抽头权重满足下面这一组 $(2N + 1)$ 个同步公式:

$$\sum_{k=-N}^N w_k c_{n-k} = \begin{cases} 1, & n = 0 \\ 0, & n \neq \pm 1, \pm 2, \dots, \pm N \end{cases}$$

其中, $c_n = e(nT)$ 。从而证明迫零均衡器是一个反向滤波器, 因为它的传输函数等于信道传输函数的倒数。

(b) 迫零均衡器的一个缺点是在存在信道噪声的情况下会出现导致系统性能下降的噪声增加现象。为了研究该现象, 考虑一个低通信道。该信道在奈奎斯特频率处有一个切口, 即 $H(f)$ 在 $f = 1/2T$ 处为零。假设信道噪声是加性白噪声, 证明均衡器输出端的噪声功率谱密度在 $f = 1/2T$ 处达到无穷大。

即使该信道在其频率响应中没有切口, 在均衡器输出端的噪声功率谱密度值也可能会很大。证明这个一般性结论。

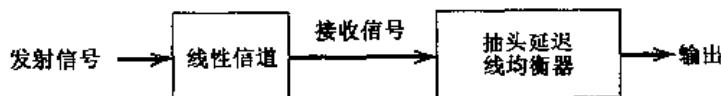


图 P4.32

4.33 式(4.108)定义了最小均方误差接收机的冲激响应。

(a) 证明式(4.109)的有效性。式(4.109)是式(4.108)的傅里叶变换。

(b) 式(4.109)中的功率谱密度 $S_q(f)$ 是时间函数 $q(t)$ 的自相关函数 $R_q(\tau_1, \tau_2)$ 的傅里叶变换。 $R_q(\tau_1, \tau_2)$ 由式(4.98)定义。由式(4.98)出发, 推导式(4.111)。

4.34 一些无线系统中存在多径失真。多径失真是由于发射机和接收机之间存在超过一条的传播路径而引起的。考虑信号 $s(t)$ 的具有如下响应输出的信道(不存在噪声):

$$x(t) = a_1 s(t - t_{01}) + a_2 s(t - t_{02})$$

其中, a_1 和 a_2 为常量, t_{01} 和 t_{02} 为传输时延。假设采用图 P4.34 中的三抽头延迟线滤波器来均衡该信道产生的多径失真。

(a) 计算该信道的传输函数。

(b) 假设 $a_2 \ll a_1$ 且 $t_{02} > t_{01}$, 计算该抽头延迟线滤波器的参数 a_1 , a_2 , t_{01} 和 t_{02} 。

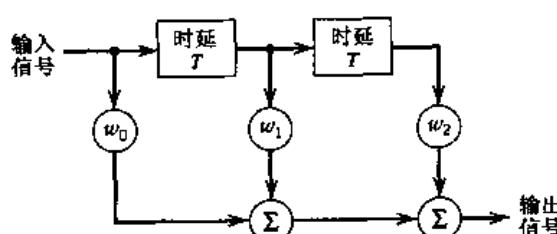


图 P4.34

- 4.35 用序列 $\{x(nT)\}$ 表示抽头延迟线均衡器的输入。试证明：如果均衡器的频率响应满足下式，那么 ISI 就可以完全消除：

$$H(f) = \frac{T}{\sum_b X(f - k/T)}$$

其中， T 为符号持续时间。

当均衡器的抽头数接近无穷大时，均衡器的频率响应就成了一个具有实系数的傅里叶级数，从而能够在间隔 $(-1/2T, 1/2T)$ 内逼近任何函数。证明均衡器的这一性质。

- 4.36 步长参数 μ 在 LMS 算法中具有关键作用。讨论以下两点：

- (a) 稳定性。如果 μ 超过了某个临界值，算法将发散（即系统变得不稳定）。
- (b) 记忆性。 μ 的倒数可以看成算法记忆性的度量；当 μ 变小时，更多的输入信号的先前样值将影响算法的工作。

- 4.37 用矢量 $w^{(1)}[n]$ 和 $w^{(2)}[n]$ 代表图 4.32 中判决反馈均衡器的前馈和反馈部分的权重。试求调整该均衡器权重的 LMS 算法公式。

计算机实验

- 4.36 在 4.11 节，我们研究了四进制 ($M = 4$) PAM 基带传输系统在噪声和带限条件下的眼图。在该实验中，假设信道是线性的，在具有随机数据的严格线性系统中，所有的眼图开口一样大。但实际的眼图常常表现出非对称性，这是由通信信道的非线性引起的。

本实验研究非线性信道对眼图的影响。特别地，我们重复图 4.34(a) 所示 $M = 4$ 的无噪眼图的计算机实验，但这次假设信道是非线性的，并具有以下输入 – 输出关系：

$$x(t) = s(t) + as^2(t)$$

其中， $s(t)$ 为信道输入， $x(t)$ 为信道输出， a 为常数。

- (a) 在 $a = 0, 0.05, 0.1, 0.2$ 时，重复该实验。
- (b) 讨论 a 的变化如何影响眼图的形状。

- 4.39 本实验研究 Chennakeshu 和 Saulnier(1993) 提出的根升余弦脉冲。该脉冲用 $p(t)$ 表示，并具有以下性质：

- 脉冲 $p(t)$ 是关于时间对称的，即 $p(-t) = p(t)$ 。
- $p(t)$ 傅里叶变换的平方，即 $P^2(f)$ ，满足式(4.60)的升余弦频谱。但是其傅里叶变换本身即 $P(f)$ 则不满足该式。
- 脉冲 $p(t)$ 满足正交约束

$$\int_{-\infty}^{\infty} p(t)p(t - nT) dt = 0, \quad n = \pm 1, \pm 2, \dots$$

其中， T 为符号间隔。

- (a) 计算滚降因子 $\alpha = 0.3$ 时，二进制数据流 101100 的基带波形。
- (b) 将(a)中波形与使用普通升余弦频谱得到的波形进行比较。

第 5 章 信号空间分析

本章讨论关于在加性高斯白噪声(AWGN)信道中进行信号传输的一些基本问题。具体来说,本章涉及了以下内容:

- 有限能量信号的几何表示,为数据传输的研究提供了很好的数学工具。
- 在加性高斯白噪声信道中进行信号检测的最大似然过程。
- 推导与前面章节中讨论过的匹配滤波接收机等价的相干接收机。
- 符号差错概率及用于其近似计算的联合边界。

这里讨论的内容很自然地引出了第 6 章对通带数据传输的研究。

5.1 引言

考虑图 5.1 中所描述的一个最基本的数字通信系统形式。信源每隔 T 秒发送一个符号,该符号属于(记为 m_1, m_2, \dots, m_M 的) M 个符号组成的字符集。例如,考虑两个数字计算机的远程连接。其中一个计算机充当信源,根据输入信号和观测结果来计算数字的输出信号。该计算机输出的结果以一串 0 和 1 表示。该输出结果通过通信信道传输到第二台计算机中。在这个例子中,字符集简单地由两个二进制符号 0 和 1 组成。第二个例子是四进制的 PCM 编码器,其中每个字母由四种可能符号组成:00, 01, 10 和 11。先验概率 p_1, p_2, \dots, p_M 规定了信源的输出情况。在缺少先验信息的情况下,习惯上假设字符集中的 M 个符号等概地出现。因此可将信源发出符号 m_i 的概率表示如下:

$$\begin{aligned} p_i &= P(m_i) \\ &= \frac{1}{M} \quad i = 1, 2, \dots, M \end{aligned} \tag{5.1}$$

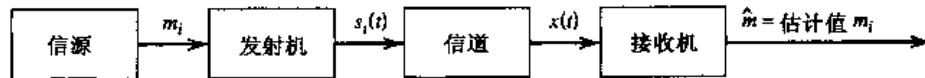


图 5.1 一般数字通信系统的组成框图

发射机获取输出的信源信息 m_i 并将其编码为不同的适于信道传输的信号 $s_i(t)$ 。信号 $s_i(t)$ 占用了分配给符号 m_i 的全部持续时间 T 。最重要的一点是, $s_i(t)$ 是一个实值能量信号(例如,有限能量信号),即

$$E_i = \int_0^T s_i^2(t) dt, \quad i = 1, 2, \dots, M \tag{5.2}$$

假设信道具有下面两个特征:

1. 信道是线性的,其带宽足以使信号 $s_i(t)$ 的传输无失真或失真可忽略。

2. 信道噪声 $w(t)$ 是一个零均值高斯白噪声过程的抽样函数。这样假设是为了使接收机的计算结果易于处理,也是对很多实际通信系统中存在的噪声类型的合理描述。

我们把这样的信道称为 AWGN 信道。因此可将接收信号 $x(t)$ 表示成

$$x(t) = s_i(t) + w(t), \quad \begin{cases} 0 \leq t \leq T \\ i = 1, 2, \dots, M \end{cases} \quad (5.3)$$

信道模型可表示为图 5.2。

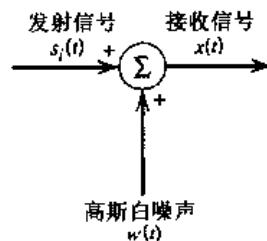


图 5.2 信道的 AWGN 模型

接收机要负责在持续时间 T 内监测接收信号 $x(t)$, 并对传输信号 $s_i(t)$ 或与之等价的符号 m_i 做出最佳估计。然而, 由于信道噪声的影响, 判决过程本质上是在统计意义上进行的, 结果是接收机会偶尔出错。因此需要对接收机进行合理设计, 使得在接收端具有最小的平均符号差错概率。平均符号差错概率的定义为

$$P_e = \sum_{i=1}^M p_i P(\hat{m} \neq m_i | m_i) \quad (5.4)$$

其中, m_i 为传输符号, \hat{m} 为接收机的估计值, $P(\hat{m} \neq m_i | m_i)$ 为已知第 i 个发送符号时的条件差错概率。这样的接收机称为最小差错概率意义下的最佳接收机。

该模型提供了一个设计最佳接收机的基础。对此, 我们将采用已知发射信号集 $\{s_i(t)\}$ 的几何表示。这种表示方法更加易于理解并可在细节上大为简化。下面我们将讨论这种表示方法。

5.2 信号的几何表示

信号几何表示^[1]的本质在于将任何一组 M 个能量信号 $\{s_i(t)\}$ 表示成 N 个正交基本函数的线性组合, 其中 $N \leq M$ 。也就是说, 已知一组实值能量信号 $s_1(t), s_2(t), \dots, s_M(t)$, 其中每一个符号的持续时间为 T 秒, 即

$$s_i(t) = \sum_{j=1}^N s_{ij} \phi_j(t), \quad \begin{cases} 0 \leq t \leq T \\ i = 1, 2, \dots, M \end{cases} \quad (5.5)$$

其中, 展开系数定义为

$$s_{ij} = \int_0^T s_i(t) \phi_j(t) dt, \quad \begin{cases} i = 1, 2, \dots, M \\ j = 1, 2, \dots, N \end{cases} \quad (5.6)$$

实值基本函数 $\phi_1(t), \phi_2(t), \dots, \phi_N(t)$ 是正交的, 也就是说

$$\int_0^T \phi_i(t) \phi_j(t) dt = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (5.7)$$

其中, δ_{ij} 为克罗内克增量(Kronecker delta)。式(5.7)中的第一个条件说明每一个基本函数都归一化为单位能量。第二个条件说明基本函数 $\phi_1(t), \phi_2(t), \dots, \phi_N(t)$ 在时间间隔 $0 \leq t \leq T$ 内是相互正交的。

显然系数集合 $\{s_{ij}\}_{j=1}^N$ 可以看成是一个 N 维矢量, 记为 s_i 。其中, 需要着重指出的一点是, 矢量 s_i 与发射信号 $s_i(t)$ 具有一一对应的关系:

- 将已知矢量 s_i 的 N 个元素(即, $s_{i1}, s_{i2}, \dots, s_{iN}$)作为输入, 就可以用图 5.3(a)所示的方案来产生信号 $s_i(t)$ 。该方案直接来自式(5.5)。该方案由一组 N 个乘法器组成, 其中每个乘法器都有各自的基本函数。乘法器后是一个加法器。这个方案可以看成是一个合成器。

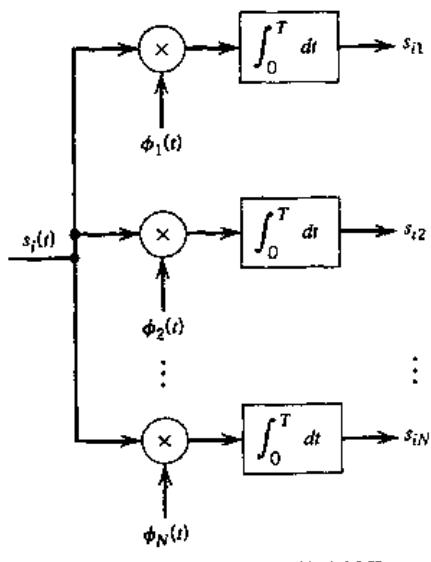
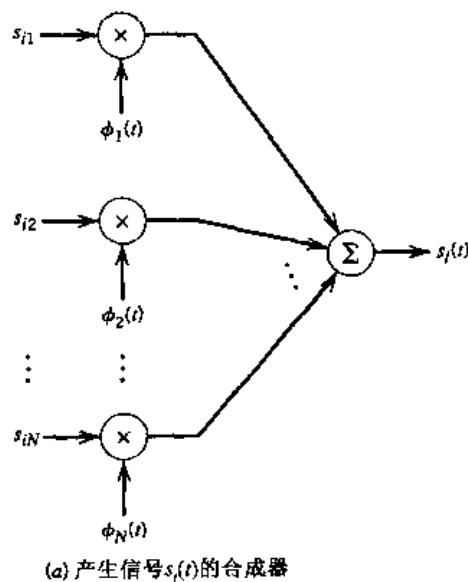


图 5.3 信号合成器与分析器

- 反之, 将已知信号 $s_i(t)$ ($i = 1, 2, \dots, M$) 作为输入, 就可用图 5.3(b) 所示的方案来计算系数 $s_{i1}, s_{i2}, \dots, s_{iN}$ 。该方案直接来自式(5.6)。该方案由一组 N 个共输入的乘积 - 积分器或相关器组成, 每个乘积 - 积分器也都有各自的基本函数。这个方案可以看成是一个分析器。

因此, 可以认为集合 $\{s_i(t)\}$ 中的每个信号完全由其系数矢量决定:

$$\mathbf{s}_i = \begin{bmatrix} s_{i1} \\ s_{i2} \\ \vdots \\ s_{iN} \end{bmatrix}, \quad i = 1, 2, \dots, M \quad (5.8)$$

矢量 \mathbf{s}_i 称为信号矢量。进一步地, 如果将此概念从习惯上的二维和三维欧氏空间拓展到 N 维欧氏空间, 就可以把信号矢量 $\{s_i | i = 1, 2, \dots, M\}$ 的集合形象化为定义在 N 维欧氏空间上的相应 M 个点的集合。该空间具有 N 个相互垂直的轴线, 分别标记为 $\phi_1, \phi_2, \dots, \phi_N$ 。该 N 维欧氏空间就称为信号空间。

这种将一组能量信号从几何上形象化的思想非常重要。这种思想为能量信号的几何表示提供了数学基础, 从而为数字通信系统的噪声分析做好了准备。这种表示如图 5.4 所示。此时, 有三个信号处于二维信号空间中, 即 $N = 2, M = 3$ 。

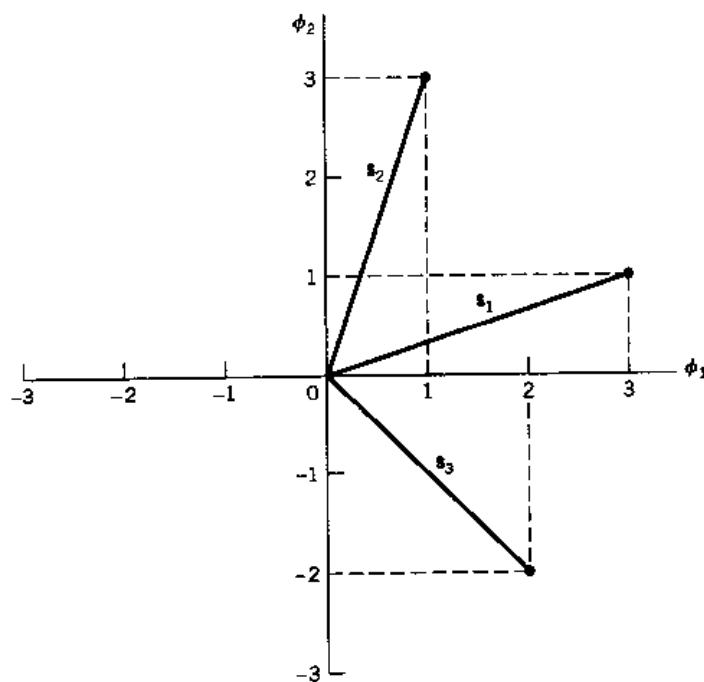


图 5.4 $N = 2, M = 3$ 时信号的几何表示

在一个 N 维欧氏空间中, 我们可以定义矢量的长度及矢量间的角度。习惯上, 标记信号矢量 \mathbf{s}_i 长度(也称为绝对值或模)为 $\|\mathbf{s}_i\|$ 。任何一个信号矢量 \mathbf{s}_i 长度的平方定义为 \mathbf{s}_i 本身的内积或点积, 即

$$\|\mathbf{s}_i\|^2 = \mathbf{s}_i^T \mathbf{s}_i = \sum_{j=1}^N s_{ij}^2, \quad i = 1, 2, \dots, M \quad (5.9)$$

其中, s_{ij} 为矢量 \mathbf{s}_i 的第 j 个元素, 上标 T 表示矩阵转置。

信号能量与其矢量表示之间存在确定的关系。由定义, 持续时间为 T 秒的信号 $s_i(t)$ 的能量为

$$E_i = \int_0^T s_i^2(t) dt \quad (5.10)$$

将式(5.5)代入式(5.10), 得到

$$E_i = \int_0^T \left[\sum_{j=1}^N s_{ij} \phi_j(t) \right] \left[\sum_{k=1}^N s_{ik} \phi_k(t) \right] dt$$

交换求和与积分的顺序, 然后重排各项, 得到

$$E_i = \sum_{j=1}^N \sum_{k=1}^N s_{ij} s_{ik} \int_0^T \phi_j(t) \phi_k(t) dt \quad (5.11)$$

由式(5.7), 由于 $\phi_j(t)$ 为一个正交集合, 式(5.11)可化简为

$$\begin{aligned} E_i &= \sum_{j=1}^N s_{ij}^2 \\ &= \| \mathbf{s}_i \| ^2 \end{aligned} \quad (5.12)$$

因此, 由式(5.9)与式(5.12)可知, 信号 $s_i(t)$ 的能量等于信号矢量 $\mathbf{s}_i(t)$ 长度的平方。

就一对分别由信号矢量 \mathbf{s}_i 与 \mathbf{s}_k 表达的信号 $s_i(t)$ 与 $s_k(t)$ 而言, 也可以证明

$$\int_0^T s_i(t) s_k(t) dt = \mathbf{s}_i^T \mathbf{s}_k \quad (5.13)$$

式(5.13)表明, 信号 $s_i(t)$ 与 $s_k(t)$ 的时域表示在时间间隔 $[0, T]$ 上的内积, 等于它们对应的矢量表示 \mathbf{s}_i 与 \mathbf{s}_k 的内积。注意, $s_i(t)$ 与 $s_k(t)$ 的内积是不会随着基本函数 $\{\phi_j(t)\}_{j=1}^N$ 的选择而改变的, 因为该内积仅由投影到各自基本函数上的信号 $s_i(t)$ 与 $s_k(t)$ 的分量决定。

与信号 $s_i(t)$ 与 $s_k(t)$ 的矢量表示有关的另一个有用的关系式如下:

$$\begin{aligned} \| \mathbf{s}_i - \mathbf{s}_k \| ^2 &= \sum_{j=1}^N (s_{ij} - s_{kj})^2 \\ &= \int_0^T (s_i(t) - s_k(t))^2 dt \end{aligned} \quad (5.14)$$

其中, $\| \mathbf{s}_i - \mathbf{s}_k \|$ 为信号矢量 \mathbf{s}_i 与 \mathbf{s}_k 所代表的点与点间的欧氏距离 d_{ik} 。

为了得到能量信号的几何表示, 需要知道信号矢量 \mathbf{s}_i 与 \mathbf{s}_k 之间的角度 θ_{ik} 。由定义, 角度 θ_{ik} 的余弦值等于这两个矢量的内积除以它们各自模的乘积, 即

$$\cos \theta_{ik} = \frac{\mathbf{s}_i^T \mathbf{s}_k}{\| \mathbf{s}_i \| \| \mathbf{s}_k \|} \quad (5.15)$$

如果内积 $\mathbf{s}_i^T \mathbf{s}_k$ 等于零, 那么称这两个矢量 \mathbf{s}_i 与 \mathbf{s}_k 相互正交或垂直, 此时角度 $\theta_{ik} = 90^\circ$ 。这也是符合直观事实的。

例 5.1 施瓦茨不等式

考虑任一对能量信号 $s_1(t)$ 与 $s_2(t)$ 。有施瓦茨不等式

$$\left(\int_{-\infty}^{\infty} s_1(t)s_2(t)dt \right)^2 \leq \left(\int_{-\infty}^{\infty} s_1^2(t)dt \right) \left(\int_{-\infty}^{\infty} s_2^2(t)dt \right) \quad (5.16)$$

当且仅当 $s_2(t) = cs_1(t)$ 时, 等号成立, 这里的 c 是任意常数。

为了证明这个重要的不等式, 将 $s_1(t)$ 与 $s_2(t)$ 用正交基本函数对 $\phi_1(t)$ 与 $\phi_2(t)$ 表示如下:

$$\begin{aligned} s_1(t) &= s_{11}\phi_1(t) + s_{12}\phi_2(t) \\ s_2(t) &= s_{21}\phi_1(t) + s_{22}\phi_2(t) \end{aligned}$$

其中, $\phi_1(t)$ 与 $\phi_2(t)$ 在整个时间间隔 $(-\infty, \infty)$ 上都满足正交条件

$$\int_{-\infty}^{\infty} \phi_i(t)\phi_j(t)dt = \delta_{ij} = \begin{cases} 1 & j = i \\ 0 & \text{其他} \end{cases}$$

在这个基础上, 可以用图 5.5 中的矢量对来分别表示信号 $s_1(t)$ 与 $s_2(t)$:

$$\begin{aligned} s_1 &= \begin{bmatrix} s_{11} \\ s_{12} \end{bmatrix} \\ s_2 &= \begin{bmatrix} s_{21} \\ s_{22} \end{bmatrix} \end{aligned}$$

从图 5.5 中不难看出, 矢量 s_1 与 s_2 之间的角度 θ 为

$$\begin{aligned} \cos \theta &= \frac{s_1^T s_2}{\|s_1\| \|s_2\|} \\ &= \frac{\int_{-\infty}^{\infty} s_1(t)s_2(t)dt}{\left(\int_{-\infty}^{\infty} s_1^2(t)dt \right)^{1/2} \left(\int_{-\infty}^{\infty} s_2^2(t)dt \right)^{1/2}} \end{aligned} \quad (5.17)$$

其中用到了式(5.15), 式(5.13)和式(5.9)。由于 $|\cos \theta| \leq 1$, 式(5.16)中的施瓦茨不等式可从式(5.17)直接得到。此外, 由式(5.17)的第一行可知, 当且仅当 $s_2 = cs_1$, 即 $s_2(t) = cs_1(t)$ 时, $|\cos \theta| = 1$ 。其中 c 是一个任意常数。

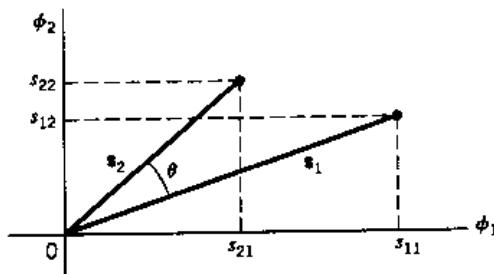


图 5.5 信号 $s_1(t)$ 与 $s_2(t)$ 的矢量表示, 为证明施瓦茨不等式提供了基础

这里对施瓦茨不等式的证明方法, 仅适用于实值信号。当然也可以很容易地将其扩展到复值信号的情况, 这样, 式(5.16)就可写为

$$\left| \int_{-\infty}^{\infty} s_1(t)s_2^*(t)dt \right|^2 \leq \left(\int_{-\infty}^{\infty} |s_1(t)|^2 dt \right) \left(\int_{-\infty}^{\infty} |s_2(t)|^2 dt \right) \quad (5.18)$$

其中, 当且仅当 $s_2(t) = cs_1(t)$ 时, 等号成立。其中 c 为常数, 参见习题 5.9。这是施瓦茨不等式的复数形式, 这种形式在第 4 章中被用于匹配滤波器的推导。

5.2.1 Gram-Schmidt 正交过程

怎样从数学上证明能量信号几何表示的优越性呢？答案就是 Gram-Schmidt 正交过程。其中，需要用到一个完全正交的基本函数集。为了将这个过程用公式进行表示，假设已有一组 M 个能量信号，记为 $s_1(t), s_2(t), \dots, s_M(t)$ 。从集合中任意地选取 $s_1(t)$ 作为起点，则第一个基本函数就定义为

$$\phi_1(t) = \frac{s_1(t)}{\sqrt{E_1}} \quad (5.19)$$

其中， E_1 为信号 $s_1(t)$ 的能量。显然有

$$\begin{aligned} s_1(t) &= \sqrt{E_1}\phi_1(t) \\ &= s_{11}\phi_1(t) \end{aligned} \quad (5.20)$$

其中，系数 $s_{11} = \sqrt{E_1}$ 且 $\phi_1(t)$ 满足单位能量的要求。

接下来要用到信号 $s_2(t)$ 。定义系数 s_{21} 为

$$s_{21} = \int_0^T s_2(t)\phi_1(t)dt \quad (5.21)$$

因而，可引入一个新的中间函数

$$g_2(t) = s_2(t) - s_{21}\phi_1(t) \quad (5.22)$$

由式(5.21)以及基本函数 $\phi_1(t)$ 具有单位能量的性质，可知该中间函数与信号 $\phi_1(t)$ 在时间间隔 $0 \leq t \leq T$ 内正交。下面定义第二个基本函数

$$\phi_2(t) = \frac{g_2(t)}{\sqrt{\int_0^T g_2^2(t)dt}} \quad (5.23)$$

将式(5.22)代入式(5.23)并化简，就得到所希望的结果

$$\phi_2(t) = \frac{s_2(t) - s_{21}\phi_1(t)}{\sqrt{E_2 - s_{21}^2}} \quad (5.24)$$

其中， E_2 为信号 $s_2(t)$ 的能量。由式(5.23)易得

$$\int_0^T \phi_2^2(t)dt = 1$$

由式(5.24)可得

$$\int_0^T \phi_1(t)\phi_2(t)dt = 0$$

即 $\phi_1(t)$ 与 $\phi_2(t)$ 构成了一个满足要求的正交对。

同理可得一般性定义

$$g_i(t) = s_i(t) - \sum_{j=1}^{i-1} s_{ij}\phi_j(t) \quad (5.25)$$

其中,系数 s_{ij} 的定义为

$$s_{ij} = \int_0^T s_i(t) \phi_j(t) dt, \quad j = 1, 2, \dots, i - 1 \quad (5.26)$$

式(5.22)是式(5.25)在 $i = 2$ 时的一个特例。当 $i = 1$ 时,函数 $g_i(t)$ 可化简为 $s_i(t)$ 。

利用 $g_i(t)$,可以定义基本函数集

$$\phi_i(t) = \frac{g_i(t)}{\sqrt{\int_0^T g_i^2(t) dt}}, \quad i = 1, 2, \dots, N \quad (5.27)$$

这就组成了一个正交集合。维数 N 是小于还是等于给定的信号数 M ,取决于信号的两种可能:

- 信号 $s_1(t), s_2(t), \dots, s_M(t)$ 组成了一个线性独立的集合,则 $N = M$;
- 信号 $s_1(t), s_2(t), \dots, s_M(t)$ 组成了一个非线性独立的集合,则 $N < M$,中间函数 $g_i(t)$ 在 $i > N$ 时为零。

值得注意的是,周期信号的传统傅里叶级数展开,是这里所描述类型的一个特例。此外,以在奈奎斯特速率下进行抽样所得样值表示的带限信号,可以看成是这种类型的另一个特例。但需注意,Gram-Schmidt 正交过程与这两个特例相比,有两个重要的不同:

1. 基本函数 $\phi_1(t), \phi_2(t), \dots, \phi_N(t)$ 的形式没有规定。即与周期信号傅里叶级数展开或带限信号的抽样表示不同的是,我们并没有限制 Gram-Schmidt 正交过程为时间的正弦函数或 sinc 函数。
2. 按有限项数展开的信号 $s_i(t)$ 不是一个只有前 N 项具有意义的近似式,而是一个有且仅有 N 项的有意义的准确表达式。

例 5.2 2B1Q 码

2B1Q 码作为 DSL 的北美线路码标准,已在第 4 章中进行了介绍。2B1Q 码代表了表 5.1 中的格雷码字符集所示的一个四进制 PAM 信号。这四种可能信号 $s_1(t), s_2(t), s_3(t)$ 和 $s_4(t)$ 是奈奎斯特脉冲的幅度分级。每个信号代表一个二位组。我们希望得到的是 2B1Q 码的矢量表示。

表 5.1 2B1Q 码的幅度电平

信号		
码元	幅度	格雷码
$s_1(t)$	-3	00
$s_2(t)$	-1	01
$s_3(t)$	+1	11
$s_4(t)$	+3	10

这个例子非常简单,通过观察就能够解决问题。用 $\phi_1(t)$ 表示归一化为单位能量的奈奎斯特脉冲。 $\phi_1(t)$ 是 2B1Q 码矢量表示的唯一基本函数。因此,该码字的信号空间表示如图 5.6 所示。它由四个信号矢量 s_1, s_2, s_3 和 s_4 组成,这四个信号矢量以关于原点对称的方式排列在 ϕ_1 轴上。在该例中, $M = 4, N = 1$ 。

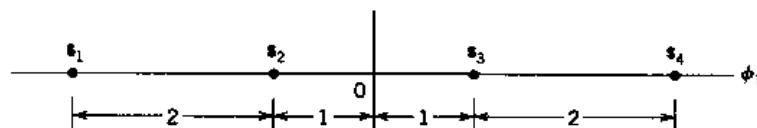


图 5.6 2B1Q 码的信号空间表示

可将图 5.6 描述的 2B1Q 码结果归纳为:一个 M 进制的 PAM 信号的信号空间图,一般来说是一维的,且 M 个信号点均匀地分布在图中唯一的轴上。

5.3 连续 AWGN 信道到矢量信道的变换

假设输入到图 5.3(b)所示的一组 N 个乘积积分器或相关器的不是发射信号 $s_i(t)$,而是按图 5.2 所示理想 AWGN 信道定义的接收信号 $x(t)$ 。即

$$x(t) = s_i(t) + w(t), \quad \begin{cases} 0 \leq t \leq T \\ i = 1, 2, \dots, M \end{cases} \quad (5.28)$$

其中, $w(t)$ 是均值为零、功率谱密度为 $N_0/2$ 的 AWGN 过程 $W(t)$ 的抽样函数。相应地,相关器 j 的输出为随机变量 X_j 的样值,即

$$\begin{aligned} x_j &= \int_0^T x(t) \phi_j(t) dt \\ &= s_{ij} + w_j, \quad j = 1, 2, \dots, N \end{aligned} \quad (5.29)$$

其中第一项 s_{ij} 是由发射信号 $s_i(t)$ 引起的决定性的量。其定义为

$$s_{ij} = \int_0^T s_i(t) \phi_j(t) dt \quad (5.30)$$

第二项 w_j 是由于信道噪声 $w(t)$ 而产生的随机变量 W_j 的样值。其定义为

$$w_j = \int_0^T w(t) \phi_j(t) dt \quad (5.31)$$

接下来,考虑一个新的随机过程 $X'(t)$,其抽样函数 $x'(t)$ 与接收信号 $x(t)$ 有如下关系:

$$x'(t) = x(t) - \sum_{j=1}^N x_j \phi_j(t) \quad (5.32)$$

将式(5.28)和式(5.29)代入式(5.32),并利用式(5.5)的展开式,得

$$\begin{aligned} x'(t) &= s_i(t) + w(t) - \sum_{j=1}^N (s_{ij} + w_j) \phi_j(t) \\ &= w(t) - \sum_{j=1}^N w_j \phi_j(t) \\ &= w'(t) \end{aligned} \quad (5.33)$$

因此,抽样函数 $x'(t)$ 仅取决于信道噪声 $w(t)$ 。基于式(5.32)和式(5.33),可将接收信号表示为

$$\begin{aligned} x(t) &= \sum_{j=1}^N x_j \phi_j(t) + x'(t) \\ &= \sum_{j=1}^N x_j \phi_j(t) + w'(t) \end{aligned} \quad (5.34)$$

因此,可将 $w'(t)$ 看成是一种为使式(5.34)左右相等而必须被包括在公式右边的余项。将式(5.34)中的接收信号 $x(t)$ 的展开式与式(5.5)中的传输信号 $s_i(t)$ 的相应展开式进行对比,是很有意义的。式(5.5)是完全确定的,而式(5.34)却是随机的,这正是我们所期望的。

5.3.1 相关器输出的统计特性

下面要求的是这组 N 个相关器输出的统计特性。用 $X(t)$ 表示一个随机过程,其抽样函数就是接收信号 $x(t)$ 。相应地,用 X_j 表示一个随机变量,其抽样值就是相关器输出 $x_j, j = 1, 2, \dots, N$ 。由图 5.2 中的 AWGN 信道模型可知,随机过程 $X(t)$ 是一个高斯过程。因此,随机变量 X_j 对于所有 j 来说,都是高斯随机变量(见 1.8 节中高斯过程的性质 1)。因此, X_j 的特性就完全地由其均值和方差来决定。均值和方差将在后面的内容中进行确定。

用 W_j 表示第 j 个相关器因响应高斯白噪声分量 $w(t)$ 而产生的样值 w_j 对应的随机变量。由定义,图 5.2 所示 AWGN 信道模型的 $w(t)$ 所对应的噪声过程 $W(t)$ 具有零均值,因此随机变量 W_j 也具有零均值。于是, X_j 的均值仅由 s_{ij} 决定,即

$$\begin{aligned}\mu_{X_j} &= E[X_j] \\ &= E[s_{ij} + W_j] \\ &= s_{ij} + E[W_j] \\ &= s_{ij}\end{aligned}\tag{5.35}$$

为得到 X_j 的方差,我们注意到

$$\begin{aligned}\sigma_{X_j}^2 &= \text{var}[X_j] \\ &= E[(X_j - \mu_{X_j})^2] \\ &= E[W_j^2]\end{aligned}\tag{5.36}$$

上式最后一行来自于式(5.29),只是将其中的 x_j 和 w_j 分别以 X_j 和 W_j 来代替。根据式(5.31),随机变量 W_j 的定义为

$$W_j = \int_0^T W(t)\phi_j(t)dt$$

因此,可将式(5.36)展开为:

$$\begin{aligned}\sigma_{X_j}^2 &= E\left[\int_0^T W(t)\phi_j(t)dt \int_0^T W(u)\phi_j(u)du\right] \\ &= E\left[\int_0^T \int_0^T \phi_j(t)\phi_j(u)W(t)W(u)dtdu\right]\end{aligned}\tag{5.37}$$

交换其中积分与数学期望的顺序

$$\begin{aligned}\sigma_{X_j}^2 &= \int_0^T \int_0^T \phi_j(t)\phi_j(u)E[W(t)W(u)]dtdu \\ &= \int_0^T \int_0^T \phi_j(t)\phi_j(u)R_W(t, u)dtdu\end{aligned}\tag{5.38}$$

其中, $R_W(t, u)$ 为噪声过程 $W(t)$ 的自相关函数。由于该噪声是平稳的,因此 $R_W(t, u)$ 仅由时间差 $t - u$ 决定。另外,由于噪声 $W(t)$ 是具有常数功率谱密度 $N_0/2$ 的白噪声,因此可根据式

(1.95), 将 $R_w(t, u)$ 表示为

$$R_w(t, u) = \frac{N_0}{2} \delta(t - u) \quad (5.39)$$

将式(5.39)代入(5.38), 再利用 $\delta(t)$ 的过滤特性, 得

$$\begin{aligned} \sigma_{X_j}^2 &= \frac{N_0}{2} \int_0^T \int_0^T \phi_j(t) \phi_j(u) \delta(t - u) dt du \\ &= \frac{N_0}{2} \int_0^T \phi_j^2(t) dt \end{aligned} \quad (5.40)$$

另外, 由定义可知, $\phi_j(t)$ 具有单位能量, 因此可得到一个最终的简单结果

$$\sigma_{X_j}^2 = \frac{N_0}{2} \quad \text{对于所有 } j \quad (5.41)$$

这个重要结果表明: 所有由 $X_j, j = 1, 2, \dots, N$ 表示的相关器的输出, 其方差都等于噪声过程 $W(t)$ 的功率谱密度 $N_0/2$ 。

此外, 由于 $\phi_j(t)$ 组成了一个正交集合, 因此 X_j 也互不相关, 即

$$\begin{aligned} \text{cov}[X_j X_k] &= E[(X_j - \mu_{X_j})(X_k - \mu_{X_k})] \\ &= E[(X_j - s_{ij})(X_k - s_{ik})] \\ &= E[W_j W_k] \\ &= E\left[\int_0^T W(t) \phi_j(t) dt \int_0^T W(u) \phi_k(u) du\right] \\ &= \int_0^T \int_0^T \phi_j(t) \phi_k(u) R_w(t, u) dt du \\ &= \frac{N_0}{2} \int_0^T \int_0^T \phi_j(t) \phi_k(u) \delta(t - u) dt du \\ &= \frac{N_0}{2} \int_0^T \phi_j(t) \phi_k(t) dt \\ &= 0, \quad j \neq k \end{aligned} \quad (5.42)$$

由于 X_j 是高斯随机变量, 所以式(5.42)意味着它们也是统计独立的(见 1.8 节中高斯过程的性质 4)。

定义有 N 个随机变量的矢量为

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} \quad (5.43)$$

其元素都是均值为 s_i , 方差为 $N_0/2$ 的独立高斯随机变量。由于矢量 \mathbf{X} 的元素都是统计独立的, 因此, 在发射信号为 $s_i(t)$ 或相应符号 m_i 的情况下, 可将矢量 \mathbf{X} 的条件概率密度函数表达为其各独立元素的条件概率密度函数的乘积, 即

$$f_{\mathbf{X}}(\mathbf{x} | m_i) = \prod_{j=1}^N f_{X_j}(x_j | m_i), \quad i = 1, 2, \dots, M \quad (5.44)$$

其中,矢量 \mathbf{x} 和标量 x_i 分别是随机矢量 \mathbf{X} 和随机变量 X_i 的样值。矢量 \mathbf{x} 称为观察矢量。相应地, x_i 就称为观察元素。满足式(5.44)的任何信道都称为无记忆信道。

因为每个 X_i 都是均值为 s_{ij} 、方差为 $N_0/2$ 的高斯随机变量,所以有

$$f_{X_j}(x_j | m_i) = \frac{1}{\sqrt{\pi N_0}} \exp \left[-\frac{1}{N_0} (x_j - s_{ij})^2 \right], \quad j = 1, 2, \dots, N \quad i = 1, 2, \dots, M \quad (5.45)$$

将式(5.45)代入式(5.44),得

$$f_{\mathbf{X}}(\mathbf{x} | m_i) = (\pi N_0)^{-N/2} \exp \left[-\frac{1}{N_0} \sum_{j=1}^N (x_j - s_{ij})^2 \right], \quad i = 1, 2, \dots, M \quad (5.46)$$

显然,随机矢量 \mathbf{X} 的元素完全决定了和式 $\sum_j X_j \phi_j(t)$ 的特性。此外,和式的样值由式(5.34)的第一项给出。但式(5.34)中还存在仅由信道噪声 $w(t)$ 决定的噪声项 $w'(t)$ 。由 $w(t)$ 表示的噪声过程 $W(t)$ 是零均值高斯过程,因此,由样值 $w'(t)$ 表示的噪声过程 $W'(t)$ 也是零均值高斯过程。此外,任何一个随机变量 $W'(t_k)$,即噪声过程 $W'(t)$ 在时刻 t_k 处的抽样,实际上都是与随机变量集 $\{X_j\}$ 统计独立的。即(见习题 5.10):

$$E[X_j W'(t_k)] = 0, \quad \begin{cases} j = 1, 2, \dots, N \\ 0 \leq t_k \leq T \end{cases} \quad (5.47)$$

由于基于残余噪声过程 $W'(t)$ 的任何一个随机变量都与随机变量集合 $\{X_j\}$ 以及发射信号集合 $\{s_i(t)\}$ 相互独立,因此式(5.47)表明:随机变量 $W'(t_k)$ 与实际发射信号的判决无关。换言之,由接收信号 $x(t)$ 决定的相关器输出是对于判决过程惟一有用的数据。因此,相关器输出为正在研究的问题提供了充分的统计数据。由定义可知,这种充分的统计总结了观察矢量提供的全部相关信息。

下面,我们用不相关定理来对这一节进行总结:

就加性高斯白噪声下的信号检测而言,只有在信号集合 $\{s_i(t)\}_{i=1}^M$ 的基本函数上有投影的噪声才会影响检测问题的充分统计,而其余噪声与此无关。

这个定理的推论是,图 5.2 所示的 AWGN 信道等价于以下述观察矢量描述的一个 N 维矢量信道:

$$\mathbf{x} = \mathbf{s}_i + \mathbf{w}, \quad i = 1, 2, \dots, M \quad (5.48)$$

其中,维数 N 是信号矢量 \mathbf{s}_i 的基本函数个数。式(5.6)和式(5.31)分别定义了信号矢量 \mathbf{s}_i 和噪声矢量 \mathbf{w} 的独立成分。不相关定理及其推论实际上是理解下述信号检测问题的基础。

5.4 似然函数

条件概率密度函数 $f_{\mathbf{X}}(\mathbf{x} | m_i)$, $i = 1, 2, \dots, M$, 是 AWGN 信道的一个非常重要的特性。在已知发送消息符号 m_i 时,由推导过程可知,条件概率密度函数与观察矢量 \mathbf{x} 函数相关。但在接收机中,情况却正好相反,即已知观察矢量 \mathbf{x} ,要估计用于生成 \mathbf{x} 的消息符号 m_i 。为了解决后一个问题,引入似然函数的思想,记为 $L(m_i)$,其定义为

$$L(m_i) = f_{\mathbf{X}}(\mathbf{x} | m_i), \quad i = 1, 2, \dots, M \quad (5.49)$$

尽管 $L(m_i)$ 与 $f_X(x|m_i)$ 恰好有相同的数学形式, 但其含义却并不相同, 认识到这一点非常重要。

实际上, 用记为 $l(m_i)$ 的对数似然函数来进行研究更为方便。 $l(m_i)$ 的定义为

$$l(m_i) = \log L(m_i), \quad i = 1, 2, \dots, M \quad (5.50)$$

对数似然函数与似然函数存在着一一对应的关系, 这是因为:

1. 根据定义, 概率密度函数总是非负的。因此, 似然函数也同样是非负的。
2. 对数似然函数是其自变量的单调递增函数。

将式(5.46)代入式(5.50), 就得到了用于 AWGN 信道的对数似然函数, 即

$$l(m_i) = -\frac{1}{N_0} \sum_{j=1}^N (x_j - s_{ij})^2, \quad i = 1, 2, \dots, M \quad (5.51)$$

其中忽略了常数项 $-(N/2)\log(\pi N_0)$, 因为它与消息符号 m_i 没有任何联系。注意, $s_{ij}, j = 1, 2, \dots, N$ 代表的是消息符号 m_i 的信号矢量 s_i 中的各元素。下面我们着手讨论接收机设计中的基本问题, 其中将用到式(5.51)。

5.5 噪声中信号的相干检测: 最大似然译码

假设在每个持续时间为 T 秒的时隙里, M 个可能的信号 $s_1(t), s_2(t), \dots, s_M(t)$, 以 $1/M$ 的概率等概地发送。就几何信号表示而言, 信号 $s_i(t), i = 1, 2, \dots, M$ 被送入一组有共同输入的相关器, 这组相关器具有 N 个适当的正交基本函数。相关器的输出决定了信号矢量 s_i 。这是因为, 要知道信号矢量 s_i , 只需要知道发射信号 $s_i(t)$ 就可以了, 反之亦然。可以用维数 $N \leq M$ 的欧氏空间中的一个点来代表信号 $s_i(t)$ 。这个点称为发射信号点或消息点。对于发射信号集 $\{s_i(t)\}_{i=1}^M$ 的消息点集合称为信号星座图。

但是, 由于加性噪声 $w(t)$ 的存在, 接收信号 $x(t)$ 的表示形式变得更为复杂。当接收信号 $x(t)$ 用于一组 N 个相关器时, 相关器的输出就决定了观察矢量 x 。由式(5.48)可知, 矢量 x 和信号矢量 s_i 的差别就在于噪声矢量 w 。噪声矢量 w 的方向是完全随机的, 其特性完全由噪声 $w(t)$ 决定。但这个结论的逆命题并不成立。噪声矢量 w 代表了噪声 $w(t)$ 中影响检测过程的那一部分。而噪声的其余部分, 记为 $w'(t)$, 则被相关器滤除了。

基于观察矢量 x , 我们可以在用于表示发射信号的同一个欧氏空间中, 用一个点来表示接收信号 $x(t)$, 这个点称为接收信号点。接收信号点完全随机地分布在消息点周围。从某种意义上说, 接收信号点可能位于以消息点为中心的高斯分布的“云图”以内的任何位置。图 5.7(a)以一个三维信号空间的例子对此进行了说明。对于一个特定的噪声矢量 w , 例如图 5.7(a)所示的随机云图内的一个特殊点, 观察矢量 x 与信号矢量 s_i 之间的关系如图 5.7(b)所示。

可将信号检测问题阐述如下:

已知观察矢量 x , 构造从 x 到发送符号 m_i 的估计值 \hat{x} 之间的映射, 从而使判决过程中的差错概率达到最小。

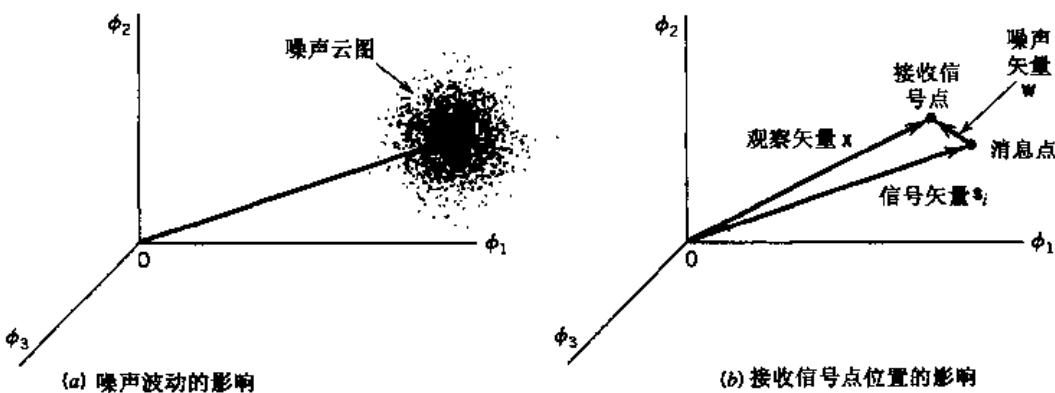


图 5.7 噪声云图与接收信号点

假设已知观察矢量 \mathbf{x} , 判决 $\hat{m} = m_i$ 。判决中的差错概率记为 $P_e(m_i | \mathbf{x})$, 即

$$\begin{aligned} P_e(m_i | \mathbf{x}) &= P(\text{不发送 } m_i | \mathbf{x}) \\ &= 1 - P(\text{发送 } m_i | \mathbf{x}) \end{aligned} \quad (5.52)$$

判决标准是在将每个观察矢量 \mathbf{x} 映射到判决值时, 使差错概率达到最小。基于式(5.52), 可将最优判决准则阐述为:

对所有 $k \neq i$, 如果 $P(\text{发送 } m_i | \mathbf{x}) \geq P(\text{发送 } m_k | \mathbf{x})$, 则有 $\hat{m} = m_i$ (5.53)

其中, $k = 1, 2, \dots, M$ 。该判决准则称为最大后验概率(MAP)准则。

式(5.53)的条件可用发射信号的先验概率和似然函数更为清楚地表达。在式(5.53)中使用贝叶斯准则(Bayes' rule), 并忽略判决过程中的可能联系, 可将 MAP 准则表述为:

如果 $k = i$ 时, $\frac{p_k f_{\mathbf{x}}(\mathbf{x} | m_k)}{f_{\mathbf{x}}(\mathbf{x})}$ 为最大值, 则有 $\hat{m} = m_i$ (5.54)

其中, p_k 为发送符号 m_k 的先验概率, $f_{\mathbf{x}}(\mathbf{x} | m_k)$ 为在发送符号 m_k 的情况下, 随机观察矢量 \mathbf{x} 的条件概率密度函数, $f_{\mathbf{x}}(\mathbf{x})$ 为 \mathbf{x} 的无条件概率密度函数。由式(5.54), 不难发现:

- 分母 $f_{\mathbf{x}}(\mathbf{x})$ 独立于发送符号。
- 当所有源符号等概率发送时, 有先验概率 $p_k = p_i$ 。
- 条件概率密度函数 $f_{\mathbf{x}}(\mathbf{x} | m_k)$ 与对数似然函数 $l(m_k)$ 存在一一对的关系。

因此, 可用 $l(m_k)$ 将式(5.54)中的判决准则简化如下:

如果 $l(m_k)$ 在 $k = i$ 时最大, 则有 $\hat{m} = m_i$ (5.55)

该判决准则称为最大似然准则, 其实现的设备相应地称为最大似然译码器。根据式(5.55), 最大似然译码器对于所有 M 个可能的消息符号都计算出对数似然函数, 经过比较, 最后根据其中最大的一个进行判决。最大似然译码器不同于最大后验译码器, 因为前者假设各消息符号是等概率的。

为便于理解, 下面对最大似然判决准则进行图形解释。用 Z 代表所有可能的观察矢量 \mathbf{x} 构成的 N 维空间, 该空间称为观察空间。由于判决准则中有 $\hat{m} = m_i$, 其中 $i = 1, 2, \dots, M$, 所以整个观察空间 Z 就相应地分为了 M 个判决区域, 分别记为 Z_1, Z_2, \dots, Z_M 。因此, 我们可以将式(5.55)的判决准则描述如下:

当 $k = i$ 时, 如果 $I(m_k)$ 最大, 则观察矢量 \mathbf{x} 位于区域 Z_i 内 (5.56)

不难发现, 除了判决区域 Z_1, Z_2, \dots, Z_M 之间的边界, 这组区域已覆盖了可能的观察矢量 \mathbf{x} 的整个空间。根据惯例, 在所有的临界处都采取随机判决, 即此时接收机只是进行简单的猜测。具体地说, 如果观察矢量 \mathbf{x} 落在了任意两个判决区域 Z_i 和 Z_k 之间的边界上, 在两个可能的判决之间进行的选择, 就可以像抛硬币一样先验地解决。显而易见, 由于在该边界上, 式(5.53)取等号, 所以这样一个事件的结果不会影响最终的差错概率值。

式(5.55)的最大似然判决准则或式(5.56)的几何诠释, 都是在信道噪声 $w(t)$ 为加性的惟一限制下得到的一般形式。接下来, 我们研究 $w(t)$ 既为白噪声又为高斯分布时的情况。

由式(5.51)定义的用于 AWGN 信道的对数似然函数, 不难发现和式

$$\sum_{j=1}^N (x_j - s_{kj})^2$$

在 $k = i$ 时为最小, 此时 $l(m_k)$ 达到最大值。因此, 可将用于 AWGN 信道的最大似然判决准则表示为

当 $k = i$ 时, 如果 $\sum_{j=1}^N (x_j - s_{kj})^2$ 为最小, 则观察矢量 \mathbf{x} 位于区域 Z_i 内 (5.57)

先前我们曾讨论过

$$\sum_{j=1}^N (x_j - s_{kj})^2 = \|\mathbf{x} - \mathbf{s}_k\|^2 \quad (5.58)$$

可将上式与式(5.14)进行比较。上式中, $\|\mathbf{x} - \mathbf{s}_k\|$ 是分别由矢量 \mathbf{x} 和 \mathbf{s}_k 表示的接收信号点与消息点之间的欧氏距离。因此, 可将式(5.57)的判决准则描述如下:

当 $k = i$ 时, 如果欧氏距离 $\|\mathbf{x} - \mathbf{s}_k\|$ 为最小, 则观察矢量 \mathbf{x} 位于区域 Z_i 内 (5.59)

式(5.59)表明, 最大似然判决准则就是仅选择离接收信号点最近的消息点, 这与直观事实也是相符的。

实际上, 式(5.59)的判决准则中可以避免平方计算, 因为

$$\sum_{j=1}^N (x_j - s_{kj})^2 = \sum_{j=1}^N x_j^2 - 2 \sum_{j=1}^N x_j s_{kj} + \sum_{j=1}^N s_{kj}^2 \quad (5.60)$$

该展开式中的第一个和项是独立于指数 k 的, 因此可忽略。第二个和项是观察矢量 \mathbf{x} 和信号矢量 \mathbf{s}_k 的内积。第三个和项为发射信号 $s_k(t)$ 的能量。由此可得等价于式(5.59)的如下判决准则:

$k = i$ 时, 如果 $\sum_{j=1}^N x_j s_{kj} - \frac{1}{2} E_k$ 为最大, 则观察矢量 \mathbf{x} 位于区域 Z_i 内 (5.61)

其中, E_k 为发射信号 $s_k(t)$ 的能量:

$$E_k = \sum_{j=1}^N s_{kj}^2 \quad (5.62)$$

由式(5.61)可得, 对于 AWGN 信道, 判决区域应是 N 维观察空间 Z 内的区域, 该区域由线性 ($N - 1$) 维的超平面边界界定。图 5.8 为信号数 $M = 4$, 维数 $N = 2$ 的判决区域的例子, 其中假设信号以等能量 E 等概率地传输。

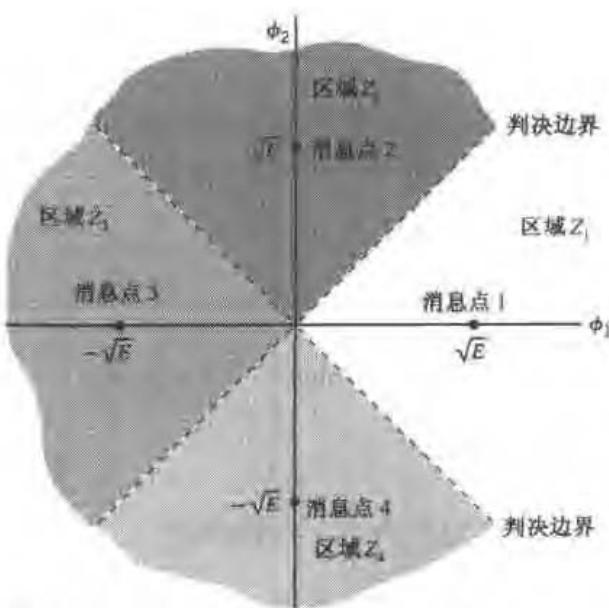


图 5.8 当 $N=2$ 及 $M=4$ 时, 将观察空间划分为判决区域的图解。这里假设 M 个发送符号等概率

5.6 相干接收机

由前面的内容不难发现, 对于 AWGN 信道, 当发射信号 $s_1(t), s_2(t), \dots, s_M(t)$ 等概率出现时, 最佳接收机由两个子系统组成, 如图 5.9 所示:

1. 图 5.9(a) 为接收机的检测器部分, 由一组 M 个乘积 - 积分器或相关器组成。其输入为一组本地生成的相应的相关参考信号或正交基本函数 $\phi_1(t), \phi_2(t), \dots, \phi_N(t)$ 。这组相关器处理接收信号 $x(t), 0 \leq t \leq T$, 用以产生观察矢量 \mathbf{x} 。
2. 图 5.9(b) 为接收机的信号传输译码器部分, 以最大似然译码器的形式来实现。即按最小化平均符号差错概率的方式处理观察矢量 \mathbf{x} , 以产生发送符号 $m_i, i = 1, 2, \dots$, 且 M 的估计值为 \hat{m} 。根据式(5.61), 观察矢量 \mathbf{x} 的 N 个元素首先与每个 M 信号矢量 $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_M$ 的 N 个对应元素相乘, 将乘积在累加器中相继求和, 以形成相应的内积 $\{\mathbf{x}^T \mathbf{s}_k | k = 1, 2, \dots, M\}$ 。由于发射信号的能量可能彼此不等, 因此需要对内积加以纠正。最后, 选择出内积中最大的那个, 这样就完成了对发送消息的恰当判决。

一般来说, 图 5.9 所示的最佳接收机称为相干接收机。

5.6.1 相干接收机与匹配滤波接收机的等价

图 5.9(a) 所示的检测器包括一组相关器。我们可以用一组匹配滤波器来代替这组相关器, 从而得到另一种检测器。匹配滤波器及其性质在 4.2 节中曾经讨论过。为了证明相关器与匹配滤波器的等价性, 现考虑一个冲激响应为 $b_j(t)$ 的线性时不变滤波器。以接收信号 $x(t)$ 作为该滤波器的输入, 滤波器输出 $y_j(t)$ 以卷积积分定义为:

$$y_j(t) = \int_{-\infty}^{\infty} x(\tau) b_j(t - \tau) d\tau \quad (5.63)$$

由 4.2 节提出的匹配滤波器的定义可知, 匹配于输入信号 $\phi_j(t)$ 的线性时不变滤波器的冲激响应 $b_j(t)$ 应是输入信号 $\phi_j(t)$ 的一个时间反转和延迟。假设有

$$b_j(t) = \phi_j(T - t) \quad (5.64)$$

那么, 滤波器的最终输出为

$$y_j(t) = \int_{-\infty}^{\infty} x(\tau) \phi_j(T - t + \tau) d\tau \quad (5.65)$$

对该输出在时刻 $t = T$ 进行抽样, 得到

$$y_j(T) = \int_{-\infty}^{\infty} x(\tau) \phi_j(\tau) d\tau$$

由定义, $\phi_j(t)$ 在时间间隔 $0 \leq t \leq T$ 之外为零, 因此 $y_j(T)$ 实际上是图 5.9(a) 所示的接收信号 $x(t)$ 在第 j 个相关器的输出 x_j , 即

$$y_j(T) = \int_0^T x(\tau) \phi_j(\tau) d\tau \quad (5.66)$$

因此, 最佳接收机的检测器部分也可以用一组匹配滤波器来实现, 如图 5.10 所示。需要着重指出的是, 图 5.9(a) 中每个相关器的输出只有在时刻 $t = T$ 上进行抽样时, 才等价于图 5.10 中的相应匹配滤波器的输出。

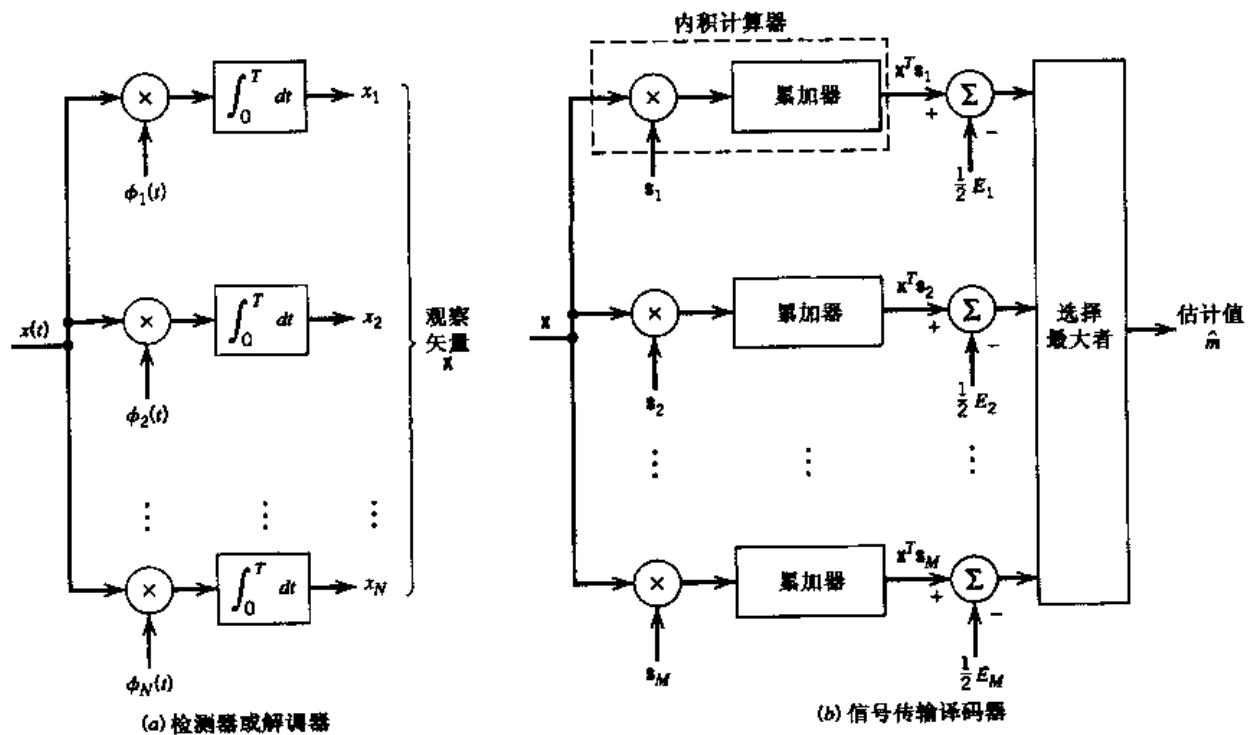


图 5.9 相干接收机

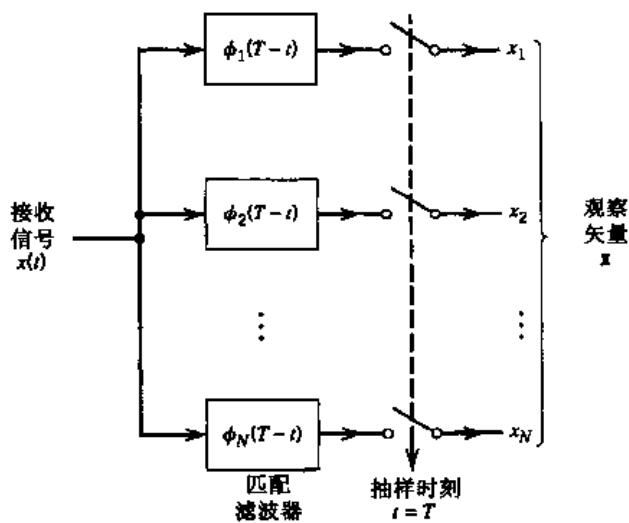


图 5.10 匹配滤波接收机的检测部分；信号传输译码器部分见图 5.9(b)

5.7 差错概率

为了得到图 5.9 中描述的相干接收机的统计特性，需要评价其噪声性能。为此，假设观察空间 \$Z\$ 按照最大似然判决准则，分为 \$M\$ 个区域 \$\{Z_i\}_{i=1}^M\$，同时假设发送符号为 \$m_i\$（或信号矢量为 \$\mathbf{s}_i\$），观察矢量为 \$\mathbf{x}\$。则当 \$\mathbf{x}\$ 表示的接收信号点没有落入消息点 \$\mathbf{s}_i\$ 对应的区域 \$Z_i\$ 时，表明判断出错。平均符号差错概率 \$P_e\$ 为

$$\begin{aligned} P_e &= \sum_{i=1}^M p_i P(\mathbf{x} \text{ 不在区域 } Z_i \text{ 内} | \text{发送 } m_i) \\ &= \frac{1}{M} \sum_{i=1}^M P(\mathbf{x} \text{ 不在区域 } Z_i \text{ 内} | \text{发送 } m_i) \\ &= 1 - \frac{1}{M} \sum_{i=1}^M P(\mathbf{x} \text{ 在 } Z_i \text{ 内} | \text{发送 } m_i) \end{aligned} \quad (5.67)$$

其中，采用标准符号来标记事件的概率及条件概率。由于 \$\mathbf{x}\$ 是随机矢量 \$\mathbf{X}\$ 的样值，因此可用似然函数将式(5.67)写为

$$P_e = 1 - \frac{1}{M} \sum_{i=1}^M \int_{Z_i} f_{\mathbf{X}}(\mathbf{x} | m_i) d\mathbf{x} \quad (5.68)$$

对一个 \$N\$ 维观察矢量，式(5.68)中的积分项同样也是 \$N\$ 维的。

5.7.1 差错概率的旋转不变性和平移不变性

在对加性高斯白噪声下的信号进行最大似然检测时，将观察空间 \$Z\$ 划分成区域 \$Z_1, Z_2, \dots, Z_M\$ 的具体方法，由所研究的信号星座图惟一决定。因此，信号星座图的方向相对于坐标轴和信号空间原点的变化，并不影响式(5.68)中定义的符号差错概率 \$P_e\$。这是因为：

1. 在最大似然检测中，符号差错概率 \$P_e\$ 仅仅由星座图中消息点之间的相对欧氏距离决定。

2. 加性高斯白噪声在信号空间中的所有方向上都是球形对称的。

首先, 考虑 P_s 关于旋转的不变性。对星座图中所有消息点旋转的效果, 等价于对所有 i , 将 N 维信号矢量 s_i 乘上一个 N 行 N 列的正交矩阵 \mathbf{Q} 。矩阵 \mathbf{Q} 满足条件

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{I} \quad (5.69)$$

其中, \mathbf{I} 为单位矩阵, 其对角线元素均为 1, 非对角线元素均为 0。根据式(5.69), 实正交矩阵的逆矩阵等于其转置形式。因此, 信号矢量 s_i 可由其旋转形式代替:

$$s_{i,\text{rotate}} = \mathbf{Q}s_i, \quad i = 1, 2, \dots, M \quad (5.70)$$

相应地, N 行 1 列的噪声矢量 w 也由其旋转形式代替:

$$w_{\text{rotate}} = \mathbf{Q}w \quad (5.71)$$

但噪声矢量的统计特性并不会受这种旋转的影响, 其原因如下:

- 由第 1 章可知, 高斯随机变量的线性组合仍是高斯型的。既然假设噪声矢量 w 是高斯型的, 那么旋转噪声矢量 w_{rotate} 也是高斯型的。
- 由于噪声矢量 w 具有零均值, 因此, 旋转噪声矢量 w_{rotate} 也具有零均值, 即

$$\begin{aligned} E[w_{\text{rotate}}] &= E[\mathbf{Q}w] \\ &= \mathbf{Q}E[w] \\ &= 0 \end{aligned} \quad (5.72)$$

- 噪声矢量 w 的协方差矩阵等于 $(N_0/2)\mathbf{I}$, 其中, $N_0/2$ 是加性高斯白噪声 $w(t)$ 的功率谱密度, 即

$$E[ww^T] = \frac{N_0}{2} \mathbf{I} \quad (5.73)$$

因此, 旋转噪声矢量 w_{rotate} 的协方差矩阵为

$$\begin{aligned} E[w_{\text{rotate}}w_{\text{rotate}}^T] &= E[\mathbf{Q}w(\mathbf{Q}w)^T] \\ &= E[\mathbf{Q}ww^T\mathbf{Q}^T] \\ &= \mathbf{Q}E[ww^T]\mathbf{Q}^T \\ &= \frac{N_0}{2} \mathbf{Q}\mathbf{Q}^T \\ &= \frac{N_0}{2} \mathbf{I} \end{aligned} \quad (5.74)$$

其中最后两行用到了式(5.73)和式(5.69)。

根据上述观察结果, 将观察矢量表示在旋转后的星座图中, 即

$$\mathbf{x}_{\text{rotate}} = \mathbf{Q}s_i + w, \quad i = 1, 2, \dots, M \quad (5.75)$$

由式(5.59)可知, 最大似然检测的判决准则是基于观察矢量 $\mathbf{x}_{\text{rotate}}$ 和旋转后的信号矢量 $s_{i,\text{rotate}} = \mathbf{Q}s_i$ 之间的欧氏距离的。比较式(5.75)与式(5.48), 易见

$$\|\mathbf{x}_{\text{rotate}} - s_{i,\text{rotate}}\| = \|\mathbf{x} - s_i\| \quad \text{对于所有 } i \quad (5.76)$$

因此, 可将旋转不变性原理阐述如下:

如果信号星座图经由一个正交变换进行旋转, 即

$$\mathbf{s}_{i,\text{rotate}} = \mathbf{Q}\mathbf{s}_i, \quad i = 1, 2, \dots, M$$

其中, \mathbf{Q} 为一个正交变换矩阵, 那么, AWGN 信道的最大似然信号检测所引起的符号差错概率 P_e 完全不变。

下面用一个例子来阐明该原理。图 5.11(b)所示的信号星座图除了被旋转了 45 度之外, 与图 5.11(a)是完全一样的。尽管这两个星座图看上去确实不同, 但旋转不变性原理告诉我们二者的 P_e 是相同的。

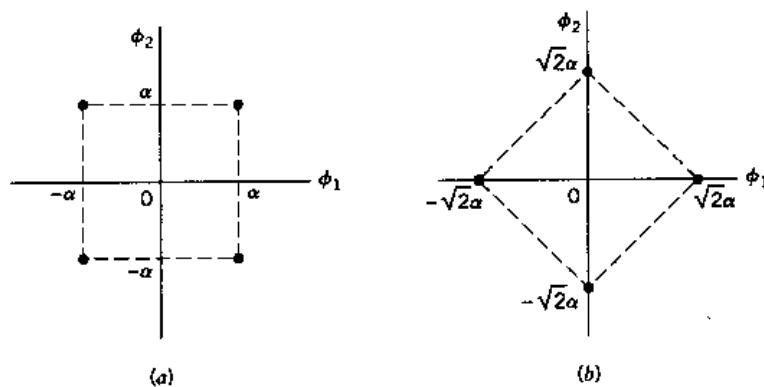


图 5.11 一对用以阐明旋转不变性原理的信号星座图

接下来考虑平移不变性。假设一个信号星座图中所有消息点都平移了一个常矢量 \mathbf{a} , 即

$$\mathbf{s}_{i,\text{translate}} = \mathbf{s}_i + \mathbf{a}, \quad i = 1, 2, \dots, M \quad (5.77)$$

相应地, 观察矢量也平移了同样的矢量值, 即

$$\mathbf{x}_{\text{translate}} = \mathbf{x} + \mathbf{a} \quad (5.78)$$

由式(5.77)和式(5.78)可知, 平移量 \mathbf{a} 在平移后的信号矢量 \mathbf{s}_i 和平移后的观察矢量 \mathbf{x} 中是相同的。因此可立即推导出

$$\|\mathbf{x}_{\text{translate}} - \mathbf{s}_{i,\text{translate}}\| = \|\mathbf{x} - \mathbf{s}_i\| \quad \text{对于所有 } i \quad (5.79)$$

从而将平移不变性原理表述如下:

如果一个信号星座图平移了一个常矢量, 那么, 由 AWGN 信道中的最大似然信号检测所引起的符号差错概率 P_e 完全不变。

作为一个例子, 考虑图 5.12 所示的两个信号星座图, 它们对应于一对不同的 4 电平 PAM 信号。图 5.12(b)所示的星座图除了沿 ϕ_1 轴右移了 $3\alpha/2$ 之外, 与图 5.12(a)是完全一样的。由平移不变性原理可知, 这两个星座图的 P_e 是一样的。

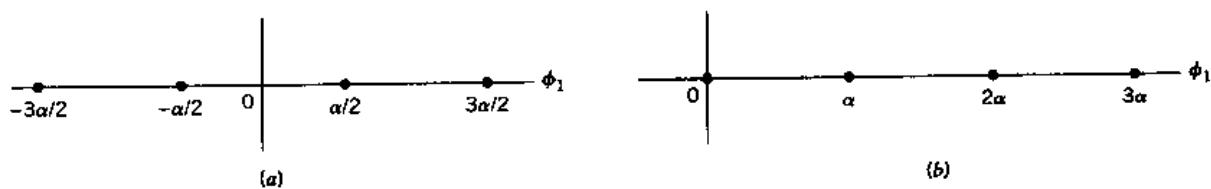


图 5.12 一对用以阐明平移不变性原理的信号星座图

5.7.2 最小能量信号

平移不变性原理的一个有效应用就是对已知的信号星座图进行平移,从而使平均能量达到最小。为了说明这个问题,考虑一组分别由信号矢量 s_1, s_2, \dots, s_M 代表的符号 m_1, m_2, \dots, m_M 。平移矢量值 \mathbf{a} 后的信号星座图的平均能量为

$$\mathcal{E}_{\text{translate}} = \sum_{i=1}^M \|s_i - \mathbf{a}\|^2 p_i \quad (5.80)$$

其中, p_i 为信源发送符号 m_i 的概率。将 s_i 与 \mathbf{a} 之间的欧氏距离的平方展开为

$$\|s_i - \mathbf{a}\|^2 = \|s_i\|^2 - 2\mathbf{a}^T s_i + \|\mathbf{a}\|^2$$

因此,可将式(5.80)展开为

$$\begin{aligned} \mathcal{E}_{\text{translate}} &= \sum_{i=1}^M \|s_i\|^2 p_i - 2 \sum_{i=1}^M \mathbf{a}^T s_i p_i + \|\mathbf{a}\|^2 \sum_{i=1}^M p_i \\ &= \mathcal{E} - 2\mathbf{a}^T E[\mathbf{s}] + \|\mathbf{a}\|^2 \end{aligned} \quad (5.81)$$

其中, \mathcal{E} 为原始信号星座图的平均能量,且

$$E[\mathbf{s}] = \sum_{i=1}^M s_i p_i \quad (5.82)$$

将式(5.81)对矢量 \mathbf{a} 求导,并令结果为零,易得最小平移量为

$$\mathbf{a}_{\min} = E[\mathbf{s}] \quad (5.83)$$

用这种方式平移的信号星座图的最小平均能量为

$$\mathcal{E}_{\text{translate},\min} = \mathcal{E} - \|\mathbf{a}_{\min}\|^2 \quad (5.84)$$

下面阐述寻找最小能量平移的过程:

已知信号星座集 $\{s_i\}_{i=1}^M$,通过将已知星座图中的每个信号矢量减去常矢量 $E[\mathbf{s}]$,可得相应的具有最小平均能量的信号星座图,其中 $E[\mathbf{s}]$ 由式(5.82)定义。

信号传输所需要的能量(或功率)是一种主要的资源,最小能量平移提供了一种使所需能量达到最小的基本方法。例如,图 5.12(a)所示的信号星座图就具有最小平均能量,而图 5.12(b)所示的信号星座图则不是这样。

5.7.3 差错概率的联合边界

对 AWGN 信道而言,平均符号差错概率 P_e 的公式从概念上讲是很直接的。通过将式(5.46)代入式(5.68),可简单地将 P_e 写为积分形式。但遗憾的是,除了几个简单(但重要)的情况之外,一般来讲,该积分的计算都是很难实现的。为了克服这种计算上的困难,可以引入边界值,这通常已经足以预测出保持规定差错率所需的信噪比(大约一个 dB 以内)。 P_e 定义中的积分近似,通过简化积分或简化积分域就能够实现。接下来就通过简化积分域来得到一种简单但实用的称为联合边界的上界,作为 AWGN 信道中一组 M 个等概率信号(符号)的平均符号差错概率的近似值。

用 A_{ik} , $(i, k) = 1, 2, \dots, M$ 代表发送符号 m_i (矢量 s_i) 时, 观察矢量 x 距离信号矢量 s_k 比距离 s_i 近这一事件。当发送符号 m_i 时, 条件符号差错概率 $P_e(m_i)$ 等于联合事件 $A_{i1}, A_{i2}, \dots, A_{i,i-1}, A_{i,i+1}, \dots, A_{i,M}$ 的概率。由概率论可知, 有限联合事件的概率不超过各组成事件的概率之和。因此, 有

$$P_e(m_i) \leq \sum_{\substack{k=1 \\ k \neq i}}^M P(A_{ik}), \quad i = 1, 2, \dots, M \quad (5.85)$$

图 5.13 中以 $M = 4$ 为例对以上关系进行了说明。图 5.13(a)中有四个信号点及相关的判决区域, 其中, s_i 代表发送符号。图 5.13(b)给出了三种组成的信号空间。其中, 发送消息点 s_i 和另一个消息点在三种信号空间中都保持不变。由图 5.13(a)可知, 条件符号差错概率 $P_e(m_i)$ 等于观察矢量 x 落入二维信号空间图阴影区域的概率。易见, 该概率小于 x 落入图 5.13(b)所示的三种信号空间阴影区域的这三个独立事件的概率之和。

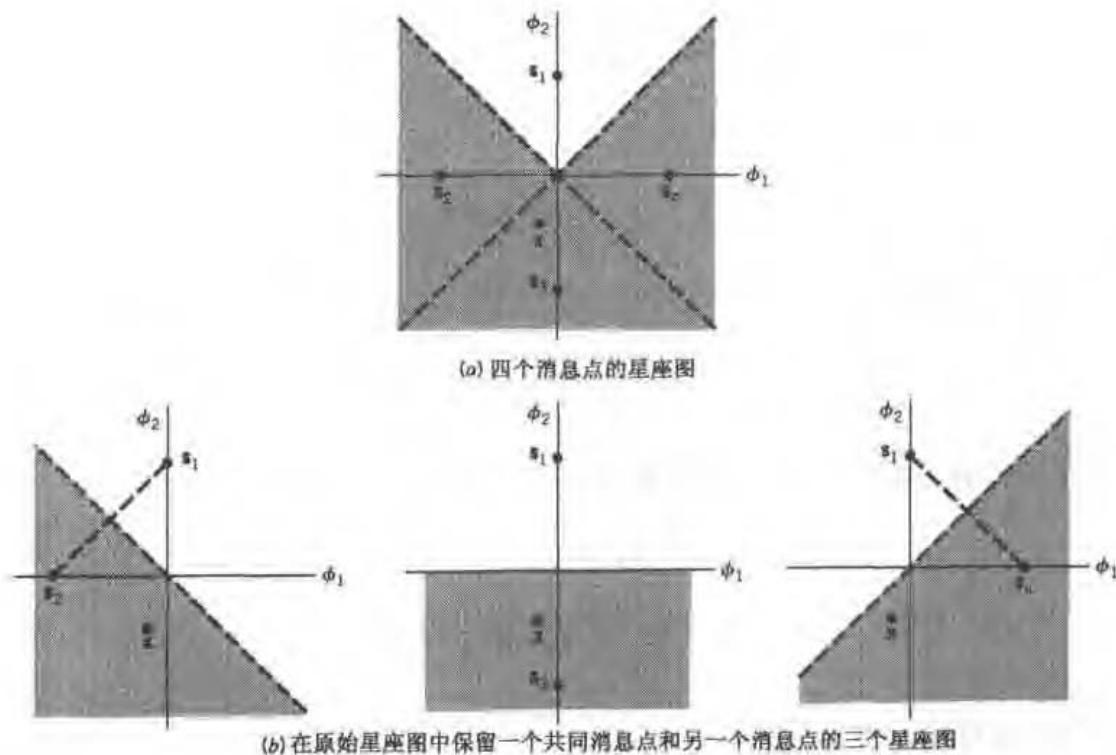


图 5.13 联合边界示意图

特别要注意的是, 一般来说, 概率 $P(A_{ik})$ 不同于概率 $P(\hat{m} = m_k | m_i)$ 。后者是当发射信号为矢量 s_i (或 m_i) 时, 观察矢量 x 更靠近信号矢量 s_k 的概率。相反, 概率 $P(A_{ik})$ 仅由两个信号矢量 s_i 和 s_k 决定。为了强调这种差别, 用 $P_2(s_i, s_k)$ 代替 $P(A_{ik})$, 将式(5.85)写为

$$P_e(m_i) \leq \sum_{\substack{k=1 \\ k \neq i}}^M P_2(s_i, s_k), \quad i = 1, 2, \dots, M \quad (5.86)$$

如果一个数字传输系统只使用一对信号 s_i 和 s_k , 那么, $P_2(s_i, s_k)$ 就是接收机将 s_k 误判为 s_i 的概率, 因此将概率 $P_2(s_i, s_k)$ 称为两两差错概率。

接下来,考虑一个简化的数字通信系统,其中采用两个分别由矢量 s_i 和 s_k 代表的等概率消息。高斯白噪声沿任一组正交轴的分布相同,因此可以暂时地选择该组中的第一条轴作为通过点 s_i 和 s_k 的轴。对图 5.13(b)中的三个例子来说,相应的判决边界由垂直于 s_i 和 s_k 连线的平分线所代表。相应的,当发送符号为 m_i (矢量 s_i)时,如果观察矢量 x 落入了 s_k 所处的那一侧,则表明判决出错。此事件的概率由下式给出:

$$\begin{aligned} P_2(s_i, s_k) &= P(\text{当发射信号为 } s_i \text{ 时, } x \text{ 离 } s_k \text{ 的距离比离 } s_i \text{ 更近}) \\ &= \int_{d_{ik}/2}^{\infty} \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{\nu^2}{N_0}\right) d\nu \end{aligned} \quad (5.87)$$

其中, d_{ik} 为 s_i 和 s_k 之间的欧氏距离,即

$$d_{ik} = \|s_i - s_k\| \quad (5.88)$$

由互补误差函数的定义,有

$$\operatorname{erfc}(u) = \frac{2}{\sqrt{\pi}} \int_u^{\infty} \exp(-z^2) dz$$

令 z 等于 $\nu/\sqrt{N_0}$, 可将式(5.87)写为

$$P_2(s_i, s_k) = \frac{1}{2} \operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right) \quad (5.89)$$

将式(5.89)代入式(5.86),得

$$P_e(m_i) \leq \frac{1}{2} \sum_{\substack{k=1 \\ k \neq i}}^M \operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right), \quad i = 1, 2, \dots, M \quad (5.90)$$

因此,所有 M 个符号的平均符号差错概率,有如下的上界:

$$\begin{aligned} P_e &= \sum_{i=1}^M p_i P_e(m_i) \\ &\leq \frac{1}{2} \sum_{i=1}^M \sum_{\substack{k=1 \\ k \neq i}}^M p_i \operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right) \end{aligned} \quad (5.91)$$

其中, p_i 为发送符号 m_i 的概率。

式(5.91)有两种对通带数据传输来说很有用的特殊形式(见第 6 章):

1. 假设信号星座图关于原点呈圆对称。则条件差错概率 $P_e(m_i)$ 对所有 i 是一样的,此时,式(5.91)简化为

$$P_e \leq \frac{1}{2} \sum_{\substack{k=1 \\ k \neq i}}^M \operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right) \quad \text{对于所有 } i \quad (5.92)$$

2. 定义信号星座图的最小距离 d_{\min} 为星座图中任意两个发射信号点的最小欧氏距离,即

$$d_{\min} = \min_{k \neq i} d_{ik} \quad \text{对于所有 } i \text{ 和 } k \quad (5.93)$$

由于互补误差函数 $\operatorname{erfc}(u)$ 随自变量 u 单调递减,因此有

$$\operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right) \leq \operatorname{erfc}\left(\frac{d_{\min}}{2\sqrt{N_0}}\right) \quad \text{对于所有 } i \text{ 和 } k \quad (5.94)$$

一般地,可将式(5.91)中平均符号差错概率的上界简化为

$$P_e \leq \frac{(M-1)}{2} \operatorname{erfc}\left(\frac{d_{\min}}{2\sqrt{N_0}}\right) \quad (5.95)$$

互补误差函数本身有如下边界:^[3]

$$\operatorname{erfc}\left(\frac{d_{\min}}{2\sqrt{N_0}}\right) \leq \frac{1}{\sqrt{\pi}} \exp\left(-\frac{d_{\min}^2}{4N_0}\right) \quad (5.96)$$

相应地,可进一步将式(5.95)中 P_e 的联合边界简化为

$$P_e \leq \frac{(M-1)}{2\sqrt{\pi}} \exp\left(-\frac{d_{\min}^2}{4N_0}\right) \quad (5.97)$$

式(5.97)表明:对一个给定的AWCN信道而言,平均符号差错概率 P_e 随最小平方距离 d_{\min}^2 呈指数级递减。

5.7.4 误比特率与符号差错概率

到目前为止,用于评价一个数字通带传输系统噪声性能的惟一性能参数就是平均符号差错概率。当发送长度 $m = \log_2 M$ 的消息(例如字母数字符号)时,该性能参数就是一个很自然的选择。但当要求传输二进制数据如数字计算机数据时,采用另一种称为误比特率的性能参数将更有意义。尽管一般来说这两个性能参数之间并没有特定的关系,但我们仍然可以从接下来所讨论的两种情况中得到二者的一些关系。

情况 1

这里假设可以进行一种从二进制到 M 进制的映射,使 M 维调制方案中任意一对临近符号相对应的二进制 M 位组,只有一位不同。通过使用格雷码就能够满足这个映射约束。当符号差错概率 P_e 小到可以接受的程度时,将一个符号误判为“最靠近”的两个符号中的任一个符号的差错概率要远远大于其他差错类型的概率。此外,如果已知一个符号出错,根据前面提到的映射约束,则最可能的比特差错数就是 1。因为每个符号有 $\log_2 M$ 个比特,所以平均符号差错概率与误比特率有下述关系:

$$\begin{aligned} P_e &= P\left(\bigcup_{i=1}^{\log_2 M} \{\text{第 } i \text{ 个比特出错}\}\right) \\ &\leq \sum_{i=1}^{\log_2 M} P(\text{第 } i \text{ 个比特出错}) \\ &= \log_2 M \cdot (\text{BER}) \end{aligned} \quad (5.98)$$

同时

$$P_e \geq P(\text{第 } i \text{ 个比特出错}) = \text{BER} \quad (5.99)$$

因此,误比特率有如下边界:

$$\frac{P_e}{\log_2 M} \leq \text{BER} \leq P_e \quad (5.100)$$

情况 2

设 $M = 2^K$, 其中 K 为整数。假设所有符号差错等概率产生, 其概率为

$$\frac{P_e}{M - 1} = \frac{P_e}{2^K - 1}$$

其中, P_e 为平均符号差错概率。那么, 一个符号中第 i 个比特出错的概率是多少呢? 符号中一个特定比特的变化对应有 2^{K-1} 种符号差错情况。另外 2^{K-1} 种符号差错情况下, 该比特不变。因此, 误比特率为

$$\text{BER} = \left(\frac{2^{K-1}}{2^K - 1} \right) P_e \quad (5.101)$$

等价地

$$\text{BER} = \left(\frac{M/2}{M - 1} \right) P_e \quad (5.102)$$

注意, 当 M 取较大的数时, 误比特率接近极限值 $P_e/2$, 这说明比特差错并不是独立的, 因为

$$P(\text{第 } i \text{ 个和第 } j \text{ 个比特同时出错}) = \frac{2^{K-2}}{2^K - 1} P_e \neq (\text{BER})^2$$

5.8 总结与讨论

本章的主要目的是明确给出在加性高斯白噪声下, 分析和设计一个数字通信接收机的过程。最大似然检测的方法中, 通过观察信道输出来判定哪个特定的发射信号最有可能是含噪信号, 由此推导出的最大似然检测器(接收机)的方法称为信号空间分析。该方法的基本思想就是用一个 N 维矢量来表示一组发射信号中的每一个, 其中 N 为发射信号的惟一几何表示所需要的正交基本函数的个数。这样形成的一组信号矢量就在一个 N 维信号空间中定义了一个信号星座图。

对一个已知的信号星座图而言, 由 AWGN 信道中最大似然信号检测所产生的(平均)符号差错概率 P_e , 具有旋转不变性和平移不变性。然而, 除了一些简单(但重要)的情况外, P_e 的数值计算很难实现。为了克服这个困难, 习惯上采用边界值的方法, 以一种直截了当的方式进行计算。本章描述了直接从信号空间图得到的联合边界。联合边界基于一个与直观事实相符的思想: 符号差错概率 P_e 是由最靠近发射信号的相邻信号所决定的。当信噪比较高时, 使用联合边界获得的结果通常是相当准确的。

信号空间分析的内容及其相关问题的解决办法, 为我们在第 6 章对通带数据传输系统的研究, 提供了非常有用的工具。

注释与参考

[1] 信号的几何表达式首先由 V.A.Kotel'nikov 于 1947 年在抗噪声最优化理论(Dover 出

版,1960)一书中提出。该书原本是1947年1月于莫斯科 Molotov Energy 学院的学术会议前提出的一篇博士论文翻译稿。该书的第二部分介绍了信号的几何表达式。该方法后来由 Wozencraft 和 Jacobs(1965)在一本经典书籍中进行了完善。信号空间分析也由 Cioffi(1998), Anderson(1999)和 Proakis(1995)等人进行过讨论。

- [2] 5.7节给出了平均符号差错概率的联合边界。有关联合边界的经典论述见 Wozencraft 和 Jacobs(1965)。要得到更紧凑的边界,可参见 Viterbi 和 Omura(1979, pp. 58–59)。
- [3] 在第4章中,我们使用了关于互补误差函数的上界

$$\operatorname{erfc}(u) < \frac{\exp(-u^2)}{\sqrt{\pi u}}$$

对一个较大的正数 u 来说,互补误差函数的第二个界限可通过忽略上式所示上界中的乘积因子 $1/u$ 来获得,即

$$\operatorname{erfc}(u) < \frac{\exp(-u^2)}{\sqrt{\pi}}$$

这就是在等式(5.97)中使用的第二个上界。

习题

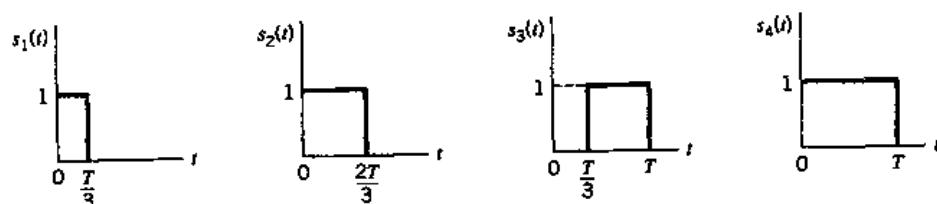
信号表示

- 5.1 3.7节中描述了用于PCM的线路码。试参考3.7节,给出下列线路码的信号星座图:
- (a) 单极性NRZ码
 - (b) 双极性NRZ码
 - (c) 单极性RZ码
 - (d) 曼彻斯特码
- 5.2 一个8电平PAM信号定义为

$$s_i(t) = A_i \operatorname{rect}\left(\frac{t}{T} - \frac{1}{2}\right)$$

其中, $A_i = \pm 1, \pm 3, \pm 5, \pm 7$ 。试给出 $\{s_i(t)\}_{i=1}^8$ 的信号星座图。

- 5.3 图P5.3给出了四个信号 $s_1(t), s_2(t), s_3(t)$ 和 $s_4(t)$ 的波形。
- (a) 使用Gram-Schmidt正交过程,找出这组信号的正交基本函数。
 - (b) 试给出相应的信号空间图。



图P5.3

- 5.4 (a) 使用 Gram-Schmidt 正交过程, 找出一组用以表示图 5.4 中的三个信号 $s_1(t)$, $s_2(t)$ 和 $s_3(t)$ 的正交基本函数。

(b) 用(a)部分找到的基本函数集表示这三个信号。

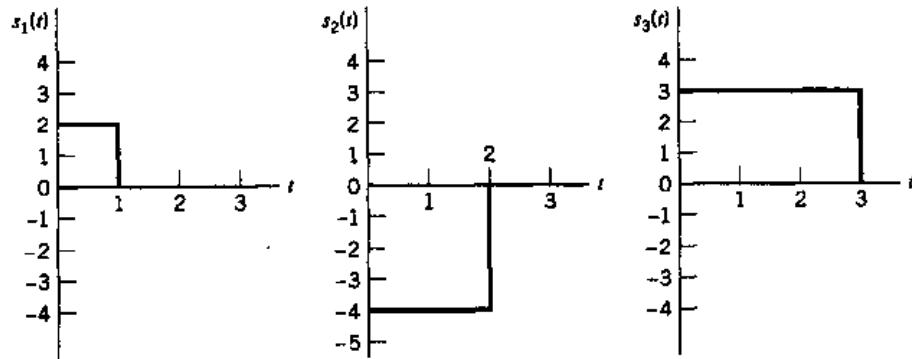


图 P5.4

- 5.5 正交信号集的特性就是其中任意一对信号的内积为零。图 P5.5 为一对满足该条件的信号 $s_1(t)$ 和 $s_2(t)$ 。试给出 $s_1(t)$ 和 $s_2(t)$ 的信号星座图。

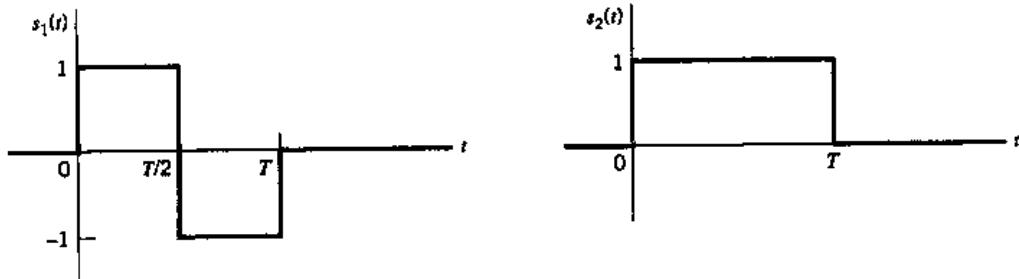


图 P5.5

- 5.6 一个信源发送记为 $\{m_i\}_{i=1}^M$ 的一组符号。现有两种备选调制方法:PDM 和 PPM。在 PDM 中, 第 i 个符号由一个幅度为 1 且宽度为 $(i/M)T$ 的脉冲表示。在 PPM 中, 第 i 个符号由一个幅度为 1 且宽度固定的窄脉冲表示, 并在时刻 $t = (i/M)T$ 发送。证明二者中只有 PPM 能够在时间间隔 $0 \leq t \leq T$ 上产生正交信号集。

- 5.7 通过提取集合中每个信号的反信号, 可将一个包含 M 个信号的正交信号集拓展为包含 $2M$ 个信号的双正交信号集。

(a) 正交信号到双正交信号的拓展能够保持信号空间的维数不变。为什么?

(b) 给出 P5.5 所示的一对正交信号的双正交信号空间图。

- 5.8 (a) 一对信号 $s_i(t)$ 和 $s_k(t)$ 具有相同的持续时间 T 。证明这对信号的内积由下式给出:

$$\int_0^T s_i(t)s_k(t)dt = \mathbf{s}_i^T \mathbf{s}_k$$

其中, \mathbf{s}_i 和 \mathbf{s}_k 分别是 $s_i(t)$ 和 $s_k(t)$ 的矢量表示。

(b) 作为(a)部分的推论, 证明

$$\int_0^T (s_i(t) - s_k(t))^2 dt = \|\mathbf{s}_i - \mathbf{s}_k\|^2$$

5.9 考虑一对分别由下式表达的复值信号 $s_1(t)$ 和 $s_2(t)$

$$\begin{aligned}s_1(t) &= a_{11}\phi_1(t) + a_{12}\phi_2(t), \quad -\infty < t < \infty \\ s_2(t) &= a_{21}\phi_1(t) + a_{22}\phi_2(t), \quad -\infty < t < \infty\end{aligned}$$

其中, 基本函数 $\phi_1(t)$ 和 $\phi_2(t)$ 都为实值, 但系数 a_{11}, a_{12}, a_{21} 和 a_{22} 为复数。证明施瓦茨不等式的复数形式

$$\left| \int_{-\infty}^{\infty} s_1(t)s_2^*(t)dt \right|^2 \leq \int_{-\infty}^{\infty} |s_1(t)|^2 dt \int_{-\infty}^{\infty} |s_2(t)|^2 dt$$

其中, * 表示复共轭。上式的等号何时成立?

随机过程

5.10 考虑一个随机过程 $X(t)$ 的展开式

$$X(t) = \sum_{i=1}^N X_i \phi_i(t) + W'(t), \quad 0 \leq t \leq T$$

其中, $W'(t)$ 是残余噪声项。 $\{\phi_i(t)\}_{i=1}^N$ 组成了在时间间隔 $0 \leq t \leq T$ 上的正交集合, X_i 的定义为

$$X_i = \int_0^T X(t)\phi_i(t)dt$$

用 $W'(t_k)$ 代表在时刻 $t = t_k$ 上观察 $W'(t)$ 所得的随机变量, 证明

$$E[X_i W'(t_k)] = 0, \quad \begin{cases} i = 1, 2, \dots, N \\ 0 \leq t_k \leq T \end{cases}$$

5.11 考虑加性高斯白噪声下, 如下正弦信号的最优检测:

$$s(t) = \sin\left(\frac{8\pi t}{T}\right), \quad 0 \leq t \leq T$$

- (a) 假设无噪声输入, 求相关器的输出。
- (b) 假设匹配滤波器有一个偶然的时延 T , 求相应的滤波器输出。
- (c) 证明这两个输出只有在时刻 $t = T$ 上才相同。

差错概率

5.12 图 P5.12 给出了一对在观察时间 $0 \leq t \leq 3T$ 上相互正交的信号 $s_1(t)$ 和 $s_2(t)$ 。接收信号定义为

$$x(t) = s_k(t) + w(t), \quad \begin{cases} 0 \leq t \leq 3T \\ k = 1, 2 \end{cases}$$

其中, $w(t)$ 为零均值且功率谱密度为 $N_0/2$ 的高斯白噪声。

- (a) 设计一个针对信号 $s_1(t)$ 或 $s_2(t)$ 判决的接收机, 其中假设这两个信号等概率。
- (b) $E/N_0 = 4$ 时, 计算该接收机的平均符号差错概率。其中 E 为信号能量。

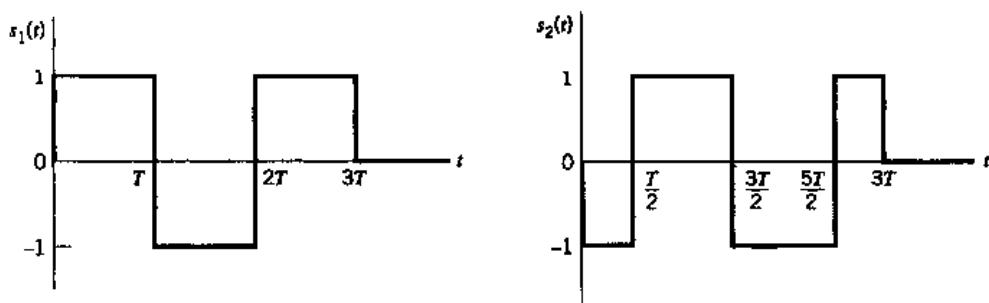


图 P5.12

- 5.13 在曼彻斯特码中,二进制符号 1 由图 P5.13 所示的双脉冲 $s(t)$ 表示,二进制符号 0 由该脉冲的反信号表示。推导在 AWGN 信道上对这种形式的信号进行最大似然检测所得到的差错概率的公式。

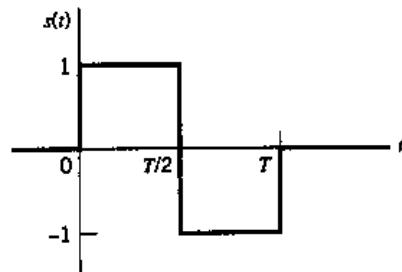


图 P5.13

- 5.14 在贝叶斯实验中,必须在两个可能性假设 H_0 和 H_1 中进行选择。为此,需要最小化冒险值 R , R 的定义为

$$\begin{aligned} R = & C_{00} p_0 P(H_0 | H_0) \\ & + C_{10} p_0 P(H_1 | H_0) \\ & + C_{11} p_1 P(H_1 | H_1) \\ & + C_{01} p_1 P(H_0 | H_1) \end{aligned}$$

其中, C_{00} , C_{10} , C_{11} 和 C_{01} 表示为四种可能结果分配的值:第一个下标代表当前选择的假设,第二个代表确定为正确的假设。假设 $C_{10} > C_{00}$ 且 $C_{01} > C_{11}$ 。 p_0 和 p_1 分别表示假设 H_0 和 H_1 的先验概率。

(a) 已知观察矢量 \mathbf{x} ,证明为了最小化冒险值 R ,将观察空间分块,可得到似然比测试:

如果 $\Lambda(\mathbf{x}) < \lambda$, 则假设 H_0
如果 $\Lambda(\mathbf{x}) > \lambda$, 则假设 H_1

其中, $\Lambda(\mathbf{x})$ 为似然率:

$$\Lambda(\mathbf{x}) = \frac{f_{\mathbf{x}}(\mathbf{x} | H_1)}{f_{\mathbf{x}}(\mathbf{x} | H_0)}$$

其中, λ 为定义如下的测试门限:

$$\lambda = \frac{p_0(C_{10} - C_{00})}{p_1(C_{01} - C_{11})}$$

(b) 使贝叶斯准则降为最小差错概率准则的代价值为多少?

旋转及平移不变性原理

- 5.15 在习题 5.1 的四种线路码中, 分辨出具有最小平均能量的和没有最小平均能量的线路码。将你的答案与 3.7 节中对这些线路码的观察结果进行比较。
- 5.16 考虑图 5.11 中的两个星座图。求将图 5.11(a) 所示的星座图转换为图 5.11(b) 所示的星座图的正交矩阵 \mathbf{Q} 。
- 5.17 (a) 图 P5.17 中的两个信号星座图具有相同的平均符号差错概率。证明该结论的正确性。
 (b) 这两个星座图中哪一个具有最小平均能量? 证明你的答案。
 可以假设与图 P5.17 中消息点相关的符号是等概率的。

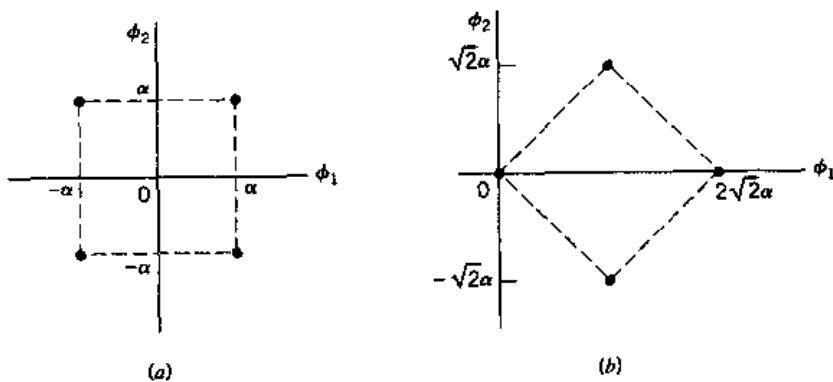


图 P5.17

- 5.18 单工(超正交)信号是等概率、高度相关且具有最负相关性的信号, 这种最负的相关性可以由一组 M 个正交信号获得。即, 该集合中任意一对信号的相关系数定义为

$$\rho_{ij} = \begin{cases} 1 & i = j \\ \frac{-1}{M-1} & i \neq j \end{cases}$$

构建超正交信号的一种方法是由一组 M 个能量均为 E 的正交信号开始, 应用最小能量平移。

考虑一组三个等概率的符号, 其信号星座图由一个等边三角形的顶点组成。证明这三个符号组成了一个超正交码。

差错概率的边界

- 5.19 本问题探究图 P5.19 所示的一对对称信号在功率谱密度为 $N_0/2$ 的加性高斯白噪声条件下的差错概率 P_e 的近似值。 P_e 的确切公式为

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right)$$

该公式将在 6.3 节中进行推导。

- (a) 用注释 3 中给出的互补误差函数的两个上界,推导 P_e 对应的近似式。
 (b) 比较(a)部分得到的 P_e 的近似式与 $E_b/N_0 = 9$ 时的确切公式。对于 P_e 的精确计算,可以使用误差函数表 A6.6。

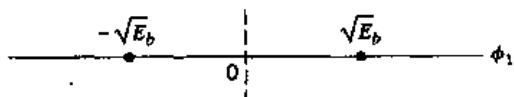


图 P5.19

5.20 考虑一个具有关于原点对称性的特殊的信号星座图。假设星座图中对应于符号 m_1, m_2, \dots, m_M 的 M 个消息点是等概率的。利用式(5.94)中给出的互补误差函数的上界,证明:星座图的平均符号差错概率有如下边界:

$$P_e \leq \frac{M_{\min}}{2\sqrt{\pi}} \exp\left(-\min_{\substack{i, k \\ i \neq k}} \left(\frac{d_{ik}^2}{4N_0}\right)\right)$$

其中, d_{ik} 为消息点 i 和 k 之间的欧氏距离, M_{\min} 是对每个 m_i 而言达到最小欧氏距离的发射信号数。

第 6 章 通带数据传输

这一章的介绍建立在第 5 章对信号空间分析所得结果的基础之上,讨论了数字数据通过线性或者非线性带通信道传输的问题。与模拟通信相似,这种模式的数据传输需要用数据流对正弦载波进行调制。

本章将具体讨论以下内容:

- 数字调制的各种不同方式,例如相移键控、正交幅度调制、频移键控,以及它们各自的变化。
- 对附加了高斯白噪声的调制信号的相干检测,这需要将接收机与发射机进行同步,使它们具备相同的载波相位和时序。
- 对附加了高斯白噪声的调制信号的非相干检测,这种检测忽略了接收信号中的相位信息。
- 使数字数据能够通过公众电话交换网发送和接收的调制解调器。
- 一些复杂的调制技术,例如无载波幅度/相位调制和离散多音。当数据通过存在严重信号间干扰的宽带信道进行传输时,比较适合于采用这些技术。
- 将接收机与发射机同步的技术。

6.1 引言

在第 4 章研究的基带脉冲传输中,数据流以离散 PAM 信号的形式直接通过低通信道传输。数字通带传输则不同,输入的数据流必须先被调制到载波(通常是正弦载波)上,而该载波恰能满足所选择的带通传输信道的频率限制。本章就研究通带数据传输。

用于通带数据传输的通信信道,可以是微波无线信道、卫星信道等。通带数据传输还有一些其他应用,例如为数字用户环路设计的通带线路码,以及用于广播的正交频分复用(OFDM)技术。总之,调制过程就是以某种方式根据输入数据来改变(键控)正弦载波的幅度、频率或者相位,从而使数据传输成为可能。共有三种基本的信号调制方式,分别是幅移键控(ASK)、频移键控(FSK)和相移键控(PSK)。这三种调制方式可以分别看成 AM, FM 和 PM 的特例。

图 6.1 以输入为二进制数据流的情况为例阐明了上述三种调制方式。值得注意的几点是:

- 对于连续波调制的情况,仅仅看波形是很难区分 PM 信号和 FM 信号的。但对于 PSK 和 FSK 信号,却不是这样。
- PSK 与 FSK 信号都有恒定的包络,这一点与 ASK 信号是不同的。

后一个特性使 PSK 和 FSK 信号在幅度为非线性的情况下仍能不受影响,这种情况在微波信道和卫星信道中较为常见。因此,在实用中,PSK 和 FSK 信号比 ASK 信号更适合于非线性信道中的通带数据传输。

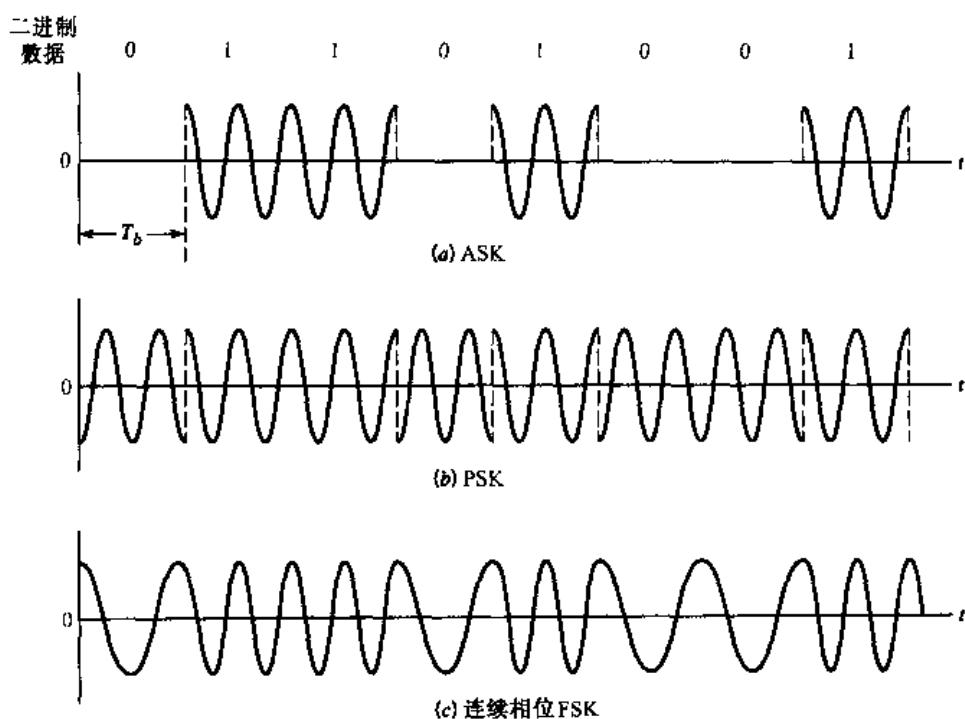


图 6.1 二进制信号波形图的三种基本形式

6.1.1 数字调制技术系列^[1]

根据接收机是否配备有相位恢复电路,可将数字调制技术分为相干和非相干两种。相位恢复电路能够保证为接收机提供逻辑载波的振荡器与为发射机提供输入数据流的振荡器之间的同步(包括频率和相位)。

正如第4章中所述,在每个信号持续时间间隔 T 内,可从一个 M 元信号图谱中任选其中一个可能的信号 $s_1(t), s_2(t), \dots, s_M(t)$ 进行发送。在绝大多数的应用中,可能的信号数 M 符合关系式 $M = 2^n$,其中 n 为整数。信号持续时间 $T = nT_b$,其中 T_b 为比特持续时间。在通带数据传输中,通过 M 个不连续的步骤,可改变正弦载波的幅度、相位或频率从而产生传输信号,并由此得到了 M 进制 ASK、 M 进制 PSK 和 M 进制 FSK 数字调制方法。另一种产生 M 进制信号的方法是将各种不同的调制方法结合起来。例如,将载波的幅度和相位都进行不连续的改变,可产生 M 进制幅度相位联合键控(APK)。 M 进制正交幅度调制(QAM)就是这种联合调制技术的一个特例。QAM 具有一些吸引人的特性。 M 进制 ASK 是 M 进制 QAM 的一个特例。

当要求以功率为代价换取带宽时, M 进制信号比二进制信号更适于在带通信道传输数字信息。实际上,几乎没有什么通信信道能够恰好满足以二进制信号传输信息的带宽需求。当信道带宽小于所需值时,采用 M 进制信号可获得最大的效率。为了证明 M 进制信号节省带宽的能力,假设传输一个比特持续时间为 T_b 的二进制序列信息。如果采用二进制 PSK,则需要的带宽和比特持续时间 T_b 成反比。但如果把 n 比特的信息合成一个块,并采用 M 进制 PSK,其中令 $M = 2^n$,信号持续时间 $T = nT_b$,则需要的带宽与 $1/nT_b$ 成正比。这表明 M 进制 PSK 的传输带宽仅为二进制 PSK 的传输带宽的 $1/n$,其中 $n = \log_2 M$ 。

M 进制 PSK 和 M 进制 QAM 都是线性调制的例子,但二者有一个重要的不同: M 进制 PSK

有恒定的包络,而 M 进制 QAM 的载波幅度是变化的。因此, M 进制 PSK 适用于在非线性通带信道中传输数字数据,而 M 进制 QAM 则适用于线性信道。

M 进制 PSK、 M 进制 QAM 和 M 进制 FSK 一般用于相干系统中。当然,在无法维持载波相位同步的非相干系统中,也可采用 ASK 和 FSK。但不存在“非相干 PSK”的说法,因为术语非相干意味着失去载波相位信息。但可以采用“伪 PSK”技术,即差分相移键控(DPSK)技术,DPSK 可以看成是一种非相干的 PSK。在实际应用中, M 进制 FSK 和 M 进制 DPSK 是非相干系统中常用的数字调制形式。

6.1.2 差错概率

通带数据传输系统的一个主要目标就是接收机的优化设计,从而在存在加性高斯白噪声的情况下使平均符号差错概率最小化。为了达到这一目标,本章的很多内容都建立在第 5 章的信号空间分析基础上。需要特别指出的是,在研究每个系统时,我们都首先将信号星座图表示为公式,并构造出对 AWGN 信道中的信号进行最大似然检测的对应区域。上述公式是估算平均符号差错概率 P_e 的基础。根据所采用数字调制方式的不同,估算 P_e 的方法也相应地为下列两种方法之一:

- 对于一些简单的调制方式,如相干二进制 PSK 和相干二进制 FSK,已经推导出 P_e 的计算公式。
- 对于一些较为复杂的调制方式,如相干 M 进制 PSK 和相干 M 进制 FSK,为了推导出 P_e 的近似计算公式,可采用联合边界的方法。

6.1.3 功率谱

为了全面地理解各种不同数字调制方式的优点,还需对调制信号的功率谱进行研究。这对于信道带宽和复用系统的信道间干扰问题,都具有极其重要的意义。

调制信号 $s(t)$ 可由其同相和正交两个分量表示:

$$\begin{aligned} s(t) &= s_I(t) \cos(2\pi f_c t) - s_Q(t) \sin(2\pi f_c t) \\ &= \operatorname{Re}[\tilde{s}(t) \exp(j2\pi f_c t)] \end{aligned} \quad (6.1)$$

其中 $\operatorname{Re}[\cdot]$ 表示取方括弧中表达式的实部。同时,有

$$\tilde{s}(t) = s_I(t) + j s_Q(t) \quad (6.2)$$

和

$$\exp(j2\pi f_c t) = \cos(2\pi f_c t) + j \sin(2\pi f_c t) \quad (6.3)$$

其中信号 $\tilde{s}(t)$ 为调制(带通)信号 $s(t)$ 的复包络(如基带信号)。 $s_I(t), s_Q(t)$ 以及 $\tilde{s}(t)$ 都是低通信号,均由带通信号 $s(t)$ 和载波频率 f_c 唯一确定。值得注意的是,信号 $s(t)$ 带宽的一半必须小于载波频率 f_c 。

用 $S_B(f)$ 表示复包络 $\tilde{s}(t)$ 的功率谱密度,即基带功率谱密度,则原始带通信号 $s(t)$ 的功率谱密度 $S_s(f)$ 就是 $S_B(f)$ 的频移,外加一个比例因子,其表达式为

$$S_s(f) = \frac{1}{4} [S_B(f - f_c) + S_B(f + f_c)] \quad (6.4)$$

这对于计算基带功率谱密度 $S_B(f)$ 来讲已经足够。由于 $s(t)$ 是低通信号, 所以 $S_B(f)$ 的计算必然比 $S_s(f)$ 更为简单(参见例 1.7)。

6.1.4 带宽效率

本书一再强调信道带宽和发射功率是两种主要的“通信资源”。为有效利用这两种资源, 必须找到有效的频谱方案。频谱有效调制的首要目标就是最大化带宽效率, 即数据速率与有效信道带宽之比。另一个目标则是以最小的平均信号功率代价实现上述带宽效率, 或者说, 在存在加性高斯白噪声干扰的信道里, 以最小的平均信噪比代价实现上述带宽效率。

用 R_b 表示数据速率, 用 B 表示有效带宽, 则带宽效率 ρ 的计算公式如下:

$$\rho = \frac{R_b}{B} \text{ b/(s·Hz)} \quad (6.5)$$

第 4 章曾经讲到带宽效率是两个相互独立的因素作用的结果: 可能采用的多级编码和频谱整形。采用多级编码时, 信息是以多个比特组成的块而不是单个比特的形式发送的。对于有效的频谱整形, 通过使用脉冲整形滤波器平滑掉传输波形中的突变尖峰, 可以减小所需信道带宽。因此, 这两个因素在确定通带数据传输系统的带宽效率时都起着重要的作用。

6.2 通带传输模型

从功能上考虑, 可以构造出如图 6.2 所示的通带数据传输系统模型。首先, 假设存在一个每 T 秒发送一个符号的消息源, 其发送符号属于一个包含有 M 个符号的字符集合。这些发送符号以 m_1, m_2, \dots, m_M 表示, 其先验概率 $P(m_1), P(m_2), \dots, P(m_M)$ 确定了信源的输出。当符号集合中的 M 个符号等概地出现时, 有

$$p_i = P(m_i) = \frac{1}{M} \quad \text{对于所有 } i \quad (6.6)$$

信源的 M 元输出被送到信号传输编码器, 相应产生由 N 个实数构成的一组矢量 s_i , 符号源的 M 个符号各自对应一个矢量 s_i , 其维数 N 小于或等于 M 。调制器以矢量 s_i 为输入, 在每个 T 秒的持续时间段中构造一个独立的信号 $s_i(t)$ 。信号 $s_i(t)$ 为能量信号, 其能量为

$$E_i = \int_0^T s_i^2(t) dt, \quad i = 1, 2, \dots, M \quad (6.7)$$

注意, 这里的 $s_i(t)$ 是实数值。每 T 秒内就有一个这样的信号被发送, 被发送的特定信号某种程度上取决于输入消息, 也可能与前面时隙中发送的信号有关。当采用正弦载波时, 调制器通过载波的幅度、频率或者相位的步进变化来区分各个信号。(有时也采用将幅度改变和相位改变相结合, 或者将幅度改变和频率改变相结合的复合调制方式。)

在图 6.2 所示的功能模型中, 假定连接发射机和接收机的通带通信信道具备两个特性:

1. 信道是线性的, 其带宽足够宽, 能够无畸变或可忽略畸变地传输调制信号 $s_i(t)$ 。
2. 信道噪声 $w(t)$ 是具有零均值且功率谱密度为 $N_0/2$ 的高斯白噪声的样本函数。

这里的假设和第 5 章进行信号空间分析时提出的假设基本一致。

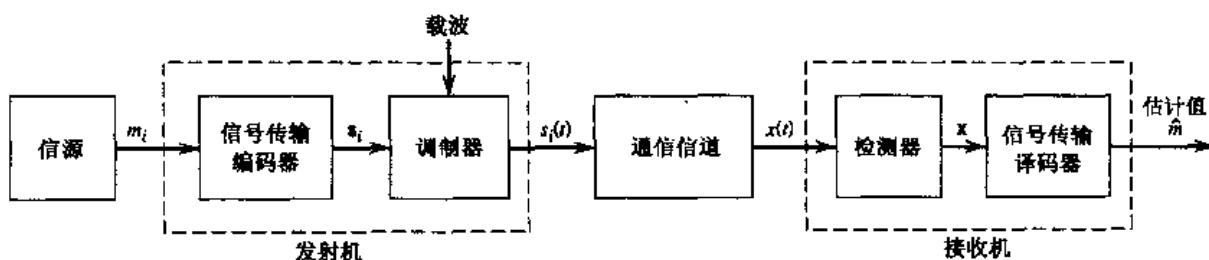


图 6.2 通带传输系统的功能模型

接收机由检测器和信号传输译码器级联而成, 具有以下两个功能:

1. 实现和发射机相反的操作。
2. 在估算发射信号 m_i 的近似值 \hat{m} 时, 使信道噪声的影响最小化。

6.3 相干 PSK

第5章介绍了对加性高斯白噪声下的信号进行相干检测的背景知识, 下面我们开始研究特定的通带数据传输系统。本节重点研究相干PSK, 即分别介绍二进制PSK、QPSK及其变型, 最后介绍M进制PSK。

6.3.1 二进制PSK

在相干二进制PSK系统中, 信号对 $s_1(t)$ 和 $s_2(t)$ 通常代表二进制符号1和0。 $s_1(t)$ 和 $s_2(t)$ 分别由下式决定:

$$s_1(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t) \quad (6.8)$$

$$s_2(t) = \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t + \pi) = -\sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t) \quad (6.9)$$

其中, $0 \leq t \leq T_b$, E_b 为每比特的传输信号能量。为了确保每个传输比特含有整数个周期的载波波形, 要求载波频率 f_c 等于 n_c/T_b , 其中, n_c 为固定的整数。一对正弦波形的惟一区别就在于它们之间有 180° 的相对相移, 如式(6.8)和式(6.9)所示。这一对正弦波形称为反相信号。

由式(6.8)和(6.9)可知, 对于二进制PSK, 单位能量的计算只有一个基本函数, 即

$$\phi_1(t) = \sqrt{\frac{2}{T_b}} \cos(2\pi f_c t), \quad 0 \leq t < T_b \quad (6.10)$$

因此可用 $\phi_1(t)$ 将发射信号 $s_1(t)$ 和 $s_2(t)$ 表示如下:

$$s_1(t) = \sqrt{E_b} \phi_1(t), \quad 0 \leq t < T_b \quad (6.11)$$

和

$$s_2(t) = -\sqrt{E_b} \phi_1(t), \quad 0 \leq t < T_b \quad (6.12)$$

因此, 相干二进制PSK系统的特点就是存在一个一维信号空间($N=1$), 其信号星座图由

两个信息点组成($M = 2$)。这两个信息点的坐标分别是

$$\begin{aligned}s_{11} &= \int_0^{T_b} s_1(t) \phi_1(t) dt \\ &= +\sqrt{E_b}\end{aligned}\quad (6.13)$$

和

$$\begin{aligned}s_{21} &= \int_0^{T_b} s_2(t) \phi_1(t) dt \\ &= -\sqrt{E_b}\end{aligned}\quad (6.14)$$

对应于 $s_1(t)$ 的信息点位于 $s_{11} = +\sqrt{E_b}$, 对应于 $s_2(t)$ 的信息点位于 $s_{21} = -\sqrt{E_b}$ 。图 6.3 为二进制 PSK 的信号空间图。图中还包括了两幅插图, 分别为代表 $s_1(t)$ 和 $s_2(t)$ 的一对反相信号的波形。值得注意的是图 6.3 的星座图是具有最小平均能量的。

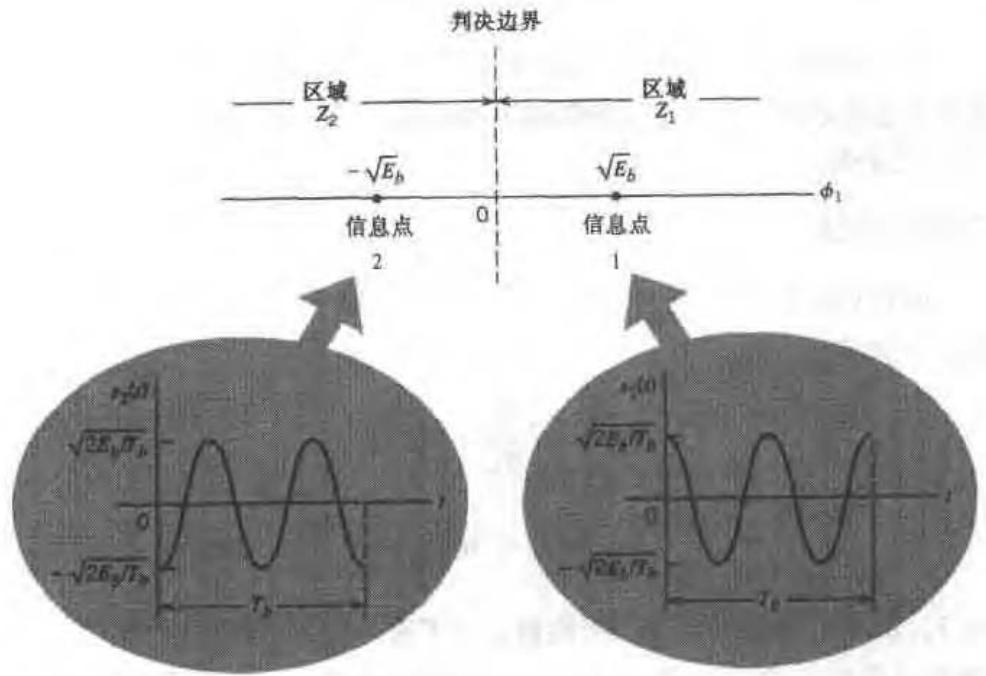


图 6.3 相干二进制 PSK 系统的信号空间图, 插图为发射信号 $s_1(t)$ 和 $s_2(t)$ 的波形, 假定 $n_c = 2$

6.3.2 二进制 PSK 的差错概率

为了得到对符号 1 和符号 0 的判决准则, 需要用到第 5 章中的式(5.59)。特别地, 将图 6.3 的信号空间分为两个区域:

- 接近 $+\sqrt{E_b}$ 处的信息点 1 的点集。
- 接近 $-\sqrt{E_b}$ 处的信息点 2 的点集。

通过确定这两个信息点连线的中点, 并由此标注出适当的判决空间, 就可以进行判决。在图 6.3 中, 这些判决空间根据围绕它们的信息点分别被标记成 Z_1 和 Z_2 。

当接收信号点落在区域 Z_1 时, 判决发送的是信号 $s_1(t)$ (如二进制符号 1)。当接收信号点落在区域 Z_2 时, 判决发送的是信号 $s_2(t)$ (如二进制符号 0)。但可能出现两种判决错误。当发射信号为 $s_2(t)$, 但噪声过大导致接收信号点落在区域 Z_1 时, 接收机将错误地判断为信号 $s_1(t)$ 。类似地, 当发射信号为 $s_1(t)$, 但噪声过大导致接收信号点落在区域 Z_2 时, 接收机将错误地判决为信号 $s_2(t)$ 。

下面计算第一种情况的差错概率。图 6.3 中与符号 1 或信号 $s_1(t)$ 对应的判决区域可以表述为

$$Z_1: 0 < x_1 < \infty$$

其中, 变量 x_1 与接收信号 $x(t)$ 有关:

$$x_1 = \int_0^{T_b} x(t)\phi_1(t) dt \quad (6.15)$$

假设符号 0 即信号 $s_2(t)$ 被发送, 则随机变量 X_1 的条件概率密度函数如下:

$$\begin{aligned} f_{X_1}(x_1 | 0) &= \frac{1}{\sqrt{\pi N_0}} \exp\left[-\frac{1}{N_0}(x_1 - s_{21})^2\right] \\ &= \frac{1}{\sqrt{\pi N_0}} \exp\left[-\frac{1}{N_0}(x_1 + \sqrt{E_b})^2\right] \end{aligned} \quad (6.16)$$

发送符号 0 而接收机判决为符号 1 的条件概率为

$$\begin{aligned} p_{10} &= \int_0^{\infty} f_{X_1}(x_1 | 0) dx_1 \\ &= \frac{1}{\sqrt{\pi N_0}} \int_0^{\infty} \exp\left[-\frac{1}{N_0}(x_1 + \sqrt{E_b})^2\right] dx_1 \end{aligned} \quad (6.17)$$

令

$$z = \frac{1}{\sqrt{N_0}}(x_1 + \sqrt{E_b}) \quad (6.18)$$

并将积分变量中的 x_1 替换为 z , 可将式(6.17)简化为

$$\begin{aligned} p_{10} &= \frac{1}{\sqrt{\pi}} \int_{-\sqrt{E_b/N_0}}^{\infty} \exp(-z^2) dz \\ &= \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \end{aligned} \quad (6.19)$$

其中, $\operatorname{erfc}(\cdot)$ 为互补误差函数。

下面计算第二种情况的差错概率。图 6.3 的信号空间是关于原点对称的, 因此, 发送符号 1, 而接收机判决为符号 0 的条件概率 p_{01} , 与式(6.19)的计算结果相同。

将条件差错概率 p_{10} 和 p_{01} 取平均, 可得相干二进制 PSK 的平均符号差错概率(或误比特率)如下(假设发送符号等概出现):

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \quad (6.20)$$

当增加每比特的发射信号功率 E_b 时,对于既定的噪声功率谱密度 N_0 ,对应于符号 1 和 0 的信息点相距更远,而平均差错概率 P_e 按式(6.20)相应减小,这显然是我们所期望的。

6.3.3 相干二进制 PSK 信号的产生和检测

由式(6.8)~式(6.10)可见,为产生二进制 PSK 信号,需用双极性信号表示输入的二进制序列,即用恒定电平 $+\sqrt{E_b}$ 和 $-\sqrt{E_b}$ 分别表示符号 1 和 0。编码过程由一个双极性不归零(NRZ)电平编码器完成。生成的二进制波形和正弦载波 $\phi_1(t)$ 被送入图 6.4(a)所示的乘积调制器,对于固定整数 n_c ,载波频率满足 $f_c = (n_c/T_b)$ 。载波和用来产生二进制波形的定时脉冲通常从共同的主时钟提取。这样,在调制器的输出端就可以得到期望的 PSK 波形。

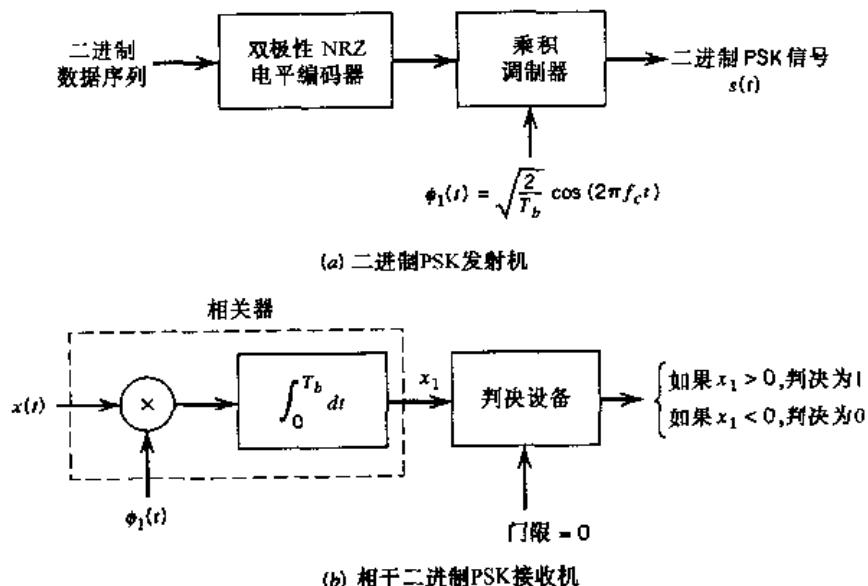


图 6.4 二进制 PSK 发射机和接收机框图

为了检测出由 1 和 0 组成的原始二进制序列,将含有噪声的 PSK 信号 $x(t)$ (在信道输出端)送入相关器,同时送入相关器的还有一个本地产生的相干参考信号 $\phi_1(t)$,如图 6.4(b)所示。相关器的输出 x_1 将和 0 V 的门限进行比较。如果 $x_1 > 0$,则接收机判决为符号 1;反之判决为符号 0。如果 x_1 恰好等于 0,则接收机通过随机猜测来决定是符号 0 还是符号 1。

6.3.4 二进制 PSK 信号的功率谱

由图 6.4(a)所示的调制器可见,二进制 PSK 波形的复包络仅含一个同相分量。进而,根据信号时间间隔 $0 \leq t \leq T_b$ 内发送的是符号 1 还是符号 0 可知,同相分量的取值为 $+g(t)$ 或 $-g(t)$ 。其中 $g(t)$ 是由下式确定的符号波形函数:

$$g(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}}, & 0 \leq t \leq T_b \\ 0, & \text{其他} \end{cases} \quad (6.21)$$

假设输入的二进制波形是随机的,符号 1 和符号 0 等概地出现,且不同的时间间隔内发送的符

号统计独立。由第1章中的例1.6可知,上述随机二进制波形的功率谱密度,等于该符号波形函数的能量谱密度除以符号持续时间。信号 $g(t)$ 的能量谱密度定义为该信号傅里叶变换的平方。因此,二进制PSK信号的基带功率谱密度等于

$$\begin{aligned} S_B(f) &= \frac{2E_b \sin^2(\pi T_b f)}{(\pi T_b f)^2} \\ &= 2E_b \operatorname{sinc}^2(T_b f) \end{aligned} \quad (6.22)$$

如图6.5所示,该功率随着频率平方值的增大而下降。

图6.5还绘制出了二进制FSK信号的基带功率谱密度,第6.5节将对其进行具体介绍。该节中还将对这两种谱进行比较。

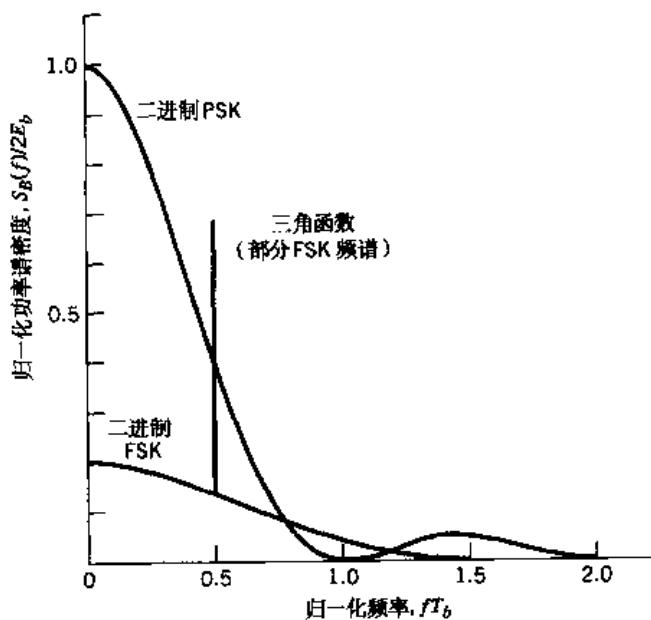


图6.5 二进制PSK和FSK信号的功率谱

6.3.5 QPSK

提供可靠的性能(例如较低的差错概率)是设计数字通信系统时的一个重要目标。另一个重要目标则是有效利用信道带宽。本节将研究一种节省带宽的调制方式,即相干四相移键控(QPSK),它是正交载波复用技术的一个例子。

QPSK与二进制PSK一样,传输信号包含的信息都存在于相位中。特别地,载波相位取四个等间隔值之一,如 $\pi/4, 3\pi/4, 5\pi/4$,和 $7\pi/4$ 。相应地,可将发射信号定义为

$$s_i(t) = \begin{cases} \sqrt{\frac{2E}{T}} \cos \left[2\pi f_c t + (2i - 1) \frac{\pi}{4} \right], & 0 \leq t \leq T \\ 0, & \text{其他} \end{cases} \quad (6.23)$$

其中, $i = 1, 2, 3, 4$; E 为发射信号每个符号的能量, T 为符号持续时间,载波频率 f_c 等于 n_c/T , n_c 为固定整数。每一个可能的相位值对应于一个特定的二位组。例如,可用前述的一组相位值来表示格雷码的一组二位组:10,00,01和11。值得注意的是,前后两个格雷码之间只有一位不同。

6.3.6 QPSK 的信号空间图

利用众所周知的三角恒等式,可用式(6.23)重新定义时间间隔 $0 \leq t \leq T$ 内的发射信号 $s_i(t)$:

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos\left[(2i-1)\frac{\pi}{4}\right] \cos(2\pi f_c t) - \sqrt{\frac{2E}{T}} \sin\left[(2i-1)\frac{\pi}{4}\right] \sin(2\pi f_c t) \quad (6.24)$$

其中 $i = 1, 2, 3, 4$ 。由式(6.24)不难发现:

- 在 $s_i(t)$ 的表达式中存在两个正交的基本函数 $\phi_1(t)$ 和 $\phi_2(t)$, 特别地, $\phi_1(t)$ 和 $\phi_2(t)$ 由一对正交载波确定:

$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), \quad 0 \leq t \leq T \quad (6.25)$$

$$\phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t), \quad 0 \leq t \leq T \quad (6.26)$$

- 存在 4 个信息点,因此,信号矢量定义如下:

$$s_i = \begin{bmatrix} \sqrt{E} \cos\left((2i-1)\frac{\pi}{4}\right) \\ -\sqrt{E} \sin\left((2i-1)\frac{\pi}{4}\right) \end{bmatrix}, \quad i = 1, 2, 3, 4 \quad (6.27)$$

信号矢量的组成部分,即 s_{i1} 和 s_{i2} 的取值见表 6.1。表 6.1 的前两列给出了二位组和对应的 QPSK 信号的相位。

表 6.1 QPSK 的信号空间特性

格雷码输入 二位组	QPSK 信号相位 (弧度)	信息点坐标	
		s_{i1}	s_{i2}
10	$\pi/4$	$+\sqrt{E/2}$	$-\sqrt{E/2}$
00	$3\pi/4$	$-\sqrt{E/2}$	$-\sqrt{E/2}$
01	$5\pi/4$	$-\sqrt{E/2}$	$+\sqrt{E/2}$
11	$7\pi/4$	$+\sqrt{E/2}$	$+\sqrt{E/2}$

因此,一个 QPSK 信号有二维的信号星座图($N=2$)和 4 个信息点($M=4$),其相位按逆时针方向增加,如图 6.6 所示。和二进制 PSK 一样,QPSK 具有最小的平均能量。

例 6.1

图 6.7 所示为产生一个 QPSK 信号的过程中包括的序列和波形。输入的二进制序列 01101000 如图 6.7(a)所示,该序列被分为两个序列,它们分别由输入序列的奇数位和偶数位组成,见图 6.7(b)和 6.7(c)顶部的标注。这两个序列对应的波形分别记为 $s_{i1}\phi_1(t)$ 和 $s_{i2}\phi_2(t)$,如图 6.7(b)和 6.7(c)所示。两个波形都可以被独立地视为二进制 PSK 信号的例子。将两个波形相加,可得图 6.7(d)所示的 QPSK 波形。

为了确定对传输的数据序列进行检测的判决准则,根据第 5 章的式(5.59),将信号空间分

成四个区域。各个区域定义为最接近由信号矢量 s_1, s_2, s_3 和 s_4 所表示的信息点的点集。连接四个信息点可构成一正方形,以正方形各边的垂直平分线来划分此区域,就可以进行判决。显然,判决区域分别是顶点位于原点的四个象限。对应于所围绕的信息点,这四个象限在图 6.6 中被分别标注为 Z_1, Z_2, Z_3 和 Z_4 。

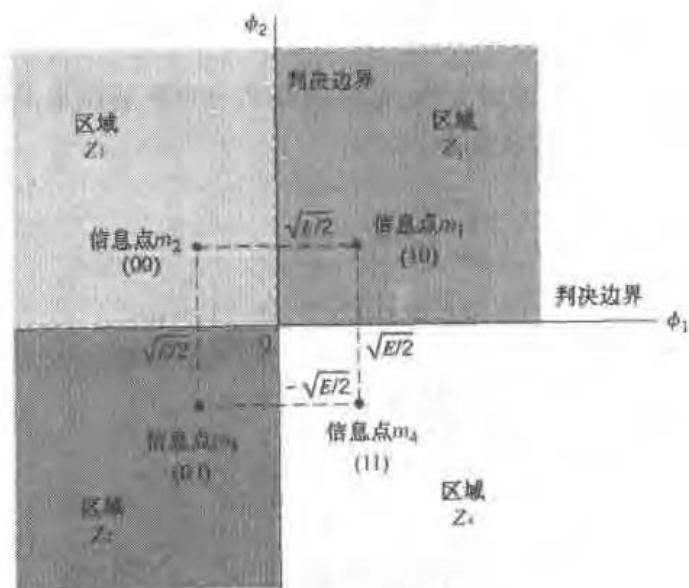


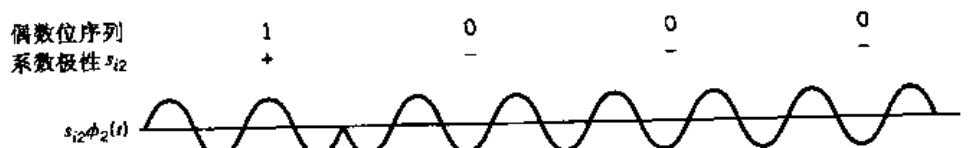
图 6.6 相干 QPSK 系统的信号空间图

输入二进制序列	$\underbrace{0 \quad 1}_{\text{二位组} 01}$	$\underbrace{1 \quad 0}_{\text{二位组} 10}$	$\underbrace{1 \quad 0}_{\text{二位组} 10}$	$\underbrace{0 \quad 0}_{\text{二位组} 00}$
---------	--	--	--	--

(a) 输入二进制序列



(b) 输入序列的奇数位及相应二进制PSK波形



(c) 输入序列的偶数位及相应二进制PSK波形

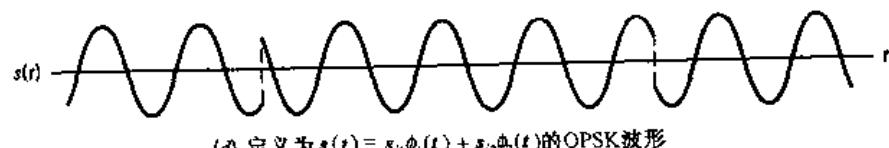
(d) 定义为 $s(t) = s_{i1}\phi_1(t) + s_{i2}\phi_2(t)$ 的QPSK波形

图 6.7 产生 QPSK 信号过程中的序列和波形

6.3.7 QPSK 的差错概率

相干 QPSK 系统的接收信号定义为

$$x(t) = s_i(t) + w(t), \quad \begin{cases} 0 \leq t \leq T \\ i = 1, 2, 3, 4 \end{cases} \quad (6.28)$$

其中, $w(t)$ 为具有零均值且功率谱密度为 $N_0/2$ 的高斯白噪声的样本函数。因此, 观察矢量 \mathbf{x} 有两个组成部分 x_1 和 x_2 , 分别定义如下:

$$\begin{aligned} x_1 &= \int_0^T x(t)\phi_1(t) dt \\ &= \sqrt{E} \cos\left[(2i - 1)\frac{\pi}{4}\right] + w_1 \\ &= \pm \sqrt{\frac{E}{2}} + w_1 \end{aligned} \quad (6.29)$$

和

$$\begin{aligned} x_2 &= \int_0^T x(t)\phi_2(t) dt \\ &= -\sqrt{E} \sin\left[(2i - 1)\frac{\pi}{4}\right] + w_2 \\ &= \mp \sqrt{\frac{E}{2}} + w_2 \end{aligned} \quad (6.30)$$

可见, 观测分量 x_1 和 x_2 分别是均值为 $\pm \sqrt{E/2}$ 和 $\mp \sqrt{E/2}$ 、方差为 $N_0/2$ 的高斯随机变量的抽样值。

现在很容易得到判决准则。若对应观察矢量 \mathbf{x} 的接收信号点落在区域 Z_1 , 则判决发送的是信号 $s_1(t)$; 若对应观察矢量 \mathbf{x} 的接收信号点落在区域 Z_2 , 则判决发送的是信号 $s_2(t)$, 以此类推。可能会出现判决错误, 例如, 当发送的信号为 $s_4(t)$, 但噪声 $w(t)$ 过大导致接收信号点落在了区域 Z_4 之外, 就会出现判决错误。

下面计算平均符号差错概率。由式(6.24)可知, 相干 QPSK 系统实际上相当于两个相位正交载波的相干二进制 PSK 系统并行工作。这说明了相干 QPSK 正交载波复用技术的特点。同相信道输出 x_1 和正交信道输出 x_2 (观察矢量 \mathbf{x} 的两个组成部分)可以看成是两个相干二进制 PSK 系统的相互独立的输出。由式(6.29)和式(6.30)可知, 这两个二进制 PSK 系统具有如下特征:

- 每比特的信号能量为 $E/2$ 。
- 噪声功率谱密度为 $N_0/2$ 。

因此, 可用式(6.20)来计算相干二进制 PSK 系统的平均误比特率。则相干 QPSK 系统每个信道的平均误比特率为

$$\begin{aligned} P' &= \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E/2}{N_0}}\right) \\ &= \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right) \end{aligned} \quad (6.31)$$

值得注意的另一点是,相干 QPSK 系统的同相和正交信道的比特差错是统计独立的。同相信道和正交信道分别利用组成 QPSK 信号(二位组)的两比特中的一个进行判决。因此,平均正确判决概率由两个信道的工作情况共同决定,即

$$\begin{aligned} P_c &= (1 - P')^2 \\ &= \left[1 - \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right)\right]^2 \\ &= 1 - \operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right) + \frac{1}{4} \operatorname{erfc}^2\left(\sqrt{\frac{E}{2N_0}}\right) \end{aligned} \quad (6.32)$$

而相干 QPSK 的平均符号差错概率则为

$$\begin{aligned} P_e &= 1 - P_c \\ &= \operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right) - \frac{1}{4} \operatorname{erfc}^2\left(\sqrt{\frac{E}{2N_0}}\right) \end{aligned} \quad (6.33)$$

在 $(E/2N_0) >> 1$ 的区域,可忽略式(6.33)右边的二次项。由此,得到相干 QPSK 平均符号差错概率的近似公式如下:

$$P_e \approx \operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right) \quad (6.34)$$

近似公式(6.34)也可以通过观察图 6.6 的信号空间图推导出来。由于图中的四个信息点关于原点对称,可利用式(5.92)进行推导。式(5.92)可写为

$$P_e \leq \frac{1}{2} \sum_{\substack{k=1 \\ k \neq i}}^4 \operatorname{erfc}\left(\frac{d_{ik}}{2\sqrt{N_0}}\right) \quad \text{对于所有 } i \quad (6.35)$$

假设发送信息点是 m_1 (相当于二位组 10),则与之最接近的信息点是 m_2 和 m_4 (相当于二位组 00 和 11)。由图 6.6 可知, m_1 与 m_2 和 m_4 的欧氏距离是相等的,即

$$d_{12} = d_{14} = \sqrt{2E}$$

假设 E/N_0 足够大,就可忽略距离 m_1 最远的信息点 m_3 (相当于二位组 01)的影响,可见利用式(6.35)得到的 P_e 近似表达式与式(6.34)的结果一致。值得注意的是,当把 m_1 误判为 m_2 或者 m_4 时,只有一位出错;但当把 m_1 误判为 m_3 时,则有两位出错。当 E/N_0 足够大时,一个符号的两位同时出错的概率远小于仅有一位出错的概率,因此若发射信号为 m_1 ,在计算 P_e 时可以忽略 m_3 的影响。

在 QPSK 系统中,每个符号对应两比特。因此,每个符号的发射信号能量是单比特信号能量的两倍,即

$$E = 2E_b \quad (6.36)$$

用比值 E_b/N_0 表示平均符号差错概率, 即

$$P_e \approx \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \quad (6.37)$$

当输入符号采用格雷码时, 由式(6.31)和式(6.36)可见, QPSK 的误比特率恰为

$$\text{BER} = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \quad (6.38)$$

综上所述, 当比特率和 E_b/N_0 的值都相同时, 相干 QPSK 系统具有与相干二进制 PSK 系统相同的平均误比特率, 但其占用的信道带宽仅为后者的一半。换言之, 对于相同的 E_b/N_0 和平均误比特率, 占用相同的信道带宽的相干 QPSK 系统的发送信息比特率是相干二进制 PSK 的两倍。对于指定的性能要求, QPSK 的信道带宽性能要优于二进制 PSK, 因此 QPSK 比二进制 PSK 应用更为广泛。

6.3.8 相干 QPSK 信号的产生和检测

下面考虑相干 QPSK 信号的产生和检测。图 6.8(a)为典型的 QPSK 发射机框图。输入的二进制数据序列首先被不归零(NRZ)电平编码器转换为极性形式, 即符号 1 和 0 分别用 $+\sqrt{E_b}$ 和 $-\sqrt{E_b}$ 表示。接着, 该二进制波形被分接器分成两个分别由输入序列的奇数位和偶数位组成的彼此独立的二进制波形, 这两个二进制波形分别用 $a_1(t)$ 和 $a_2(t)$ 表示。容易注意到, 在任一信号时间间隔内, $a_1(t)$ 和 $a_2(t)$ 的幅度恰好分别等于 s_{11} 和 s_{12} , 即由发送的二位组决定。这两个二进制波形 $a_1(t)$ 和 $a_2(t)$ 被用来调制一对正交载波或者说正交基本函数: $\phi_1(t) = \sqrt{2/T} \cos(2\pi f_c t)$, $\phi_2(t) = \sqrt{2/T} \sin(2\pi f_c t)$ 。这样就得到一对二进制 PSK 信号。 $\phi_1(t)$ 和 $\phi_2(t)$ 的正交性使这两个信号可以被独立地检测。最后, 将这两个二进制 PSK 信号相加, 从而得到期望的 QPSK 信号。

如图 6.8(b)所示, QPSK 接收机由一对共输入的相关器组成。这两个相关器分别提供本地产生的相干参考信号 $\phi_1(t)$ 和 $\phi_2(t)$ 。相应于接收信号 $x(t)$, 相关器输出的 x_1 和 x_2 被用来与门限值 0 进行比较。如果 $x_1 > 0$, 则判决同相信道的输出为符号 1; 如果 $x_1 < 0$, 则判决同相信道的输出为符号 0。类似地, 如果 $x_2 > 0$, 则判决正交信道的输出为符号 1; 如果 $x_2 < 0$, 则判决正交信道的输出为符号 0。最后, 同相信道和正交信道输出的这两个二进制数据序列被复接器合并, 重新得到原始的二进制序列。在 AWGN 信道中, 判决结果具有最小的符号差错概率。

6.3.9 QPSK 信号的功率谱密度

假设调制器输入的二进制波形是随机的, 符号 1 和 0 等概地出现, 相邻时间间隔内发送的符号统计独立, 则对于 QPSK 信号的同相和正交分量, 有以下结论:

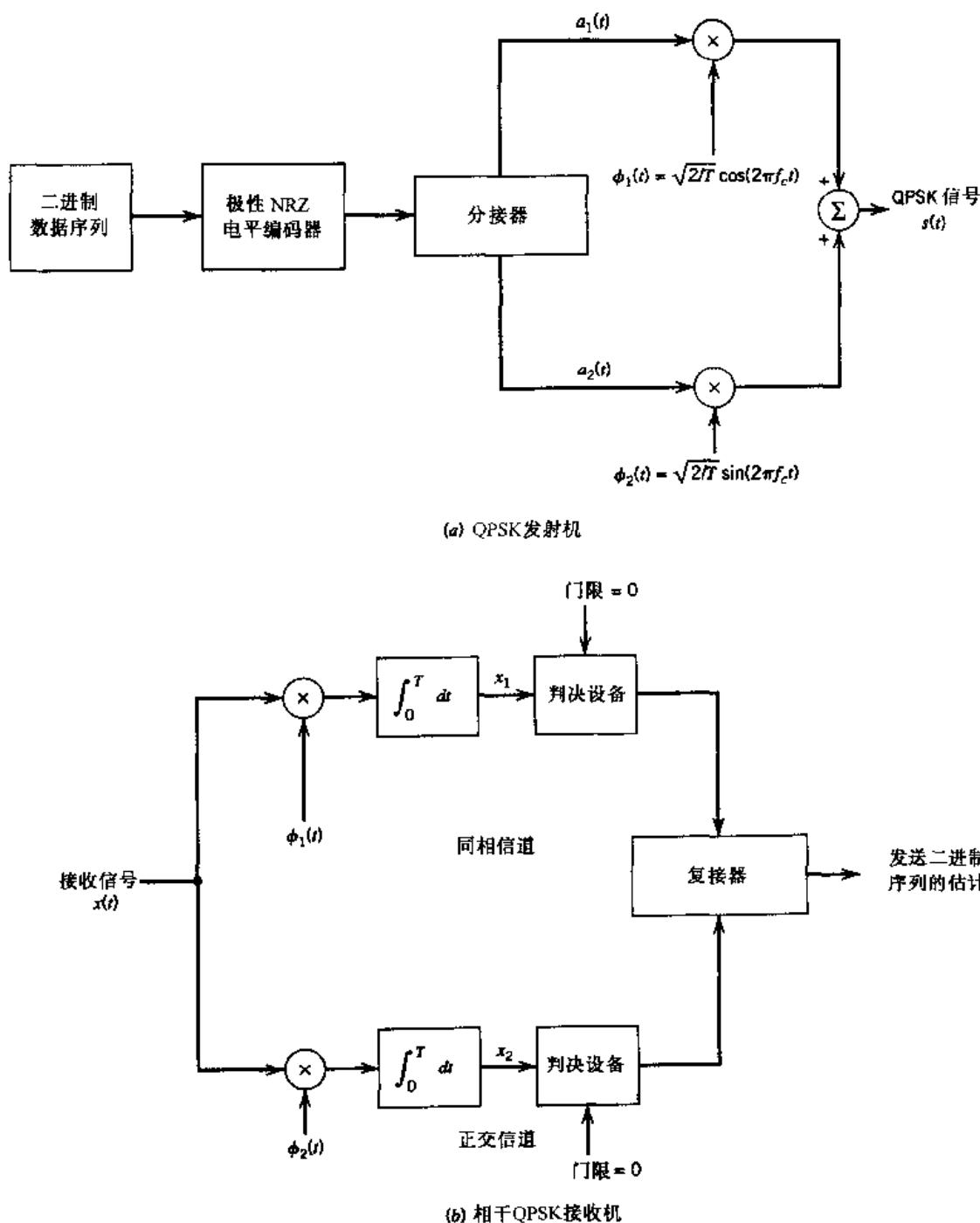


图 6.8 QPSK 发射机和接收机框图

- 根据信号时间间隔 $-T_b \leq t \leq T_b$ 内发送的二位组的值, 同相分量等于 $+g(t)$ 或 $-g(t)$, 正交分量与之类似。 $g(t)$ 表示符号波形函数, 即

$$g(t) = \begin{cases} \sqrt{\frac{E}{T}}, & 0 \leq t \leq T \\ 0, & \text{其他} \end{cases} \quad (6.39)$$

因此, 同相分量和正交分量具有相同的功率谱密度, 即 $E \operatorname{sinc}^2(Tf)$ 。

2. 同相分量和正交分量统计独立。因此, QPSK 信号的基带功率谱密度等于同相和正交分量各自功率谱密度之和, 即

$$\begin{aligned} S_B(f) &= 2E \operatorname{sinc}^2(Tf) \\ &= 4E_b \operatorname{sinc}^2(2T_b f) \end{aligned} \quad (6.40)$$

图 6.9 为按 $4E_b$ 归一化的 $S_B(f)$ 相对于归一化频率 fT_b 的变化趋势。该图还描绘了一种称为最小频移键控的二进制 FSK 的基带功率谱密度曲线。6.5 节将对最小频移键控进行具体介绍, 并对这两种谱进行比较。

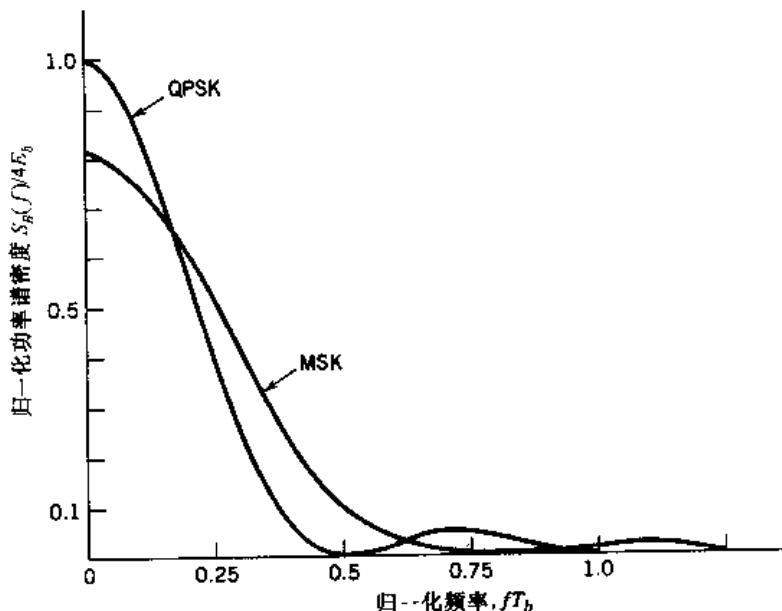


图 6.9 QPSK 和 MSK 信号的功率谱

6.3.10 偏移 QPSK

图 6.10(a)所示的信号空间图包含了产生 QPSK 信号过程中所有可能出现的相位转换。特别地, 考察图 6.7 中例 6.1 的 QPSK 波形, 可以看出:

1. 当 QPSK 信号的同相分量和正交分量同时改变符号时, 载波相位将有 $\pm 180^\circ$ 的变化。图 6.7 中当输入二进制序列从 01 变成 10 时, 就属于这种情况。
2. 当同相分量或正交分量改变符号时, 载波相位将有 $\pm 90^\circ$ 的变化。图 6.7 中当输入二进制序列从 10 变成 00, 即同相分量改变符号而正交分量符号不变时, 就属于这种情况。
3. 当同相分量和正交分量的符号都保持不变时, 载波相位不变。图 6.7 中当两个相继的信号时间间隔内都发送 10 时, 就属于这种情况。

QPSK 信号进行传输时, 在解调检测之前要进行滤波。尤其需要注意的是情况 1 以及次之的情况 2。特别地, 载波相位 180° 和 90° 的变化将带来载波幅度 (QPSK 信号的包络) 的变化, 从而在检测时带来额外的符号差错。

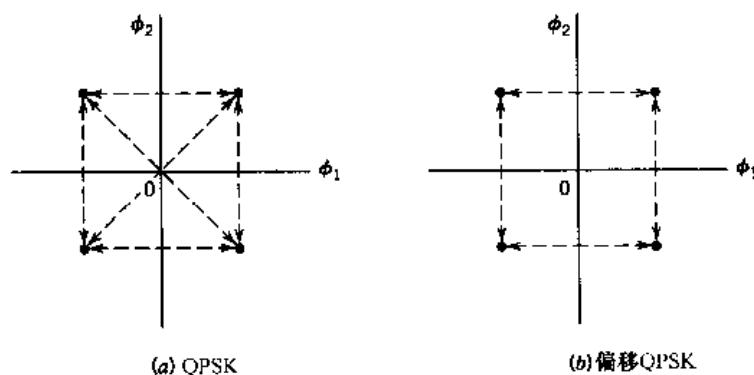


图 6.10 QPSK 和偏移 QPSK 信息点之间可能的转移路径

采用偏移 QPSK^[2]可减小 QPSK 信号幅度的波动。在偏移 QPSK 中,产生正交分量的比特流相对于产生同相分量的比特流延迟(偏移)了半个信号时间间隔。偏移 QPSK 的两个基本函数为

$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), \quad 0 \leq t \leq T \quad (6.41)$$

$$\phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t), \quad \frac{T}{2} \leq t \leq \frac{3T}{2} \quad (6.42)$$

由图 6.10(b)所示的信号空间图可见,偏移 QPSK 可能产生的相位变化被限制在 $\pm 90^\circ$ 之间,这一点与 QPSK 截然不同。偏移 QPSK 中相位变化的发生频率是 QPSK 的 2 倍,而变化强度则是 QPSK 的 $1/2$ 。由于 QPSK 中除了 $\pm 90^\circ$ 的相位变化以外,还存在 $\pm 180^\circ$ 的相位变化,因此偏移 QPSK 中由滤波引起的幅度波动的振幅比 QPSK 要小。

比较式(6.42)和式(6.26)可知,除了基本函数 $\phi_2(t)$ 存在 $T/2$ 的时延外,偏移 QPSK 和 QPSK 在 AWGN 信道中具有完全一致的符号差错概率。这两种相移键控方式获得同等的噪声性能的前提是采用相干检测。二者能够具有同等噪声性能的原因是其同相分量和正交分量具有统计独立性。由此可见,相干偏移 QPSK 接收机的同相分量及正交分量的差错概率仍然等于 $(1/2)\text{erfc}\sqrt{E/2N_0}$,因此,式(6.34)在偏移 QPSK 中依然适用。

6.3.11 $\pi/4$ 移相 QPSK

如图 6.11(a)和 6.11(b)所示,一个 QPSK 信号可能位于两个常用星座图里的任一个中,这两个星座图之间有 $\pi/4$ 的相位差。在 $\pi/4$ 移相 QPSK^[3] 中,用来传送连续符号(二位组)的载波相位从图 6.11 所示的两个 QPSK 星座图中交替选择。因此 $\pi/4$ 移相 QPSK 信号可能处于 8 个可能的相位状态之中的任一个,如图 6.12 所示。

表 6.2 总结了 $\pi/4$ 移相 QPSK 的相位变化与输入的格雷码二位组之间的一组可能关系。例如,当调制器处于图 6.11(b)所示的相位状态之一时,二位组 00 的接收将使相位逆时针旋转 $\pi/4$,从而处于图 6.11(a)所示的相位状态之一。

$\pi/4$ 移相 QPSK 具有以下优点:

- 相邻符号之间的相位变化限制在 $\pm \pi/4 \sim \pm 3\pi/4$ 的范围内,这比 QPSK 中的 $\pm \pi/2$ 到 $\pm \pi$

的变化范围要窄。因此, $\pi/4$ 移相 QPSK 中由于滤波造成的包络变化比 QPSK 中要小得多。

- 和偏移 QPSK 不同的是, $\pi/4$ 移相 QPSK 信号可以进行非相干检测, 这就大大简化了接收机的设计。此外, 与 QPSK 信号类似, $\pi/4$ 移相 QPSK 也可以进行差分编码, 这种编码方法相应地称为 $\pi/4$ 移相 DQPSK。

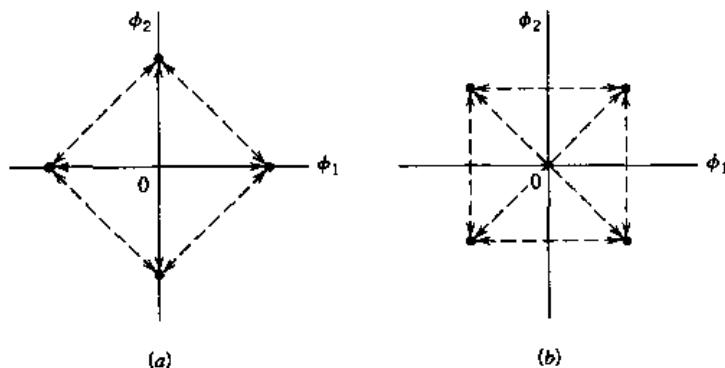


图 6.11 QPSK 的两种常用信号空间图;箭头所指为 QPSK 调制器改变状态的路径

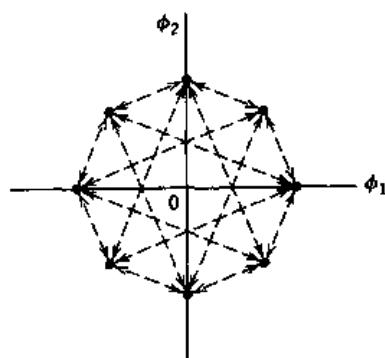


图 6.12 $\pi/4$ 移相 QPSK 调制器的 8 种可能的相位状态

$\pi/4$ 移相 DQPSK 符号可用符号对 (I, Q) 表示, 符号对的产生由下面的一对关系式描述(见习题 6.13):

$$\begin{aligned} I_k &= \cos(\theta_{k-1} + \Delta\theta_k) \\ &= \cos \theta_k \end{aligned} \quad (6.43)$$

$$\begin{aligned} Q_k &= \sin(\theta_{k-1} + \Delta\theta_k) \\ &= \sin \theta_k \end{aligned} \quad (6.44)$$

其中, θ_{k-1} 为符号 $k - 1$ 的绝对相角, $\Delta\theta_k$ 为按照表 6.2 定义的差分编码的相位变化。

表 6.2 输入二位组与 $\pi/4$ 移相 DQPSK 相位变化的对应关系

格雷码输入二位组	相位变化, $\Delta\theta$ (弧度)
00	$\pi/4$
01	$3\pi/4$
11	$-3\pi/4$
10	$-\pi/4$

例 6.2

本例仍采用例 6.1 中的输入二进制序列 01101000。假设 $\pi/4$ 移相 DQPSK 调制器的初始相位状态是图 6.11(b) 所示星座图中的相角 $\theta_0 = \pi/4$ 。以表 6.2 所示的二位组序列作为输入，可得到表 6.3 所示的结果。

表 6.3 例 6.2 所示 $\pi/4$ 移相 DQPSK 的结果

步 骤	相位 θ_{k-1} (弧度)	输入二位组	相位变化 $\Delta\theta_k$ (弧度)	发送相位 θ_k (弧度)
1	$\pi/4$	00	$\pi/4$	$\pi/2$
2	$\pi/2$	10	$-\pi/4$	$\pi/4$
3	$\pi/4$	10	$-\pi/4$	0
4	0	01	$3\pi/4$	$3\pi/4$

6.3.12 $\pi/4$ 移相 DQPSK 信号的检测

下面介绍 $\pi/4$ 移相 DQPSK 信号的差分检测。对于含有噪声的信道输出 $x(t)$ ，接收机首先计算出 $x(t)$ 在基本函数 $\phi_1(t)$ 和 $\phi_2(t)$ 上的投影，然后将输出结果（分别由 I 和 Q 表示）送入图 6.13 所示的差分检测器。该差分检测器的组成如下：

- 用来提取信道输出（接收信号）相角 θ 的反正切计算器。
- 用来确定相角 θ 在一个信号时间间隔内的相位变化的相位差计算器。
- 用来纠正可能由相角绕回实轴而引起的错误的模 2π 纠错逻辑。

在上述最后一条中，假设符号 k 和 $k-1$ 对应的信道输出相角分别为 θ_k 和 θ_{k-1} ， $\Delta\theta_k$ 为 θ_k 和 θ_{k-1} 之间的相位差，则模 2π 纠错逻辑按照如下方式工作：

$$\begin{aligned} &\text{如果 } \Delta\theta_k < -180^\circ, \text{ 则 } \Delta\theta_k = \Delta\theta_k + 360^\circ \\ &\text{如果 } \Delta\theta_k > 180^\circ, \text{ 则 } \Delta\theta_k = \Delta\theta_k - 360^\circ \end{aligned} \quad (6.45)$$

为了说明这种相位纠错的必要性，考虑图 6.14 所示的情况，即 $\theta_{k-1} = 350^\circ$ 和 $\theta_k = 60^\circ$ ，均为逆时针方向。由图可知，经过相位纠错，相位变化 $\Delta\theta_k$ 为 70° 。但如果不行相位纠错，相位变化 $\Delta\theta_k$ 的计算结果则为 $60^\circ - 350^\circ = -290^\circ$ 。由式(6.45)的第一行，模 -2π 纠错逻辑能够补偿相角靠近正实轴而引起的差错，纠错结果为

$$\Delta\theta_k = -290^\circ + 360^\circ = 70^\circ$$

图 6.13 中用来解调 $\pi/4$ 移相 DQPSK 信号的正切型差分检测器实现起来相对简单。这种检测器在恒定多径环境下的瑞利衰落信道中能够提供令人满意的检测效果。但是，仿真结果表明，在多径环境如商用数字无线通信系统中，采用这种检测器将引起接收机性能的快速下降。^[4]

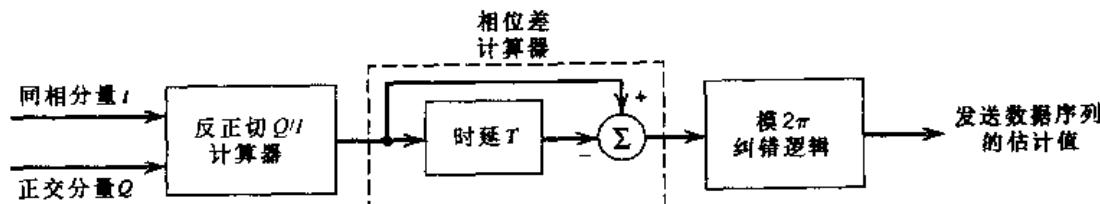


图 6.13 $\pi/4$ 移相 DQPSK 检测器框图

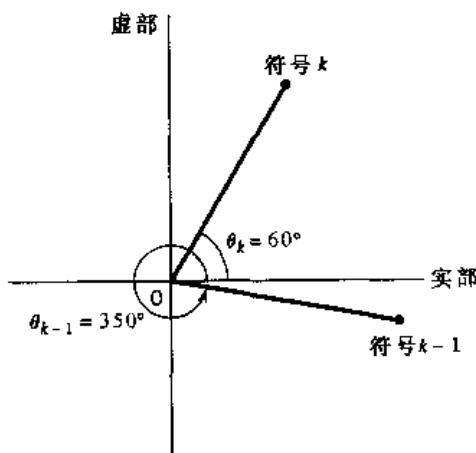


图 6.14 可能的靠近正实轴的相角例子

6.3.13 M 进制 PSK

QPSK 是 M 进制 PSK 的一个特例。 M 进制 PSK 的载波相位取 M 个可能值之一, 即 $\theta_i = 2(i-1)\pi/M$, 其中 $i = 1, 2, \dots, M$ 。因此, 每个持续时间为 T 的信号时间间隔内发送的信号为

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos\left(2\pi f_c t + \frac{2\pi}{M}(i-1)\right), \quad i = 1, 2, \dots, M \quad (6.46)$$

其中, E 为单个符号的信号能量。载波相位 $f_c = n_c/T$ 由固定整数 n_c 确定。

每个 $s_i(t)$ 可以分别按照式(6.25)和式(6.26)定义的基本函数 $\phi_1(t)$ 和 $\phi_2(t)$ 展开。因此, M 进制 PSK 的信号星座图也是二维的。 M 个信息点等距离分布在半径为 \sqrt{E} , 圆心在原点的圆周上, 图 6.15(a) 为八相 PSK 星座图的例子 ($M = 8$)。

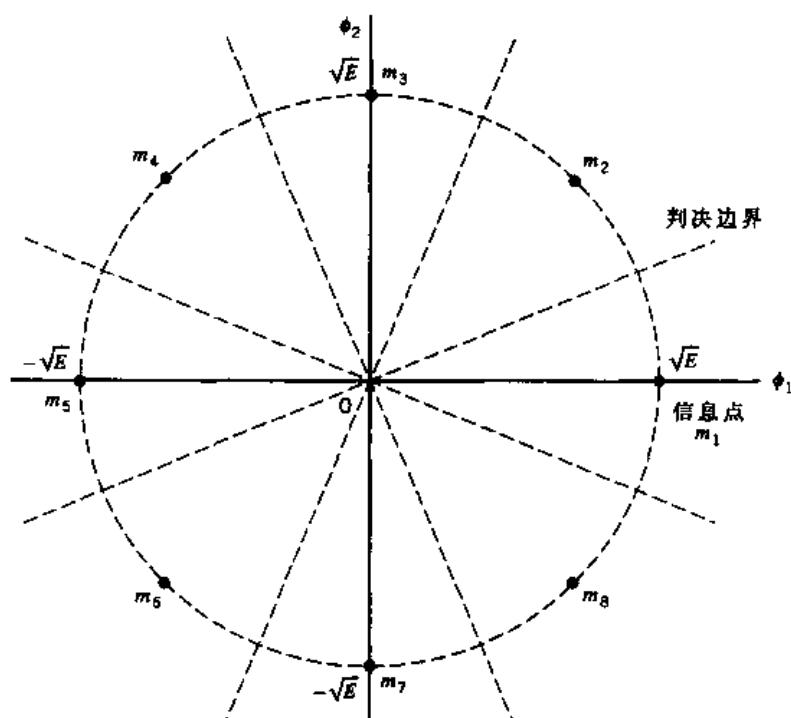
由图 6.15(a) 可知信号空间图是环绕对称的, 因此, 可利用式(5.92)以及联合边界, 推导出 M 进制 PSK 平均符号差错概率的近似公式。假设发射信号对应于信息点 m_1, m_1 在 ϕ_1 轴和 ϕ_2 轴的坐标分别为 $+\sqrt{E}$ 和 0。假设信噪比 E/N_0 足够大, 从而, 仅有分布在 m_1 两侧并与之最接近的两个信息点可能因信道噪声而被错误判决为 m_1 。图 6.15(b) 即为当 $M = 8$ 时上述假设对应的情况。其中上述两个信息点与 m_1 之间的欧氏距离为

$$d_{12} = d_{18} = 2\sqrt{E} \sin\left(\frac{\pi}{M}\right)$$

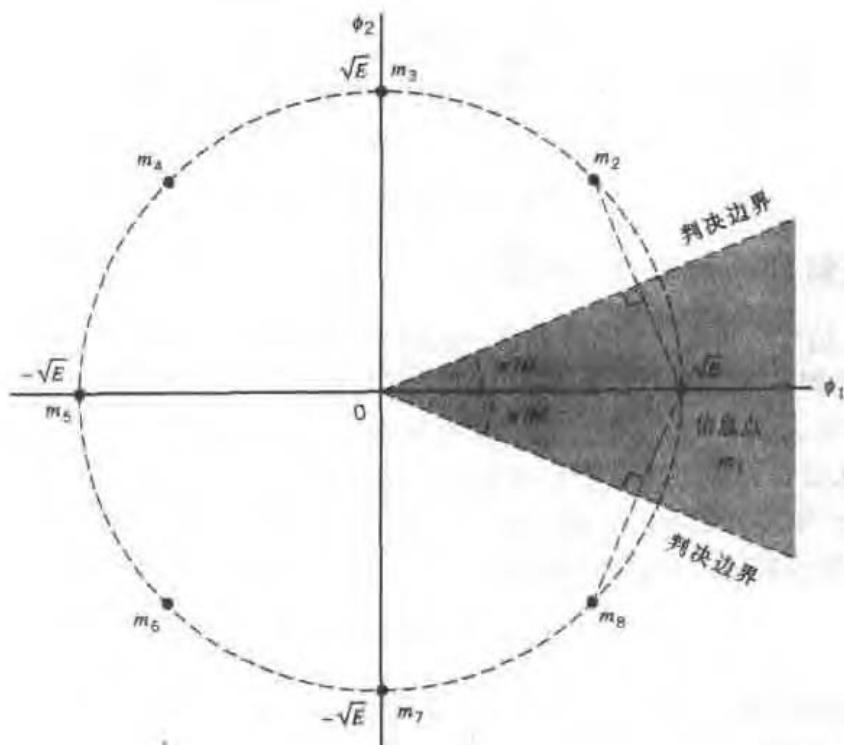
由第 5 章的式(5.92)可得相干 M 进制 PSK 的平均符号差错概率为

$$P_e \approx \operatorname{erfc}\left(\sqrt{\frac{E}{N_0}} \sin\left(\frac{\pi}{M}\right)\right) \quad (6.47)$$

其中, 假设 $M \geq 4$ 。当 M 固定而 E/N_0 增大时, 近似值和真实值之间的差别将减小。当 $M = 4$ 时, 式(6.47)可简化为式(6.34)中 QPSK 的形式。



(a) 八相PSK的信号空间图(M=8), 判决边界用虚线表示



(b) 八相PSK联合边界的信号空间图

图 6.15 八相 PSK 及其联合边界的信号空间图

6.3.14 M 进制 PSK 信号的功率谱密度

M 进制 PSK 的符号持续时间定义如下：

$$T = T_b \log_2 M \quad (6.48)$$

$$\rho = \frac{R_b}{B} = \frac{\log_2 M}{2} \quad (6.52)$$

表 6.4 给出了利用式(6.52), 在不同 M 值下计算所得的 ρ 值。

表 6.4 M 进制 PSK 信号的带宽效率

M	2	4	8	16	32	64
$\rho(\text{bs}^{-1}\text{Hz}^{-1})$	0.5	1	1.5	2	2.5	3

根据式(6.47)和表 6.4, 对 M 进制 PSK 可得出如下结论: 随着状态数 M 的增大, 带宽效率也得到提高, 但同时会带来误码性能的恶化。为了确保误码性能不发生恶化, 就必须提高 E_b/N_o , 以补偿 M 增大带来的影响。

6.4 幅度/相位混合调制方案

在 M 进制 PSK 系统中, 调制信号的同相分量和正交分量之间的相互关系使得信号包络保持恒定, 这从信息点的圆形星座图中可以清楚地看出。如果取消包络恒定的约束, 同相分量和正交分量就能够彼此独立, 从而得到一种新的调制方式, 即 M 进制 QAM。QAM 是一种联合调制方式, 因为载波同时受到幅度和相位的调制。

当根据符号速率 $1/T$ 任意选取载波频率 f_c 时, M 进制 QAM 的通带基本函数可能并不具备周期性, 但这种非周期性通常并不会引起注意。通过以某种方式重新表示发射信号, 就可以忽略连续进行符号发送时基本函数的时间变化, 这也就简化了发射机的实现。特别地, 当完全保持幅度相位混合调制过程的本质时, 发射机看上去是“无载波”的。这实际上就是无载波幅度/相位调制(CAP)的基本思想。

QAM 和 CAP 信号星座图除了在实现细节上有所差别外, 是完全相同的。因此, 对于既定的接收机复杂度, 二者的性能基本相同。下面先介绍 QAM, 随后再介绍 CAP。

6.4.1 M 进制 QAM

第 4 和第 5 章研究了 M 进制 PAM。PAM 是一维的。 M 进制 QAM 则是 M 进制 PAM 的二维推广形式, 包含了两个正交的通带基本函数, 如下所示:

$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), \quad 0 \leq t \leq T \quad (6.53)$$

$$\phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t), \quad 0 \leq t \leq T \quad (6.54)$$

假设将 (ϕ_1, ϕ_2) 平面内的第 i 个信息点 s_i 记为 $(a_i d_{\min}/2, b_i d_{\min}/2)$, 其中 d_{\min} 为星座图上任意两个信息点之间的最小距离, a_i 和 b_i 为整数, $i = 1, 2, \dots, M$ 。假设 $d_{\min}/2 = \sqrt{E_0}$, 其中 E_0 为具有最小幅度的信号的能量。符号 k 对应的 M 进制 QAM 信号的定义如下:

$$s_k(t) = \sqrt{\frac{2E_0}{T}} a_k \cos(2\pi f_c t) + \sqrt{\frac{2E_0}{T}} b_k \sin(2\pi f_c t), \quad 0 \leq t \leq T \quad k = 0, \pm 1, \pm 2, \dots \quad (6.55)$$

信号 $s_k(t)$ 由两个相位正交的离散 AM 载波构成, 因此这种调制方式称为 QAM。

对应于各种可能的符号数 M , 有两种不同的 QAM 星座图: 当每个符号对应于偶数个比特时, 星座图是正方形的; 当每个符号对应于奇数个比特时, 星座图是十字形的。这两种情况将在下面依次讨论。

6.4.2 QAM 正方形星座图

当每个符号对应于偶数个比特时, 有

$$L = \sqrt{M} \quad (6.56)$$

其中, L 为正整数。因此, 可将 M 进制 QAM 的正方形星座图看成是一维 L 元 PAM 信号与其自身的笛卡儿积。根据定义, 两组坐标(代表了一对一维星座图)的笛卡儿积由一组可能出现的有序坐标对组成。其中, 各个坐标对的第一个坐标取自笛卡儿积的第一组, 第二个坐标取自笛卡儿积(Cartesian Product)的第二组。

对于具有正方形星座图的 QAM, 有序坐标对通常构成一个正方形的矩阵

$$\{a_i, b_i\} = \begin{bmatrix} (-L+1, L-1) & (-L+3, L-1) & \cdots & (L-1, L-1) \\ (-L+1, L-3) & (-L+3, L-3) & \cdots & (L-1, L-3) \\ \vdots & \vdots & & \vdots \\ (-L+1, -L+1) & (-L+3, -L+1) & \cdots & (L-1, -L+1) \end{bmatrix} \quad (6.57)$$

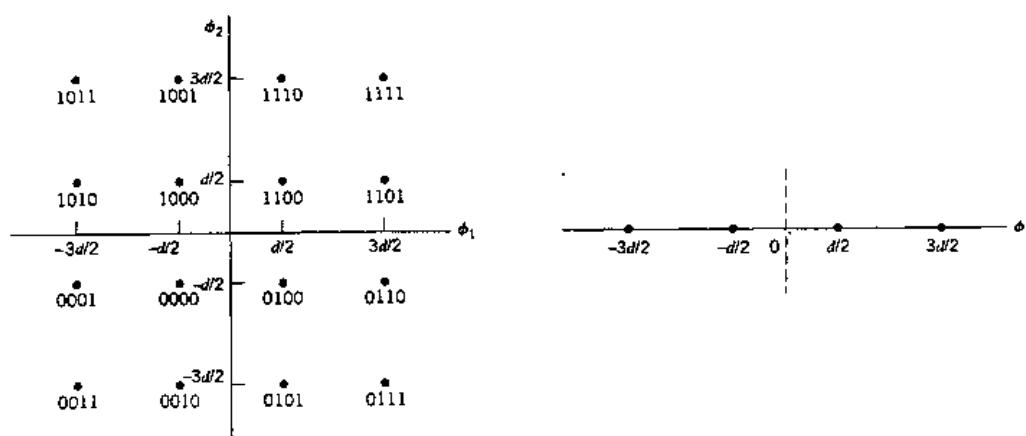
例 6.3

考虑如图 6.17(a) 所示的 16-QAM 信号星座图, 图中信息点的编码如下:

- 最左端的两位, 指定了信息点在 (ϕ_1, ϕ_2) 平面内所处的象限。也就是说, 从第一象限起逆时针旋转, 经过的四个象限分别用二位组 11, 10, 00 和 01 表示。
- 剩下的两位用来表示 (ϕ_1, ϕ_2) 平面各个象限内的四个可能符号之一。

值得注意的是对四个象限的编码以及对各个象限内符号的编码都应依照格雷码的规则进行。

本例中 $L=4$, 因此, 图 6.17(a) 所示的正方形星座图是图 6.17(b) 所示 4-PAM 星座图与其自身的笛卡儿积。此外, 式(6.57)的矩阵取值如下:



(a) M 进制 QAM 在 $M=16$ 时的信号空间图。每个象限的信息点由格雷码的四位组确定

(b) 对应的 4-PAM 信号空间图

图 6.17 16-QAM 与 4-PAM 信号空间图

$$\{a_i, b_i\} = \begin{bmatrix} (-3, 3) & (-1, 3) & (1, 3) & (3, 3) \\ (-3, 1) & (-1, 1) & (1, 1) & (3, 1) \\ (-3, -1) & (-1, -1) & (1, -1) & (3, -1) \\ (-3, -3) & (-1, -3) & (1, -3) & (3, -3) \end{bmatrix}$$

在 M 进制 QAM 的符号差错概率的计算中, 需要用到 QAM 正方形星座图可被分解为相应的 PAM 星座图与其自身笛卡儿积的特性。计算过程如下:

1. M 进制 QAM 的正确检测概率为

$$P_c = (1 - P'_e)^2 \quad (6.58)$$

其中, P'_e 为相应 L 进制 PAM 的符号差错概率, $L = \sqrt{M}$ 。

2. 符号差错概率 P'_e 的定义如下:

$$P'_e = \left(1 - \frac{1}{\sqrt{M}}\right) \operatorname{erfc}\left(\sqrt{\frac{E_0}{N_0}}\right) \quad (6.59)$$

要注意的是, M 进制 QAM 中的 $L = \sqrt{M}$ 对应于习题 4.27 中的 M 进制 PAM 中的 M 。

3. M 进制 QAM 的符号差错概率是

$$\begin{aligned} P_e &= 1 - P_c \\ &= 1 - (1 - P'_e)^2 \\ &\approx 2P'_e \end{aligned} \quad (6.60)$$

其中, 假设 P'_e 远小于 1, 从而可忽略平方项。

将式(6.58)和式(6.59)代入式(6.60), 可得 M 进制 QAM 的符号差错概率的近似值为

$$P_e \approx 2\left(1 - \frac{1}{\sqrt{M}}\right) \operatorname{erfc}\left(\sqrt{\frac{E_0}{N_0}}\right) \quad (6.61)$$

由于 M 进制 QAM 的瞬时值取决于特定的发送符号, 因此其发射能量也是变化的。因此用发送能量的平均值来表示 P_e , 比用 E_0 更为合理。假设同相分量或正交分量的 L 个幅度值是等概率出现的, 即有

$$E_{av} = 2 \left[\frac{2E_0}{L} \sum_{i=1}^{L/2} (2i - 1)^2 \right] \quad (6.62)$$

方括号外的乘法因子 2 说明了同相分量和正交分量的作用是相同的。方括号里求和公式上的上、下限以及乘法因子 2 说明了零点两侧相关幅度的对称性。将式(6.62)中的级数求和, 得

$$\begin{aligned} E_{av} &= \frac{2(L^2 - 1)E_0}{3} \\ &= \frac{2(M - 1)E_0}{3} \end{aligned} \quad (6.63)$$

因此, 可用 E_{av} 的形式将式(6.61)改写为

$$P_e \approx 2\left(1 - \frac{1}{\sqrt{M}}\right) \operatorname{erfc}\left(\sqrt{\frac{3E_{av}}{2(M - 1)N_0}}\right) \quad (6.64)$$

这正是我们所期望的结果。

$M = 4$ 的情况是我们特别关注的,此时的信号星座图和 QPSK 的星座图一样。实际上,在式(6.64)中设 $M = 4$ (此时 $E_s = E$,其中 E 为单个符号的能量),得到的符号差错概率公式与式(6.34)是一致的,这和理论结果是相符合的。

6.4.3 QAM 十字形星座图

为产生每个符号对应于奇数个比特的 M 进制 QAM 信号,需要采用十字形星座图。如图 6.18 所示,按照下述步骤,就可以构造出每个符号对应于 n 比特的信号星座图:

- 首先给出每个符号对应于 $n - 1$ 比特的 QAM 正方形星座图。
- 通过在各边都增加 2^{n-3} 个符号来拓展 QAM 正方形星座图。
- 拓展时,忽略正方形的四角。

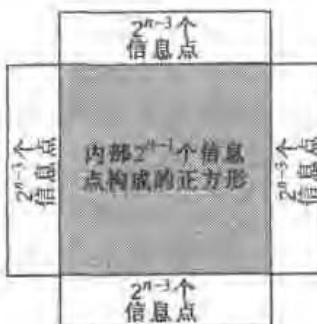


图 6.18 将正方形 QAM 星座图扩展为 QAM 十字形星座图

内部的正方形代表了 2^{n-1} 个符号,四边的拓展共增加了 $4 \times 2^{n-3} = 2^{n-1}$ 个符号。因此,十字形星座图中的符号总数为 $2^{n-1} + 2^{n-1} = 2^n$,这就代表了每个符号对应于 n 比特的情况。

与 QAM 正方形星座图不同的是,QAM 十字形星座图不能表示为一个 PAM 星座图与其自身的笛卡儿积。因此,具有十字形星座图的 M 进制 QAM 的符号差错概率 P_e 的确定比较复杂。这里我们只是不加证明地给出 P_e 的计算公式:

$$P_e \approx 2 \left(1 - \frac{1}{\sqrt{2M}} \right) \operatorname{erfc} \left(\sqrt{\frac{E_0}{N_0}} \right), E_0/N_0 \text{ 很大} \quad (6.65)$$

这与正方形星座图对应的式(6.61)的表达是基本一致的,只是在星座图的每一维都包含有额外的 0.5 比特。^[5]值得注意的是,对 QAM 十字形星座图是无法很好地进行格雷编码的。

6.4.4 CAP

式(6.53)和式(6.54)所示的通带基本函数中,假设使用矩形脉冲作为脉冲波形函数。下面用通用脉冲波形函数 $g(t)$ 重新定义式(6.55)中的 M 进制 QAM 信号为

$$s_k(t) = a_k g(t - kT) \cos(2\pi f_c t) - b_k g(t - kT) \sin(2\pi f_c t), \quad 0 \leq t \leq T \\ k = 0, \pm 1, \pm 2, \dots \quad (6.66)$$

假设载波频率 f_c 与符号速率 $1/T$ 无关。根据式(6.66),可将发送的 M 进制 QAM 信号 $s(t)$ 表示为符号的无穷级数,即

$$\begin{aligned}s(t) &= \sum_{k=-\infty}^{\infty} s_k(t) \\&= \sum_{k=-\infty}^{\infty} [a_k g(t - kT) \cos(2\pi f_c t) - b_k g(t - kT) \sin(2\pi f_c t)]\end{aligned}\quad (6.67)$$

这个公式表明,对于任意的 f_c ,通带函数 $g(t - kT) \cos(2\pi f_c t)$ 和 $g(t - kT) \sin(2\pi f_c t)$ 都是非周期的,并随着符号的变化而变化。

如何才能消除这些通带基本函数随着符号变化而发生的时间上的改变呢?为了回答这个问题,可将对应的公式由实数形式转为复数形式。特别地,将式(6.67)写为下面的等价形式:

$$\begin{aligned}s(t) &= \operatorname{Re} \left\{ \sum_{k=-\infty}^{\infty} (a_k + jb_k) g(t - kT) \exp(j2\pi f_c t) \right\} \\&= \operatorname{Re} \left\{ \sum_{k=-\infty}^{\infty} A_k g(t - kT) \exp(j2\pi f_c t) \right\}\end{aligned}\quad (6.68)$$

其中, A_k 为由下式定义的复数:

$$A_k = a_k + jb_k \quad (6.69)$$

$\operatorname{Re}\{\cdot\}$ 表示取括号内复数的实部。显然,如果将式(6.68)中的被加数乘以模为 1 的复指数 $\exp(-j2\pi f_c kT)$ 与其复共轭 $\exp(j2\pi f_c kT)$ 之积,该公式不变。于是式(6.68)可以重新写为下式:

$$\begin{aligned}s(t) &= \operatorname{Re} \left\{ \sum_{k=-\infty}^{\infty} A_k g(t - kT) \exp(j2\pi f_c t) \exp(-j2\pi f_c kT) \exp(j2\pi f_c kT) \right\} \\&= \operatorname{Re} \left\{ \sum_{k=-\infty}^{\infty} A_k \exp(j2\pi f_c kT) g(t - kT) \exp(j2\pi f_c (t - kT)) \right\}\end{aligned}\quad (6.70)$$

定义

$$\tilde{A}_k = A_k \exp(j2\pi f_c kT) \quad (6.71)$$

$$g_+(t) = g(t) \exp(j2\pi f_c t) \quad (6.72)$$

比例因子 \tilde{A}_k 只是 (ϕ_1, ϕ_2) 平面内第 k 个发送符号坐标的复数表达形式的旋转。在对复值信号 $g_+(t)$ 做出解释之前,先假设脉冲波形函数 $g(t)$ 为一个高频分量小于载波频率 f_c 的低通信号。这样,根据附录 2,可将 $g_+(t)$ 看成是带通信号 $g(t) \cos(2\pi f_c t)$ 的解析信号或预包络。为明确起见,将 $g_+(t)$ 展开如下:

$$\begin{aligned}g_+(t) &= g(t) \cos(2\pi f_c t) + jg(t) \sin(2\pi f_c t) \\&= p(t) + j\hat{p}(t)\end{aligned}\quad (6.73)$$

其中, $p(t)$ 和 $\hat{p}(t)$ 的定义为

$$p(t) = g(t) \cos(2\pi f_c t) \quad (6.74)$$

和

$$\hat{p}(t) = g(t) \sin(2\pi f_c t) \quad (6.75)$$

可以说,解析信号 $g_+(t)$ 的正交(虚)分量 $\hat{p}(t)$ 是其同相(实)分量的希尔伯特变换。不难发现,尽管脉冲波形函数 $g(t)$ 是基带函数,但相应解析信号 $g_+(t)$ 的同相分量和正交分量都是通带函

数。因此, $g(t)$ 被称为基带脉冲, 而 $p(t)$ 和 $\hat{p}(t)$ 则分别被称为通带同相和通带正交脉冲。

利用以上对 \tilde{A}_k 和 $g_+(t)$ 的定义, 可将式(6.70)的发射信号简单地重新定义如下:

$$s(t) = \operatorname{Re} \left\{ \sum_{k=-\infty}^{\infty} \tilde{A}_k g_+(t - kT) \right\} \quad (6.76)$$

从上式中可以得到三个值得注意的重要结论:

- 发射信号 $s(t)$ 是无载波的。
- 由于式(6.76)中对 $s(t)$ 的表述实际上等同于式(6.67)中对 M 进制 QAM 信号的表述, $s(t)$ 的新公式完全保留了原始 M 进制 QAM 信号混合幅度相位调制的特点。
- 发射信号 $s(t)$ 代表了该混合调制过程的符号时不变实现。

对于既定的载波频率 f_c , 式(6.71)中涉及到的旋转因子序列是确定的。因此, 接收机仅需检测出 \tilde{A}_k , 然后对其进行下述逆向旋转就可以计算出相应的 A_k 值:

$$A_k = \tilde{A}_k \exp(-j2\pi f_c k T), \quad k = 0, \pm 1, \pm 2, \dots$$

但是, 在实际应用中, 由于旋转因子对混合调制系统的工作或者说性能并没有什么影响, 因此, 它们常被忽略。只有在研究 QAM 的等同性时, 才必须用到旋转因子。因此, 可忽略式(6.71), 而将式(6.76)重新简单定义如下:

$$\begin{aligned} s(t) &= \operatorname{Re} \left\{ \sum_{k=-\infty}^{\infty} A_k g_+(t - kT) \right\} \\ &= \operatorname{Re} \left\{ \sum_{k=-\infty}^{\infty} (a_k + jb_k)(p(t - kT) + j\hat{p}(t - kT)) \right\} \\ &= \sum_{k=-\infty}^{\infty} [a_k p(t - kT) - b_k \hat{p}(t - kT)] \end{aligned} \quad (6.77)$$

其中, 如前所述, $\hat{p}(t)$ 是 $p(t)$ 的希尔伯特变换。式(6.77)中的发射信号就称为无载波幅度/相位调制(CAP)。^[6]

6.4.5 通带同相和正交脉冲的性质

从式(6.74)和式(6.75)对通带同相和正交脉冲的定义中, 可以得到以下性质:

性质 1

假设基带脉冲 $g(t)$ 是时间 t 的偶函数, 则通带同相脉冲 $p(t)$ 和正交脉冲 $\hat{p}(t)$ 分别为时间 t 的偶函数和奇函数。

这一性质可由式(6.74)和式(6.75)直接得到。

性质 2

通带脉冲 $p(t)$ 和 $\hat{p}(t)$ 在整个时间间隔 $(-\infty, \infty)$ 内构成了正交集合, 即

$$\int_{-\infty}^{\infty} p(t)\hat{p}(t) dt = 0 \quad (6.78)$$

在式(6.78)中, 不把积分时间间隔限制在符号时间 T 内的原因是, 对于高性能的 CAP 系统, 脉

冲 $p(t)$ 和 $\hat{p}(t)$ 都具有很高的带宽效率。为了证明性质 2, 利用傅里叶变换将式(6.78)变为如下频域形式:

$$\int_{-\infty}^{\infty} p(t)\hat{p}(t) dt = \int_{-\infty}^{\infty} P(f)\hat{P}^*(f) df \quad (6.79)$$

其中 $p(t) \Leftrightarrow P(f)$ 。在 $\hat{p}(t) \Leftrightarrow \hat{P}(f)$ 中, * 表示复共轭。(频率函数 $\hat{P}^*(f)$ 不应当看成是 $P(f)$ 的希尔伯特变换, 而应当根据傅里叶分析的常用术语, 将其视为 $\hat{p}(t)$ 的傅里叶变换。)傅里叶变换式 $\hat{P}(f)$ 通过下式与傅里叶变换式 $P(f)$ 联系起来(见附录 2):

$$\hat{P}(f) = -j \operatorname{sgn}(f)P(f) \quad (6.80)$$

其中, $\operatorname{sgn}(f)$ 是符号函数。傅里叶变换 $P(f)$ 和 $\hat{P}(f)$ 有相同的幅度谱, 但二者的相位谱在频率为负时相差 $+90^\circ$, 在频率为正时相差 -90° 。因此, 可将式(6.79)写为

$$\begin{aligned} \int_{-\infty}^{\infty} p(t)\hat{p}(t) dt &= j \int_{-\infty}^{\infty} P(f)P^*(f) \operatorname{sgn}(f) df \\ &= j \int_{-\infty}^{\infty} |P(f)|^2 \operatorname{sgn}(f) df \end{aligned} \quad (6.81)$$

注意, 幅度响应 $|P(f)|$ 是频率 f 的偶函数, 而符号函数 $\operatorname{sgn}(f)$ 则是频率 f 的奇函数, 因此式(6.81)的积分值为 0, 性质 2 得证。

接着, 假设通带脉冲 $p(t)$ 和 $\hat{p}(t)$ 通过一个冲激响应函数为 $h(t)$ 的线性时不变信道, 得到下列通带输出:

$$u(t) = p(t) \star h(t) \quad (6.82)$$

和

$$\hat{u}(t) = \hat{p}(t) \star h(t) \quad (6.83)$$

其中 \star 表示卷积。

性质 3

由式(6.82)和式(6.83), 通带脉冲 $u(t)$ 和 $\hat{u}(t)$ 构成了希尔伯特变换对, 因此对于任意的 $h(t)$, $u(t)$ 和 $\hat{u}(t)$ 在整个时间间隔 $(-\infty, \infty)$ 内是正交的。

正是这一重要性质, 使得 CAP 接收机在给定信道输出时, 能够区分发送的实符号 a_k 和虚符号 b_k 。为了证明性质 3, 仍然利用傅里叶变换得到

$$\begin{aligned} \int_{-\infty}^{\infty} u(t)\hat{u}(t) dt &= \int_{-\infty}^{\infty} U(f)\hat{U}^*(f) df \\ &= \int_{-\infty}^{\infty} (P(f)H(f))(-j \operatorname{sgn}(f)P(f)H(f))^* df \\ &= \int_{-\infty}^{\infty} j|P(f)|^2|H(f)|^2 \operatorname{sgn}(f) df \\ &= 0 \end{aligned}$$

例 6.4 提高带宽效率的频谱整形

式(6.77)中的 CAP 信号利用复数因子 A_k 进行了多级编码, 从而具有较高的频谱效率。可通过采用基带脉冲 $g(t)$ 的有效频谱形式进一步提高 CAP 的带宽效率。在选择 $g(t)$

时,可利用第4章中讨论的升余弦脉冲波形函数。

例如,考虑一个滚降因子 $\alpha=0.2$ 的基带升余弦波形滤波器(界外带宽占20%)。基带脉冲 $g(t)$ 的公式为^①

$$g(t) = \langle \text{sinc}(t) \rangle \frac{\cos(\pi\alpha t)}{1 - 4\alpha^2 t^2} \quad (6.84)$$

其中,时间 t 已经按照符号持续时间 T 进行了归一化。因此,将 $\alpha=0.2$ 代入式(6.84),可得到图6.19所示的结果。利用式(6.74)和式(6.75),可计算出如图6.20(a)和6.20(b)所示的通常同相脉冲 $p(t)$ 和正交脉冲 $\hat{p}(t)$,此时归一化载波频率 $f_c T = 0.5(1 + \alpha) = 0.6$ 。图6.20的波形表明 $p(t)$ 和 $\hat{p}(t)$ 分别是时间 t 的偶函数和奇函数,这和性质1是一致的;而它们在时间间隔 $(-\infty, \infty)$ 内正交,这则与性质2一致。

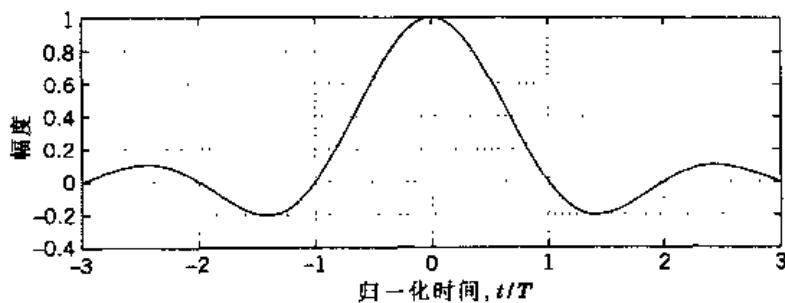


图6.19 滚降因子 $\alpha=0.2$ 的基带脉冲 $g(t)$

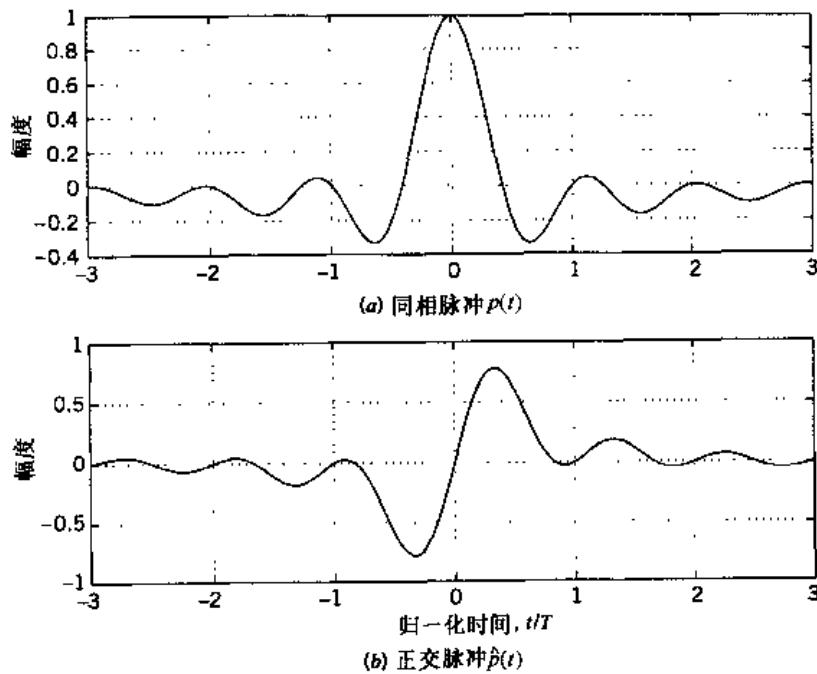


图6.20 滚降因子 $\alpha=0.2$ 的同相脉冲和正交脉冲

^① 式(6.84)通过在第4章式(4.62)中用归一化时间 t 代替 $2\pi t$ 而得到。需注意,奈奎斯特带宽 W 等于符号(比特)速率的0.5倍。同时注意不要将第4章中的 $p(t)$ 和本章中的 $p(t)$ 混为一谈。

6.4.6 CAP 发射机基本结构

图 6.21 是根据式(6.74)、式(6.75)和式(6.77)绘制的 CAP 发射机的功能框图。发射机由一个多级编码器和一对带通滤波器组成，多级编码器将输入的连续数据流分割成由 n 比特组成的连续的数据块。这些数据块随即被映射为多级符号 a_k 和 b_k ，其中 k 表示第 k 个符号周期。通带同相和正交滤波器并行地分别处理符号流 $\{a_k\}$ 和 $\{b_k\}$ 。对于选定的基带脉冲 $g(t)$ ，这两个滤波器的冲激响应 $p(t)$ 和 $\hat{p}(t)$ 根据式(6.74)和式(6.75)进行设计。将这两个滤波器的输出相减，就得到了式(6.77)中的发射 CAP 信号 $s(t)$ 。

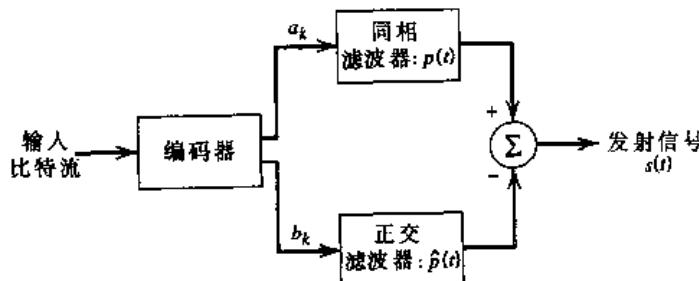


图 6.21 CAP 发射机框图

发射信号 $s(t)$ 通过冲激响应为 $h(t)$ 、附加噪声为 $w(t)$ 的信道传播，产生的信道输出为

$$\begin{aligned}
 x(t) &= s(t) \star h(t) + w(t) \\
 &= \sum_{k=-\infty}^{\infty} [a_k p(t - kT) \star h(t) - b_k \hat{p}(t - kT) \star h(t)] + w(t) \\
 &= \sum_{k=-\infty}^{\infty} [a_k u(t - kT) - b_k \hat{u}(t - kT)] + w(t)
 \end{aligned} \tag{6.85}$$

其中 $u(t)$ 和 $\hat{u}(t)$ 分别按照式(6.82)和式(6.83)进行定义。假设输入信号为 $x(t)$ ，接收机的功能就是以一种最佳的方式一个接一个地恢复得到的符号 a_k 和 b_k 。

此处选用的是一个二维最佳接收机，其具体情况取决于信道的干扰。特别地，需要提到两种不同的情况：

- 仅存在加性高斯白噪声的干扰。在这种理想情况下，最佳 CAP 接收机为如图 6.22 所示的二维匹配滤波器。最佳同相滤波器是与通带同相脉冲 $p(t)$ 相匹配的滤波器，而最佳正交滤波器则是与通带正交脉冲 $\hat{p}(t)$ 相匹配的滤波器。
- 信道干扰包括 ISI 和加性高斯白噪声。在这种更为现实的情况下，最佳 CAP 接收机根据 4.9 节介绍的最佳线性接收机理论设计。特别地，CAP 接收机由一个二维匹配滤波器和一对相同的均衡器，以及一对同步抽样器组成，如图 6.23 所示。均衡器（即抽头延时线）补偿了信道弥散的影响。

从性质 3 可见，通带正交脉冲通过同相匹配滤波器的输出是 0，反之亦然。因此，在这两种情况下，发送符号 a_k 和 b_k 能够被二维最佳 CAP 接收机分别检测出来。

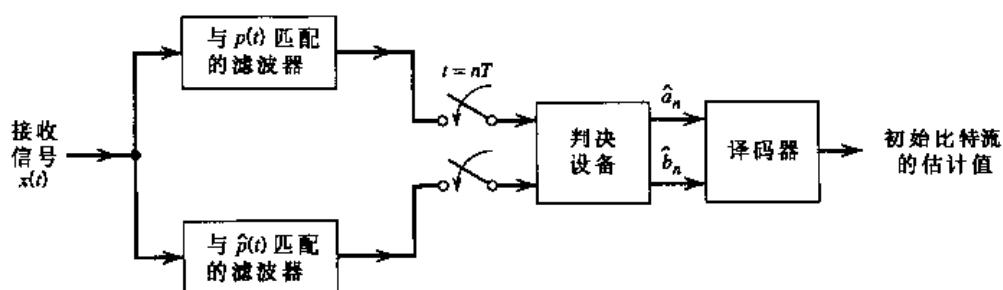


图 6.22 理想高斯白噪声信道情况下, 使用二维匹配滤波器的 CAP 接收机框图

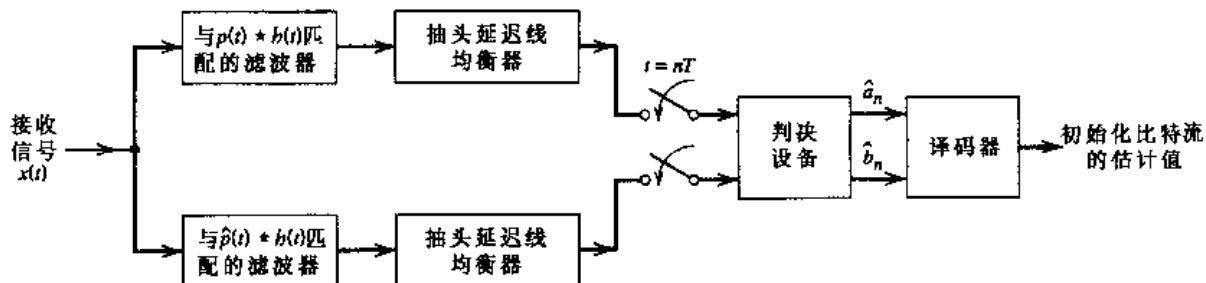


图 6.23 噪声弥散信道下使用一对最佳线性接收机的 CAP 接收机框图

6.4.7 CAP 接收机的数字实现

图 6.23 的接收机结构可进一步简化。首先将该结构重新表示为图 6.24(a), 其中在接收机输入端引入了 A/D(模 - 数)转换器。由于图 6.24(a)中的匹配滤波器和均衡器都是线性系统, 因此, 图 6.24(a)中 A、B 点的抽样值是输入抽样值 $x[n] = x(nT_s)$ 的线性组合, 其中 T_s 为抽样周期。因此, 可将图 6.24(a)所示的各条接收路径的匹配滤波器和均衡器的组合都用一个速率为 $1/T_s$ 的有限冲激响应滤波器(FIR)来代替, 如图 6.24(b)所示。在同输入的情况下, 图 6.24(b)中的这两个 FIR 滤波器共用一组位延迟, 但它们各有一组系数。同时, 不难注意到, 图 6.24(b)所示的接收机结构是图 6.21 发射机结构的镜像。

最后还要注意一点。在实际应用中, 图 6.24(b)所示接收机中的两个 FIR 滤波器都是自适应的, 因此可以适应 CAP 接收机在未知环境中正常工作的需要。由第 4 章可知, 在选择自适应均衡(滤波)时有两个选择: 线性均衡和判决反馈均衡(DFE)。当信道的频率响应近似平坦时, 使用线性均衡就已经足够了。但当信道的频率响应和/或噪声功率谱密度不近似平坦时, 使用 DFE 能够在很大程度上提高 CAP 接收机的性能。

6.4.8 CAP 的一种应用

CAP(使用 32 点或 64 点星座图)的一个重要应用就是通过长度小于 100 m 的双绞线进行数字数据的通带传输, 如在虚拟分配系统的局域网(LAN)中的应用。(所有的现代局域网标准都将传输线长限制在 100 m 以内。如果有必要的话, CAP 系统实际可以用于更长的回路。)当可用信道带宽严格限制在 30 MHz 以内时, 数据速率范围为从 51 Mbps 到 155 Mbps。

利用双绞线电缆的双工无线电收发机的主要干扰是传播损耗和近端串话。这些干扰在

4.8节中介绍 DSL 时已经讨论过。

在这里描述的环境中,一个实际问题就是如何使接收机适应双绞线电缆的广泛变化。为了适应这一环境,在接收机中选用自适应滤波器来实现最佳同相和正交滤波器,这在前面已经提到了。双绞线电缆的变化范围越大,这些自适应滤波器也就越复杂。

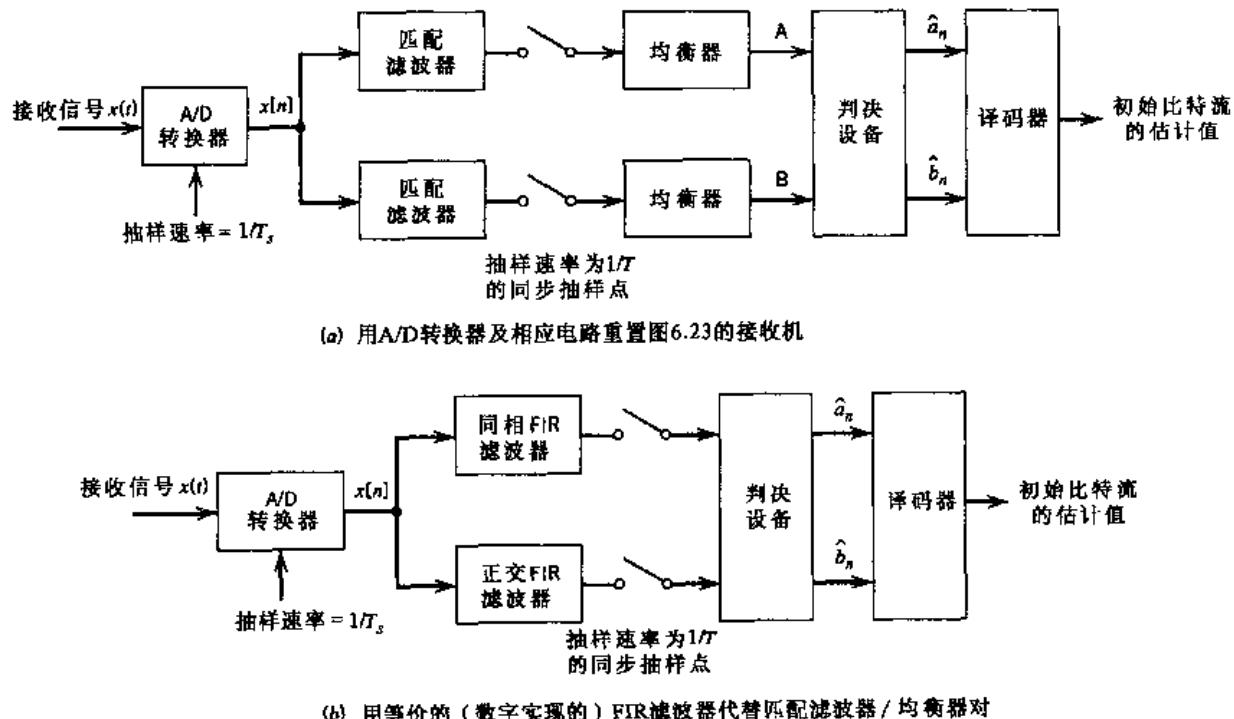


图 6.24 CAP 接收机的数字实现

6.5 相干 FSK

M 进制 PSK 和 M 进制 QAM 都是线性调制的例子。本节研究一种非线性通带数据传输方式,即相干 FSK。首先考虑二进制 FSK 的简单情况。

6.5.1 二进制 FSK

在二进制 FSK 系统中,通过发送两个有固定频率差的正弦信号之一来区分符号 1 和 0。一对典型的正弦信号如下式描述:

$$s_i(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_i t), & 0 \leq t \leq T_b \\ 0, & \text{其他} \end{cases} \quad (6.86)$$

其中, $i = 1, 2$; E_b 为每比特的发射信号能量。发射频率为

$$f_i = \frac{n_c + i}{T_b} \quad n_c \text{ 为固定整数, } i = 1, 2 \quad (6.87)$$

于是, 符号 1 用 $s_1(t)$ 表示, 而符号 0 用 $s_2(t)$ 表示。这里介绍的 FSK 信号被称为桑德(Sunde's)FSK。由于这种信号即使在位变换期间也总是保持相位的连续, 因此是一种连续相位信号。这种形式的数字调制是连续相位 FSK(CPFSK)的一个例子, 这一节的后面将对此进行详述。

由式(6.86)和式(6.87)可知, 信号 $s_1(t)$ 和 $s_2(t)$ 是正交的, 但并没有归一化到单位能量。由此得出这组正交基本函数最有用的形式为

$$\phi_i(t) = \begin{cases} \sqrt{\frac{2}{T_b}} \cos(2\pi f_i t), & 0 \leq t \leq T_b \\ 0, & \text{其他} \end{cases} \quad (6.88)$$

其中, $i = 1, 2$ 。因此, 系数 s_{ij} ($i = 1, 2; j = 1, 2$) 的定义为

$$\begin{aligned} s_{ij} &= \int_0^{T_b} s_i(t) \phi_j(t) dt \\ &= \int_0^{T_b} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_i t) \sqrt{\frac{2}{T_b}} \cos(2\pi f_j t) dt \\ &= \begin{cases} \sqrt{E_b}, & i = j \\ 0, & i \neq j \end{cases} \end{aligned} \quad (6.89)$$

与相干二进制 PSK 不同的是, 相干二进制 FSK 系统的特征是具有由两个信息点($M = 2$)组成的如图 6.25 所示的二维($N = 2$)信号空间。两个信息点的定义为

$$\mathbf{s}_1 = \begin{bmatrix} \sqrt{E_b} \\ 0 \end{bmatrix} \quad (6.90)$$

和

$$\mathbf{s}_2 = \begin{bmatrix} 0 \\ \sqrt{E_b} \end{bmatrix} \quad (6.91)$$

这两个信息点之间的欧氏距离等于 $\sqrt{2E_b}$ 。图 6.25 中还有两幅插图, 分别为信号 $s_1(t)$ 和 $s_2(t)$ 的波形。

6.5.2 二进制 FSK 的差错概率

观察矢量 \mathbf{x} 有两个部分 x_1 和 x_2 , 其定义分别为

$$x_1 = \int_0^{T_b} x(t) \phi_1(t) dt \quad (6.92)$$

和

$$x_2 = \int_0^{T_b} x(t) \phi_2(t) dt \quad (6.93)$$

其中 $x(t)$ 为接收信号, 其形式取决于发送的符号。假设符号 1 被发送, $x(t)$ 就等于 $s_1(t) + w(t)$, 其中 $w(t)$ 是具有零均值且功率谱密度为 $N_0/2$ 的高斯白噪声的样本函数。如果发送的是符号 0, 则 $x(t)$ 等于 $s_2(t) + w(t)$ 。

应用式(5.59)的判决准则, 可将观测空间分为两个判决区域, 分别在图 6.25 中标注为 Z_1

和 Z_2 。分割 Z_1 和 Z_2 的判决边界, 是两个信息点连线的垂直平分线。如果由观察矢量 \mathbf{x} 表示的接收信号点落在区域 Z_1 内, 则接收机判决其为符号 1, 当 $x_1 > x_2$ 时, 就属于上述情况; 反之, 当 $x_1 < x_2$ 时, 接收信号点落在区域 Z_2 内, 因而接收机将其判决为符号 0。在判决边界, 有 $x_1 = x_2$, 此时接收机随机猜测发送的是符号 1 还是符号 0。

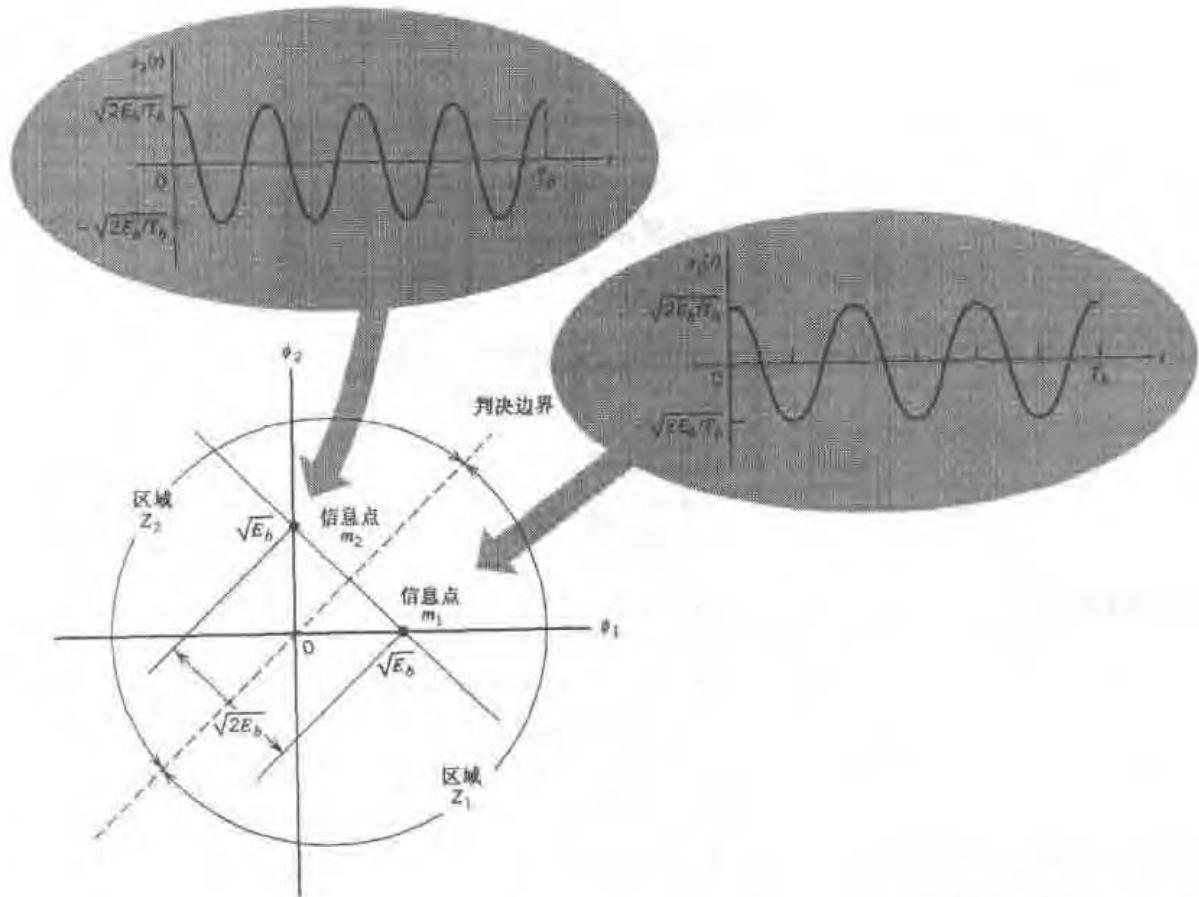


图 6.25 二进制 FSK 系统的信号空间图, 该图还包括两个调制信号 $s_1(t)$ 和 $s_2(t)$ 的波形

定义一个新的高斯随机变量 Y , 其抽样值 y 等于 x_1 和 x_2 的差值, 即

$$y = x_1 - x_2 \quad (6.94)$$

随机变量 Y 的平均值取决于发送的二进制符号。假设符号 1 被发送, 则高斯随机变量 X_1 和 X_2 (抽样值分别为 x_1 和 x_2) 平均值分别为 $\sqrt{E_b}$ 和 0。因此, 发送符号 1 时, 随机变量 Y 的条件平均值为

$$\begin{aligned} E[Y|1] &= E[X_1|1] - E[X_2|1] \\ &= +\sqrt{E_b} \end{aligned} \quad (6.95)$$

反之, 如果符号 0 被发送, 则随机变量 X_1 和 X_2 的平均值分别为 0 和 $\sqrt{E_b}$ 。因此, 发送符号 0 时, 随机变量 Y 的条件平均值为

$$E[Y|0] = E[X_1|0] - E[X_2|0] = -\sqrt{E_b} \quad (6.96)$$

随机变量 Y 的变化与发送的二进制符号无关。由于随机变量 X_1 和 X_2 是统计独立的, 它们的变化范围都是 $N_0/2$, 故有

$$\text{var}[Y] = \text{var}[X_1] + \text{var}[X_2] = N_0 \quad (6.97)$$

假设已知发送的是符号 0, 则随机变量 Y 的条件概率密度函数如下:

$$f_Y(y|0) = \frac{1}{\sqrt{2\pi N_0}} \exp\left[-\frac{(y + \sqrt{E_b})^2}{2N_0}\right] \quad (6.98)$$

由于条件 $x_1 > x_2$, 或者说 $y > 0$, 将使接收机判决发送符号为 1, 因此得到发送符号 0 时的条件差错概率为

$$\begin{aligned} p_{10} &= P(y > 0 | \text{发送符号 } 0) \\ &= \int_0^\infty f_Y(y|0) dy \\ &= \frac{1}{\sqrt{2\pi N_0}} \int_0^\infty \exp\left[-\frac{(y + \sqrt{E_b})^2}{2N_0}\right] dy \end{aligned} \quad (6.99)$$

令

$$\frac{y + \sqrt{E_b}}{\sqrt{2N_0}} = z \quad (6.100)$$

接着, 将积分变量 y 改为 z , 可将式(6.99)重新写为

$$\begin{aligned} p_{10} &= \frac{1}{\sqrt{\pi}} \int_{\sqrt{E_b/2N_0}}^\infty \exp(-z^2) dz \\ &= \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right) \end{aligned} \quad (6.101)$$

类似地, 可证明 p_{01} , 即发送符号 1 时的条件差错概率, 与式(6.101)的结果是一致的。据此, 将 p_{10} 和 p_{01} 进行平均, 可得平均误比特率或相干二进制 FSK 的误比特率(假设符号等概地发送)为

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{2N_0}}\right) \quad (6.102)$$

比较式(6.20)和式(6.102)可知, 在相干二进制 FSK 系统中, 为了保持与相干二进制 PSK 系统相同的误比特率, 必须提供两倍于 PSK 系统的比特信噪比 E_b/N_0 。这一结果与图 6.3 及图 6.25 的信号空间图是完全一致的。由图可见, 二进制 PSK 系统两个信息点之间的欧氏距离等于 $2\sqrt{E_b}$, 而二进制 FSK 系统两个信息点之间的欧氏距离则是 $\sqrt{2E_b}$ 。因此, 对于给定的 E_b , 二进制 PSK 的最小距离 d_{\min} 是二进制 FSK 相应距离的 $\sqrt{2}$ 倍。第 5 章中曾经提到过, 差错概率随着 d_{\min}^2 指数级衰减, 因此造成了式(6.20)和式(6.102)之间的差别。

6.5.3 相干二进制 FSK 信号的产生和检测

可采用图 6.26(a)所示的方法来产生二进制 FSK 信号。输入的二进制数据序列首先被送入开关电平编码器, 编码器输出端将用恒定的 $\sqrt{E_b}$ 伏电平表示符号 1, 而用恒定的 0 伏电平表

示符号 0。在图 6.26(a)中的下方通道中采用反相器,就能够确保输入符号 1 时,上方通道中频率为 f_1 的振荡器打开,而下方通道中频率为 f_2 的振荡器关闭,其结果就是发射频率为 f_1 。反之,当输入符号 0 时,上方通道的振荡器关闭,而下方通道的振荡器打开,其结果就是发送频率为 f_2 。如式(6.87)所示,这两个频率 f_1 和 f_2 等于比特速率 $1/T_b$ 的不同整数倍。

在图 6.26(a)所示的发射机中,假设两个振荡器是同步的,其输出满足式(6.88)中对两个正交基本函数 $\phi_1(t)$ 和 $\phi_2(t)$ 的要求。也可用单键值(压控)振荡器来代替这两个振荡器。在任一情况下,调制波形的频率和输入二进制波形都应当一致,以确保相位变化是连续的。

为了在接收到的噪声干扰信号 $x(t)$ 中检测出原始二进制序列,可采用图 6.26(b)所示的接收机。该接收机由两个共输入的相关器组成,两个相关器的输入分别为本地参考信号 $\phi_1(t)$ 和 $\phi_2(t)$ 。将相关器的两个输出相减,得到的差值 y 将与 0 V 的门限进行比较。如果 $y > 0$,接收机判决为符号 1;反之判决为符号 0。如果 y 恰好等于 0,接收机将随机判决为符号 1 或符号 0。

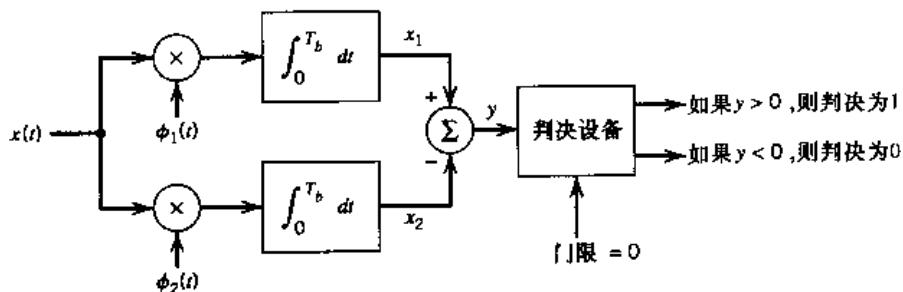
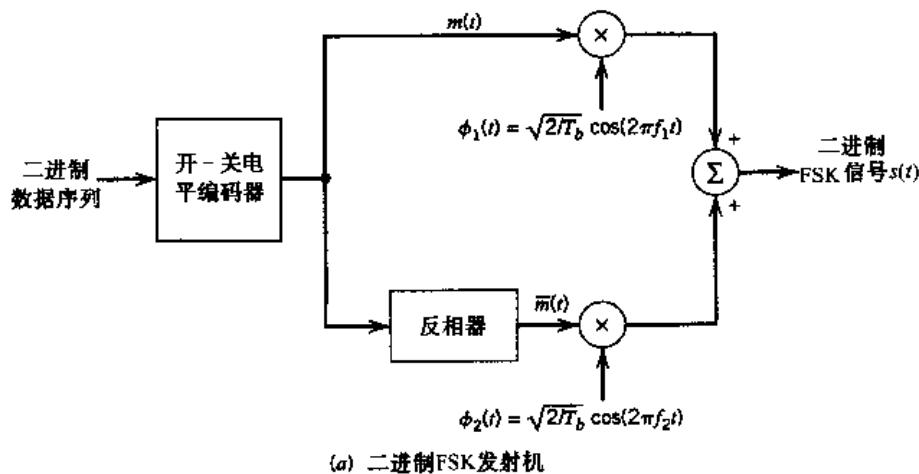


图 6.26 二进制 FSK 发射机和接收机框图

6.5.4 二进制 FSK 信号的功率谱

在桑德 FSK 的情况下,两个发射频率 f_1 和 f_2 的差值等于比特速率 $1/T_b$, 算术平均值等于载波频率 f_c 。这样,即使在位变换期间也能保持相位的连续性。可将这种特殊的二进制 FSK

信号表示为：

$$s(t) = \sqrt{\frac{2E_b}{T_b}} \cos\left(2\pi f_c t \pm \frac{\pi t}{T_b}\right), \quad 0 \leq t \leq T_b \quad (6.103)$$

利用三角恒等式，可得

$$\begin{aligned} s(t) &= \sqrt{\frac{2E_b}{T_b}} \cos\left(\pm \frac{\pi t}{T_b}\right) \cos(2\pi f_c t) - \sqrt{\frac{2E_b}{T_b}} \sin\left(\pm \frac{\pi t}{T_b}\right) \sin(2\pi f_c t) \\ &= \sqrt{\frac{2E_b}{T_b}} \cos\left(\frac{\pi t}{T_b}\right) \cos(2\pi f_c t) \mp \sqrt{\frac{2E_b}{T_b}} \sin\left(\frac{\pi t}{T_b}\right) \sin(2\pi f_c t) \end{aligned} \quad (6.104)$$

在式(6.104)的最后一行，加号对应于发送符号 0，而减号对应于发送符号 1。和前面一样，假设调制器输入端的随机二进制波形中，符号 1 和 0 等概地出现，并假设相邻时间间隔内发送的符号是统计独立的。根据式(6.104)，对于连续相位二进制 FSK 信号的同相分量和正交分量，有：

1. 同相分量与输入的二进制波形完全无关。在任意时刻 t ，同相分量都等于 $\sqrt{2E_b/E_b} \cos(\pi t/T_b)$ 。因此，该分量的功率谱密度由两个冲激函数组成。两个冲激函数的加权因子为 $E_b/2T_b$ ，分别位于 $f = \pm 1/2T_b$ 处。
2. 正交分量与输入的二进制波形直接关联。在信号时间间隔 $0 \leq t \leq T_b$ 内，当发送符号为 1 时，正交分量等于 $-g(t)$ ；而当发送符号为 0 时，则等于 $+g(t)$ 。符号波形函数 $g(t)$ 的定义如下：

$$g(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \sin\left(\frac{\pi t}{T_b}\right), & 0 \leq t \leq T_b \\ 0, & \text{其他} \end{cases} \quad (6.105)$$

符号波形函数的能量谱密度为

$$\Psi_g(f) = \frac{8E_b T_b \cos^2(\pi T_b f)}{\pi^2 (4T_b^2 f^2 - 1)^2} \quad (6.106)$$

正交分量的能量谱密度等于 $\Psi_g(f)/T_b$ 。显然，二进制 FSK 信号的同相分量和正交分量彼此独立。因此，桑德 FSK 信号的基带功率谱密度等于这两个分量的功率谱密度之和，即

$$S_B(f) = \frac{E_b}{2T_b} \left[\delta\left(f - \frac{1}{2T_b}\right) + \delta\left(f + \frac{1}{2T_b}\right) \right] + \frac{8E_b \cos^2(\pi T_b f)}{\pi^2 (4T_b^2 f^2 - 1)^2} \quad (6.107)$$

将式(6.107)代入式(6.4)，可见二进制 FSK 信号的功率谱包含两个离散的频率分量。这两个频率分量分别位于 $(f_c + 1/2T_b) = f_1$ 和 $(f_c - 1/2T_b) = f_2$ 处，其平均值等于二进制 FSK 信号总功率的一半。这两个离散频率分量的存在，为接收机提供了一种与发射机保持同步的方法。

不难注意到，连续相位二进制 FSK 信号的基带功率谱密度随着频率值的四次方的倒数不断衰减。在 f 接近于无穷大时，对式(6.107)取极限，也可得到上述结论。但是，如果 FSK 信号在位变换期间不具有相位连续性（当使用频率分别为 f_1 和 f_2 的两个振荡器，且振荡器彼此工

作不独立时就会产生这种现象), 功率谱密度将随着频率值平方的倒数不断衰减, 见习题 6.23。因此, 与相位不连续的 FSK 信号不同的是, 连续相位 FSK 信号在信号带外并不会受到太大的影响。

图 6.5 为式(6.22)和式(6.107)对应的基带功率谱。(为了简化问题, 仅给出了正频率的结果。)在两种情况下, $S_B(f)$ 都按照 $2E_b$ 进行了归一化, 而频率都按照比特速率 $R_b = 1/T_b$ 进行了归一化。这两种谱衰减速率的差别可以用脉冲波形 $g(t)$ 来解释。脉冲越平滑, 谱线衰减到 0 的速度越快。由于二进制 FSK(连续相位)的脉冲波形较为平滑, 其旁瓣也就相应地比二进制 PSK 小。

6.5.5 MSK

在对二进制 FSK 信号进行相干检测时, 包含在接收信号中的相位信息并没有得到充分利用, 而仅用于实现接收机与发射机之间的同步。下面将证明, 通过在检测时适当地利用该相位信息, 可以显著提高接收机的噪声性能。不过, 这种性能的提高, 是以增加接收机复杂性为代价换来的。

考虑一个定义在 $0 \leq t \leq T_b$ 时间间隔内的连续相位 FSK(CPFSK)信号:

$$s(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_1 t + \theta(0)] & \text{对于符号 1} \\ \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_2 t + \theta(0)] & \text{对于符号 0} \end{cases} \quad (6.108)$$

其中, E_b 为每比特的发射信号能量, T_b 为比特持续时间。相位 $\theta(0)$ 表示在时刻 $t = 0$ 时的相位值, 该相位值等于直到 $t = 0$ 时刻的所有调制过程引起的相位变化的总和。频率 f_1 和 f_2 分别对应于输入调制器的符号 1 和 0。

另一种有效表示 CPFSK 信号 $s(t)$ 的方法是用下述角度调制信号的通式:

$$s(t) = \sqrt{\frac{2E_b}{T_b}} \cos[2\pi f_c t + \theta(t)] \quad (6.109)$$

其中, $\theta(t)$ 为 $s(t)$ 的相位。当相位 $\theta(t)$ 为连续时间函数时, 调制信号 $s(t)$ 本身(即使在位变换期间)也是时间连续的。在每个比特持续时间 T_b 秒内, CPFSK 信号的相位 $\theta(t)$ 随着时间线性地增大或减小, 即

$$\theta(t) = \theta(0) \pm \frac{\pi b}{T_b} t, \quad 0 \leq t \leq T_b \quad (6.110)$$

其中“+”对应于发送符号 1, “-”对应于发送符号 0。下面对参数 b 进行定义。将式(6.110)代入式(6.109), 然后将余弦函数的角度与式(6.108)进行比较, 得到下面的关系对:

$$f_c + \frac{b}{2T_b} = f_1 \quad (6.111)$$

$$f_c - \frac{b}{2T_b} = f_2 \quad (6.112)$$

由式(6.111)和式(6.112)可解得 f_c 和 b :

$$f_c = \frac{1}{2} (f_1 + f_2) \quad (6.113)$$

和

$$b = T_b (f_1 - f_2) \quad (6.114)$$

标称载波频率 f_c 为频率 f_1 和 f_2 的算术平均值。频率 f_1 和 f_2 之间的差值, 按照比特速率 $1/T_b$ 归一化之后, 定义了无量纲参数 b , 参数 b 称为偏移率。

6.5.6 相位网格图

由式(6.110)可见, 在时刻 $t = T_b$, 有

$$\theta(T_b) - \theta(0) = \begin{cases} \pi b & \text{对于符号 1} \\ -\pi b & \text{对于符号 0} \end{cases} \quad (6.115)$$

也就是说, 符号 1 的发送将使 CPFSK 信号 $s(t)$ 的相位增加 πb 弧度, 而符号 0 的发送将使其减小 πb 弧度。

相位 $\theta(t)$ 随时间 t 的变化沿着一系列直线构成的路径进行, 路径的斜率代表了频率的变化。图 6.27 描绘了从时刻 $t = 0$ 起的全部可能路径。图 6.27 称为相位树图。树图的引入使我们可以清楚地看出输入数据序列时间间隔边界上的相位变换。更进一步, 从图 6.27 可以清楚看到, 对应于比特持续时间 T_b 的奇数或偶数倍的时刻, CPFSK 信号的相位分别为 πb 弧度的奇数或偶数倍。

图 6.27 中的相位树图明显具有相位连续性, 这也正是 CPFSK 信号的内在特性。为了更好地理解相位连续性的概念, 我们再回顾一下前面介绍过的桑德 FSK 这种特殊的 CPFSK 方式。此时, 偏移率 b 恰好等于 1, 因此, 根据图 6.27, 一个比特持续时间内的相位变化为 $\pm \pi$ 弧度。但是, 由于模为 2π , $+\pi$ 弧度的相位变化与 $-\pi$ 弧度的相位变化实质上是一样的。由此得出结论: 桑德 FSK 是无记忆性的。也就是说, 前一比特持续时间内发生的特定变化, 对当前比特持续时间内的信号是没有影响的。

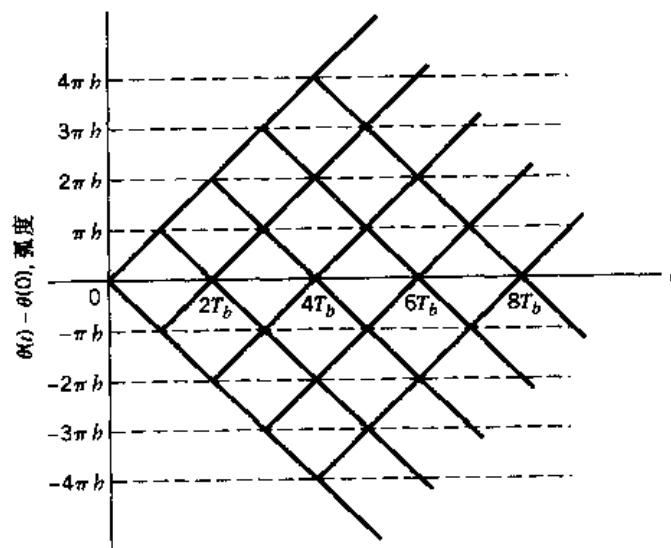


图 6.27 相位树图

与之对照,当偏移率 b 取特定值 $1/2$ 时,情况则完全不同。从图6.28中可以看到,在 T_b 的奇数倍时刻,相位值只能取 $\pm\pi/2$ 。在 T_b 的偶数倍时刻,相位值只能取 0 和 π 。图6.28称为相位网格图。“网格”是一种带有重新合并的分支的树型结构。图6.28所示网格里每一条从左到右的路径都对应于一个特定的输入二进制序列。例如,图6.28中的粗折线表示的路径对应于 $\theta(0) = 0$ 、输入二进制序列为1101000时的情况。此后,我们都假设 $b = 1/2$ 。

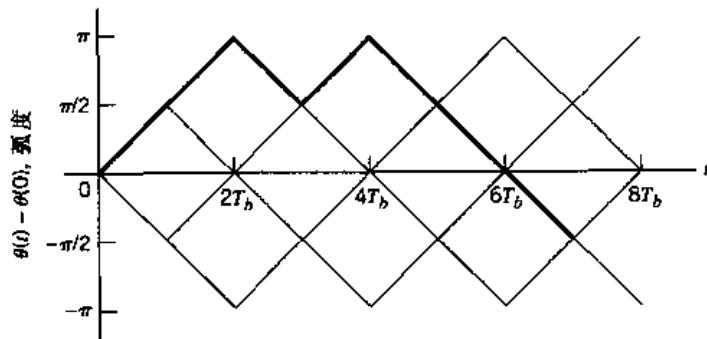


图6.28 相位网格图,粗折线所示路径代表序列1101000

当 $b = 1/2$ 时,由式(6.114)可知,频率偏移(两个信号频率 f_1 和 f_2 之间的差值)等于比特速率的一半。这个频率偏移就是当代表符号1和0的两个FSK信号为如式(6.108)所定义的那样相干正交时,在检测过程中互不干扰所允许的最小频率间隔。因此偏移率为 $1/2$ 的CPFSK信号通常称为最小频移键控(MSK)。^[7]

6.5.7 MSK的信号空间图

在式(6.109)中使用三角恒等式,可将CPFSK信号 $s(t)$ 用其同相分量和正交分量表示如下:

$$s(t) = \sqrt{\frac{2E_b}{T_b}} \cos[\theta(t)] \cos(2\pi f_c t) - \sqrt{\frac{2E_b}{T_b}} \sin[\theta(t)] \sin(2\pi f_c t) \quad (6.116)$$

首先考虑同相分量 $\sqrt{2E_b/T_b} \cos[\theta(t)]$ 。当偏移率 $b = 1/2$ 时,由式(6.110)可得

$$\theta(t) = \theta(0) \pm \frac{\pi}{2T_b} t, \quad 0 \leq t \leq T_b \quad (6.117)$$

其中“+”对应于符号1,“-”对应于符号0。在时间间隔 $-T_b \leq t \leq 0$ 内, $\theta(t)$ 也有相似的取值,只是在这两个时间间隔内,式中的符号不尽相同。经过调制,相位 $\theta(0)$ 的取值相应地为0或 π 。不难发现,在时间间隔 $-T_b \leq t \leq T_b$ 内, $\cos[\theta(t)]$ 的极性仅取决于 $\theta(0)$,而与时刻 $t=0$ 前后发送的0,1序列无关。因此,时间间隔 $-T_b \leq t \leq T_b$ 内,同相分量由下式定义的半周期余弦脉冲组成:

$$\begin{aligned} s_I(t) &= \sqrt{\frac{2E_b}{T_b}} \cos[\theta(t)] \\ &= \sqrt{\frac{2E_b}{T_b}} \cos[\theta(0)] \cos\left(\frac{\pi}{2T_b} t\right) \\ &= \pm \sqrt{\frac{2E_b}{T_b}} \cos\left(\frac{\pi}{2T_b} t\right), \quad -T_b \leq t \leq T_b \end{aligned} \quad (6.118)$$

其中“+”对应于 $\theta(0) = 0$, “-”对应于 $\theta(0) = \pi$ 。类似地, 可以证明, 在时间间隔 $0 \leq t \leq 2T_b$ 内, 正交分量 $s_Q(t)$ 由半周期正弦脉冲组成, 脉冲的极性仅取决于 $\theta(T_b)$, 即

$$\begin{aligned} s_Q(t) &= \sqrt{\frac{2E_b}{T_b}} \sin[\theta(t)] \\ &= \sqrt{\frac{2E_b}{T_b}} \sin[\theta(T_b)] \sin\left(\frac{\pi}{2T_b} t\right) \\ &= \pm \sqrt{\frac{2E_b}{T_b}} \sin\left(\frac{\pi}{2T_b} t\right), \quad 0 \leq t \leq 2T_b \end{aligned} \quad (6.119)$$

其中“+”对应于 $\theta(T_b) = \pi/2$, “-”对应于 $\theta(T_b) = -\pi/2$ 。

从前面的讨论可知, 由于相位状态 $\theta(0)$ 和 $\theta(T_b)$ 各有两种可能的取值, 因此, 可能出现的情况有 4 种:

- 相位 $\theta(0) = 0, \theta(T_b) = \pi/2$, 对应于发送符号 1。
- 相位 $\theta(0) = \pi, \theta(T_b) = \pi/2$, 对应于发送符号 0。
- 相位 $\theta(0) = \pi, \theta(T_b) = -\pi/2$ (或模为 2π 时的 $3\pi/2$), 对应于发送符号 1。
- 相位 $\theta(0) = 0, \theta(T_b) = -\pi/2$, 对应于发送符号 0。

这也就意味着根据 $\theta(0)$ 和 $\theta(T_b)$ 的取值的不同, MSK 信号自身可能取 4 种形式之一。

由式(6.116)的展开式可推导出, MSK 正交基本函数 $\phi_1(t)$ 和 $\phi_2(t)$ 由一组正弦调制的正交载波组成:

$$\phi_1(t) = \sqrt{\frac{2}{T_b}} \cos\left(\frac{\pi}{2T_b} t\right) \cos(2\pi f_c t), \quad 0 \leq t \leq T_b \quad (6.120)$$

$$\phi_2(t) = \sqrt{\frac{2}{T_b}} \sin\left(\frac{\pi}{2T_b} t\right) \sin(2\pi f_c t), \quad 0 \leq t \leq T_b \quad (6.121)$$

因此, 可将 MSK 信号展开为如下形式:

$$s(t) = s_1 \phi_1(t) + s_2 \phi_2(t), \quad 0 \leq t \leq T_b \quad (6.122)$$

其中系数 s_1 和 s_2 分别与相位状态 $\theta(0)$ 和 $\theta(T_b)$ 相对应。为了估计 s_1 的值, 将乘积 $s(t)\phi_1(t)$ 在区间 $-T_b$ 到 T_b 之间进行积分, 即

$$\begin{aligned} s_1 &= \int_{-T_b}^{T_b} s(t) \phi_1(t) dt \\ &= \sqrt{E_b} \cos[\theta(0)], \quad -T_b \leq t \leq T_b \end{aligned} \quad (6.123)$$

类似地, 为了估计 s_2 的值, 将乘积 $s(t)\phi_2(t)$ 在区间 0 到 $2T_b$ 之间进行积分, 即

$$\begin{aligned} s_2 &= \int_0^{2T_b} s(t) \phi_2(t) dt \\ &= -\sqrt{E_b} \sin[\theta(T_b)], \quad 0 \leq t \leq 2T_b \end{aligned} \quad (6.124)$$

不难注意到, 在式(6.123)和式(6.124)中:

- 两个积分都在长度为比特持续时间两倍的时间间隔内进行。
- 用来估算系数 s_1 的乘积积分的下限和上限, 相对于用来估算系数 s_2 的乘积积分的下限和上限, 有等于比特持续时间 T_b 的偏移。
- 两个积分都包括了时间间隔 $0 \leq t \leq T_b$, 即相位状态 $\theta(0)$ 和 $\theta(T_b)$ 定义的区间。

因此, MSK 信号的信号星座图是二维($N = 2$)的, 有四个可能的信息点($M = 4$), 如图 6.29 所示。信息点的坐标按照逆时针方向分别为 $(+\sqrt{E_b}, +\sqrt{E_b})$, $(-\sqrt{E_b}, +\sqrt{E_b})$, $(-\sqrt{E_b}, -\sqrt{E_b})$ 和 $(+\sqrt{E_b}, -\sqrt{E_b})$ 。图 6.29 中也标注出了对应于这四个信息点的 $\theta(0)$ 和 $\theta(T_b)$ 的可能取值。MSK 的信号空间图和 QPSK 的信号空间图都具有四个信息点, 但二者之间存在微妙的差别: 在 QPSK 中, 发送符号用 4 个信息点之中的任一个表示; 而在 MSK 中, 任一时刻, 发送符号则根据 $\theta(0)$ 的取值用两个信息点之一表示。

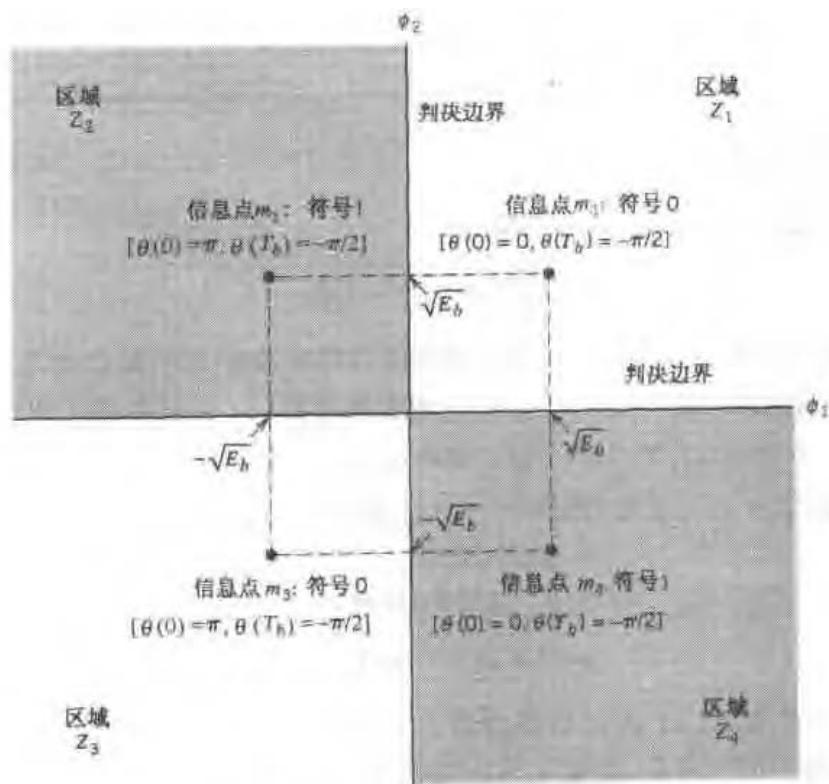


图 6.29 MSK 系统的信号空间图

表 6.5 分别列出了在时间间隔 $-T_b \leq t \leq T_b$ 和时间间隔 $0 \leq t \leq 2T_b$ 内, $\theta(0)$ 和 $\theta(T_b)$ 的取值以及相应的 s_1 和 s_2 的值。表的第一列表明在时间间隔 $0 \leq t \leq T_b$ 内发送的是符号 1 还是符号 0。注意, 当在时间间隔内发送的是符号 1 时, 信息点 s_1 和 s_2 的坐标异号; 当发送的是符号 0 时, 二者的坐标则同号。因此, 对于给定的输入数据序列, 可用表 6.5 的内容来一位一位地推导出衡量 $\phi_1(t)$ 和 $\phi_2(t)$ 所需的两个系数序列, 从而确定 MSK 信号 $s(t)$ 。

例 6.5

图 6.30 为产生与二进制序列 1101000 对应的 MSK 信号的过程中所出现的序列和波形。

输入二进制序列如图 6.30(a)。两个调制频率分别为: $f_1 = 5/4 T_b$ 和 $f_2 = 3/4 T_b$ 。假设在时刻 $t=0$, 相位 $\theta(0)$ 等于 0, 则模为 2π 的相位状态序列如图 6.30 所示。用于衡量时间函数 $\phi_1(t)$ 和 $\phi_2(t)$ 系数的两个序列的极性分别标注在图 6.30(b) 和 6.30(c) 的顶部。注意, 这两个序列之间的时间间隔偏移等于比特持续时间 T_b 。 $s(t)$ 的两个分量 $s_1 \phi_1(t)$ 和 $s_2 \phi_2(t)$ 的波形也显示在图 6.30(b) 和 6.30(c) 中。将这两个调制波形相加, 就得到了图 6.30(d) 所示的 MSK 信号 $s(t)$ 。

表 6.5 MSK 的信号-空间特性

发送的二进制符号 $0 \leq t \leq T_b$	相位状态(弧度)		信息点坐标	
	$\theta(0)$	$\theta(T_b)$	s_1	s_2
0	0	$-\pi/2$	$+\sqrt{E_b}$	$+\sqrt{E_b}$
1	π	$-\pi/2$	$-\sqrt{E_b}$	$+\sqrt{E_b}$
0	π	$+\pi/2$	$-\sqrt{E_b}$	$-\sqrt{E_b}$
1	0	$+\pi/2$	$+\sqrt{E_b}$	$-\sqrt{E_b}$

6.5.8 MSK 的差错概率

在 AWGN 信道的情况下, 接收信号为

$$x(t) = s(t) + w(t)$$

其中 $s(t)$ 为发送的 MSK 信号, $w(t)$ 为具有零均值且功率谱密度为 $N_0/2$ 的高斯白噪声过程的样本函数。为了判决在时间间隔 $0 \leq t \leq T_b$ 内发送的是符号 1 还是符号 0, 需要建立用 $x(t)$ 来检测相位状态 $\theta(0)$ 和 $\theta(T_b)$ 的程序。为了使检测到的 $\theta(0)$ 最优化, 首先应确定时间间隔 $-T_b \leq t \leq T_b$ 内的接收信号 $x(t)$ 在参考信号 $\phi_1(t)$ 上的投影:

$$\begin{aligned} x_1 &= \int_{-T_b}^{T_b} x(t) \phi_1(t) dt \\ &= s_1 + w_1, \quad -T_b \leq t \leq T_b \end{aligned} \quad (6.125)$$

其中 s_1 由式(6.123)定义, w_1 为具有零均值且方差为 $N_0/2$ 的高斯随机变量的抽样值。由图 6.29 所示的信号空间图可见, 如果 $x_1 > 0$, 接收机选择估计值 $\hat{\theta}(0) = 0$ 。反之, 如果 $x_1 < 0$, 接收机则选择估计值 $\hat{\theta}(0) = \pi$ 。

类似地, 为了对 $\theta(T_b)$ 进行最优检测, 应确定时间间隔 $0 \leq t \leq 2T_b$ 内, 接收信号 $x(t)$ 在第二个参考信号 $\phi_2(t)$ 上的投影, 即

$$\begin{aligned} x_2 &= \int_0^{2T_b} x(t) \phi_2(t) dt \\ &= s_2 + w_2, \quad 0 \leq t \leq 2T_b \end{aligned} \quad (6.126)$$

其中 s_2 由式(6.124)确定, w_2 为另一个独立的、具有零均值且方差为 $N_0/2$ 的高斯随机变量的抽样值。再次参考图 6.29, 可见如果 $x_2 > 0$, 接收机选择估计值 $\hat{\theta}(T_b) = -\pi/2$; 而如果 $x_2 < 0$, 接收机则选择估计值 $\hat{\theta}(T_b) = \pi/2$ 。

为了重构原始的二进制序列, 将上述两组相位判决相结合起来并表述如下(见表 6.5):

- 如果有估计值 $\hat{\theta}(0) = 0, \hat{\theta}(T_b) = -\pi/2$, 或者有估计值 $\hat{\theta}(0) = \pi$ 和 $\hat{\theta}(T_b) = \pi/2$, 则接收机判决为符号 0。
- 如果有估计值 $\hat{\theta}(0) = \pi$ 和 $\hat{\theta}(T_b) = -\pi/2$, 或者有估计值 $\hat{\theta}(0) = 0$ 和 $\hat{\theta}(T_b) = \pi/2$, 则接收机判决为符号 1。

参考图 6.29 可知, 接收机对符号 1 的判决结果为信息点 m_1 或 m_3 , 对符号 0 的判决结果为信息点 m_2 或 m_4 。因此, 对于 $\theta(0)$ 为 0 或 π , 以及 $\theta(T_b)$ 为 $-\pi/2$ 或 $+\pi/2$ 的判决(位判决), 分别在接收机的 I 通道和 Q 通道中交替进行, 每个通道的输入信号持续 $2T_b$ 秒。在各个通道里, 其他位的信号不影响接收机对当前位的判决。当 I 通道对 $\theta(0)$ 计算错误, 或者 Q 通道对 $\theta(T_b)$ 计算错误时, 接收机将做出错误的判决。因此, 利用式(6.125)和式(6.126)定义的两个通道乘积积分器的输出 x_1 和 x_2 取值的统计特征, 可得相干 MSK 的误比特率为

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \quad (6.127)$$

这与二进制 PSK 以及 QPSK 的结果完全一致。然而值得注意的是, MSK 信号有上述优良性能的原因是, 接收机对信号的检测是基于持续时间为 $2T_b$ 秒的观察结果的。

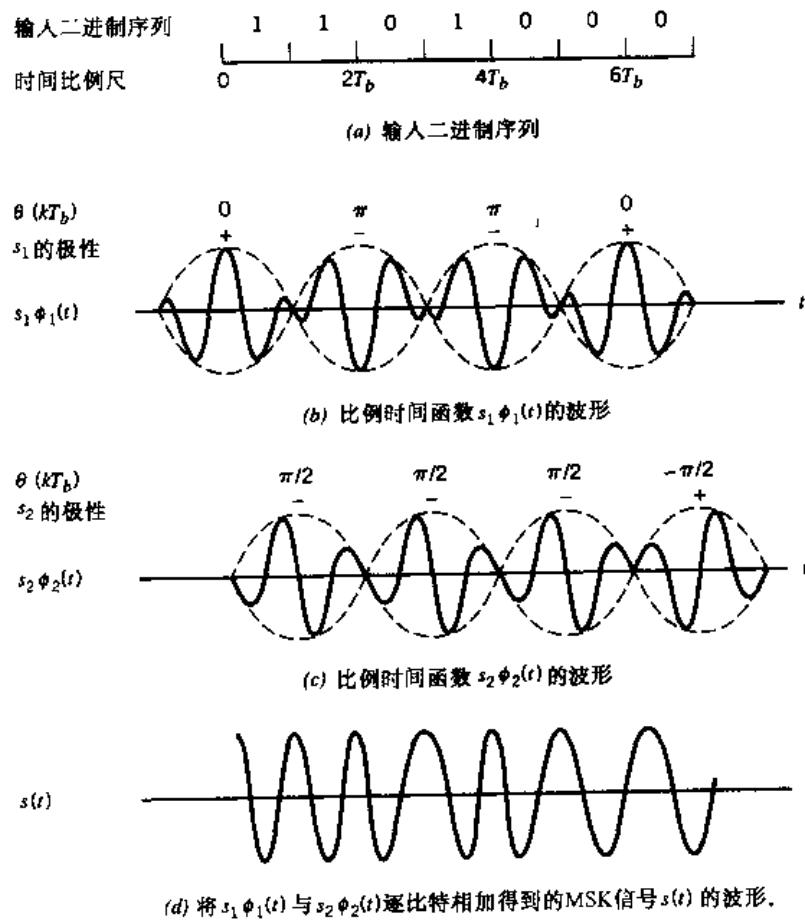


图 6.30 MSK 信息

6.5.9 MSK 信号的产生和检测

下面讨论 MSK 信号的产生和检测。图 6.31(a)为典型的 MSK 发射机框图。这种产生 MSK 信号方法的优点是,信号的相干性和偏移率几乎不受输入数据速率变化的影响。两个输入正弦信号首先被送入乘积调制器,其中一个频率为 $f_c = n_e/4T_b$,另一个频率为 $1/4T_b$, n_e 为特定整数。这就产生了两个频率分别为 f_1 和 f_2 的相位相干的正弦信号,根据式(6.111)和式(6.112), f_1 和 f_2 与载波频率 f_c 以及比特速率 $1/T_b$ 相关,这里 $b = 1/2$ 。这两个正弦信号由两个中心频率分别为 f_1 和 f_2 的窄带滤波器分开。将滤波器的输出进行线性组合,从而得到正交载波对,或正交基本函数 $\phi_1(t)$ 和 $\phi_2(t)$ 。最后, $\phi_1(t)$ 和 $\phi_2(t)$ 分别和两个比特速率等于 $1/2T_b$ 的二进制波形 $a_1(t)$ 和 $a_2(t)$ 相乘,这两个二进制序列是按照例 6.5 中的方法从输入二进制序列中提取的。

图 6.31(b)为典型的 MSK 接收机框图。接收信号 $x(t)$ 分别与本地产生的相干参考信号 $\phi_1(t)$ 和 $\phi_2(t)$ 进行相关。需注意积分时间间隔都是 $2T_b$ 秒,正交信道的积分比同相信道的积分延迟 T_b 秒。得到的同相和正交信道的相关输出 x_1 和 x_2 ,将与门限值 0 进行比较,并按照上述方法推导出相位 $\theta(0)$ 和 $\theta(T_b)$ 的估计值。最后,将这些相位判决结合,就能重构出原始的输入二进制序列。这种判决方法在 AWGN 信道中具有最小的平均符号差错概率。

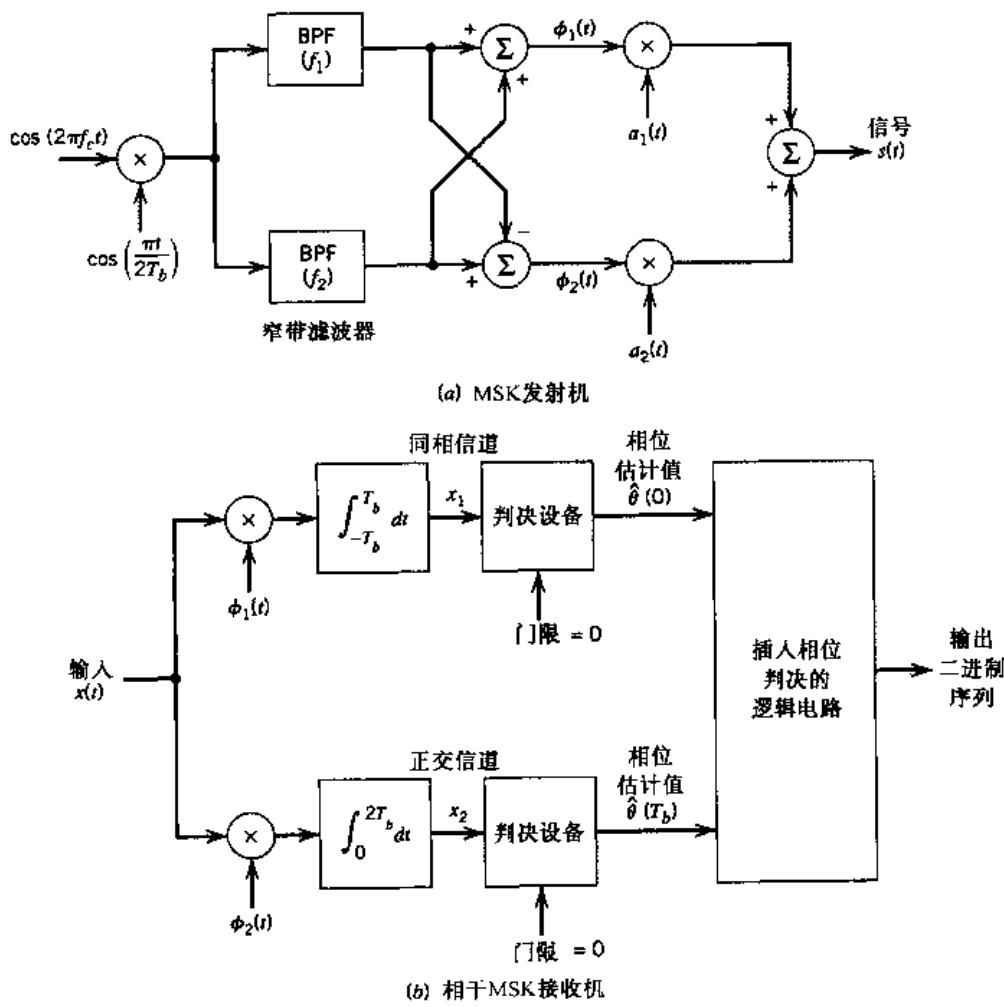


图 6.31 MSK 发射机和接收机框图

6.5.10 MSK 信号的功率谱密度

和讨论二进制 FSK 信号时一样,假设输入的二进制波形是随机的,符号 1 和 0 等概地出现,不同时间间隔内发送的符号彼此统计独立。在这样的情况下,有以下结论:

- 根据相位状态 $\theta(0)$ 的取值,同相分量等于 $+g(t)$ 或 $-g(t)$,其中

$$g(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos\left(\frac{\pi t}{2T_b}\right), & -T_b \leq t \leq T_b \\ 0, & \text{其他} \end{cases} \quad (6.128)$$

符号波形函数的能量谱密度为

$$\psi_g(f) = \frac{32E_bT_b}{\pi^2} \left[\frac{\cos(2\pi T_b f)}{16T_b^2 f^2 - 1} \right]^2 \quad (6.129)$$

因此,同相分量的功率谱等于 $\psi_g(f)/2T_b$ 。

- 根据相位状态 $\theta(T_b)$ 的取值,正交分量等于 $+g(t)$ 或 $-g(t)$,其中

$$g(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \sin\left(\frac{\pi t}{2T_b}\right), & 0 \leq t \leq 2T_b \\ 0, & \text{其他} \end{cases} \quad (6.130)$$

这个符号波形函数的能量谱密度也由式(6.129)确定。因此,同相和正交分量有相同的功率谱密度。

- MSK 信号的同相和正交分量也是统计独立的。因此,MSK 信号的基带功率谱密度为

$$\begin{aligned} S_B(f) &= 2 \left[\frac{\psi_g(f)}{2T_b} \right] \\ &= \frac{32E_b}{\pi^2} \left[\frac{\cos(2\pi T_b f)}{16T_b^2 f^2 - 1} \right]^2 \end{aligned} \quad (6.131)$$

式(6.131)的基带功率谱密度如图 6.9 所示,其中,功率谱已按照 $4E_b$ 进行了归一化,频率 f 已按照比特速率 $1/T_b$ 进行了归一化。图 6.9 中还包括了式(6.40)中 QPSK 信号的相应曲线。对于 $f >> 1/T_b$ 的情况,MSK 信号的基带功率谱密度随着频率四次方的倒数迅速下降。而对于 QPSK 信号,基带功率谱密度则按照频率平方的倒数下降。因此,MSK 在我们关注的信号通带外所受的干扰小于 QPSK。这是 MSK 的一个令人满意的特性,其优越性在有带宽限制的数字通信系统中尤其明显。

6.5.11 高斯滤波 MSK

通过以上对 MSK 的详细讨论,可将 MSK 的优点总结如下:

- 恒定包络
- 相对较窄的带宽
- 与 QPSK 相当的相干检测性能

虽然 MSK 信号的带外功率特性已经很好,但仍不能满足无线通信等应用的严格限制要求。由式(6.131)可知,在 $fT_b = 0.5$ 处,MSK 信号的基带功率谱密度比中心带宽处仅下降了 $10 \log_{10} 9 = 9.54$ dB。因此,当 MSK 信号的发射带宽为 $1/T_b$ 时,采用 MSK 的无线通信系统中相近信道的干扰还不够小,还不能满足这样的多用户通信环境的实际需要。

可通过对压控振荡器(VCO)进行直接的频率调制来产生 MSK 信号。因此,可在保持 MSK 信号包络恒定的同时,将其功率谱调制为一种紧凑的形式,从而克服 MSK 的这种严重缺陷。这种调制可通过采用预调制低通滤波器来实现。这种滤波器称为基带脉冲整形滤波器。脉冲整形滤波器应具有如下特点:

1. 频率响应具有窄通带和快速衰减的特性。
2. 冲激响应的过冲较小。
3. 和 MSK 的相位网格图相似,调制信号的载波相位在 T_b 的奇数倍时刻取值为 $\pm \pi/2$,在 T_b 的偶数倍时刻取值为 0 和 π 。

利用条件 1,可抑制发射信号的高频分量。利用条件 2,可防止 FM 信号瞬时频率过大偏移。最后,条件 3 确保了可用与 MSK 信号一样的方法来进行对 FM 信号进行相干检测,或用与简单二进制 FSK 信号一样的方法来对其进行非相干检测。

通过将不归零二进制数据流通过一个基带脉冲整形滤波器,就可以获得上述优点。该基带脉冲整形滤波器的冲激响应(和频率响应)由高斯函数定义。这样的二进制序列调制方法就称为高斯滤波 MSK 或者 GMSK。^[8]

用 W 表示脉冲整形滤波器的 3 dB 基带带宽。可将脉冲整形滤波器的传输函数 $H(f)$ 以及冲激响应 $b(t)$ 分别定义为:

$$H(f) = \exp\left(-\frac{\log 2}{2}\left(\frac{f}{W}\right)^2\right) \quad (6.132)$$

和

$$b(t) = \sqrt{\frac{2\pi}{\log 2}} W \exp\left(-\frac{2\pi^2}{\log 2} W^2 t^2\right) \quad (6.133)$$

该高斯滤波器对持续时间为 T_b (中心在原点)的单位幅度方波的响应为(见习题 6.28):

$$\begin{aligned} g(t) &= \int_{-T_b/2}^{T_b/2} b(t - \tau) d\tau \\ &= \sqrt{\frac{2\pi}{\log 2}} W \int_{-T_b/2}^{T_b/2} \exp\left(-\frac{2\pi^2}{\log 2} W^2(t - \tau)^2\right) d\tau \end{aligned} \quad (6.134)$$

该响应可表示为两个互补误差函数的差值,即

$$g(t) = \frac{1}{2} \left[\operatorname{erfc}\left(\pi \sqrt{\frac{2}{\log 2}} WT_b \left(\frac{t}{T_b} - \frac{1}{2}\right)\right) - \operatorname{erfc}\left(\pi \sqrt{\frac{2}{\log 2}} WT_b \left(\frac{t}{T_b} + \frac{1}{2}\right)\right) \right] \quad (6.135)$$

冲激响应 $g(t)$ 构成了 GMSK 调制器的频率整形脉冲,其中无量纲的时间-带宽积 WT_b 是一个关键的设计参数。

如式(6.135)所示的频率整形脉冲 $g(t)$ 是非因果响应,因为它在 $t < -T_b/2$ 时为非零,其

中 $t = -T_b/2$ 是输入方波(关于原点对称)被送入高斯滤波器的时刻。对于因果响应,必须将 $g(t)$ 截断并按时间进行移位。图 6.32 中所示的 $g(t)$ 已经在 $t = \pm 2.5T_b$ 处被截断,并按时间移位了 $2.5T_b$ 。其中,分别给出了 $WT_b = 0.2, 0.25$ 和 0.3 时 $g(t)$ 的曲线。值得注意的是,随着 WT_b 的减小,频率整形脉冲的时间展宽则相应增加。

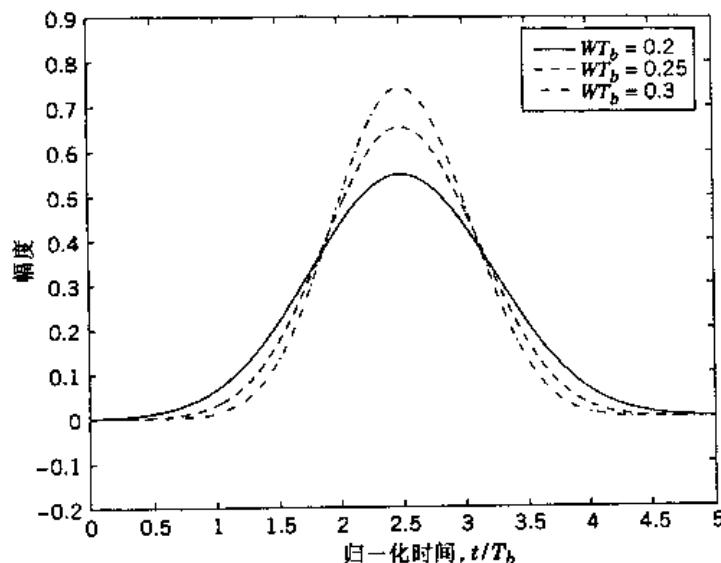


图 6.32 不同时间 – 带宽积 WT_b 下,式(6.135)对应的频率整形脉冲 $g(t)$ 。其中 $g(t)$ 按时间移位 $2.5T_b$,并在 $\pm 2.5T_b$ 处截断

图 6.33 为用计算机计算出的 MSK 信号的谱功率(用 dB 表示)相对于归一化频差 $(f - f_c)T_b$ 的曲线,其中, f_c 为中间频率, T_b 为比特持续时间。^[9]图 6.33 中的结果对应着不同的时间 – 频率积 WT_b 。从该图可得如下结论:

- $WT_b = \infty$ 条件下的曲线对应于普通 MSK。
- 当 WT_b 小于 1 时,GMSK 信号的通带内集中的发射功率显著增大。

GMSK 的一个缺点就是,不归零二进制数据通过高斯滤波器后产生的调制信号将不再像普通 MSK 那样限制在单个比特持续时间内,如图 6.32 所示。换言之,脉冲整形滤波器的高斯冲激响应的尾部将使调制信号延伸到临近的符号持续时间里。其直接结果就是产生了 ISI,并且干扰的程度随 WT_b 的减小而增大。根据这一发现,以及对图 6.33 所示 GMSK 信号谱功率观察的结果,可以说,对时间 – 带宽积 WT_b 的选择需要在频谱收敛性和性能损失之间进行权衡。

下面研究系统性能下降的问题。考虑在加性高斯白噪声的环境里使用相干检测的 GMSK 的差错概率为 P_e 。由于 GMSK 是一种特殊的二进制频率调制,因此可用经验公式将 P_e 表示为

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{\alpha E_b}{2N_0}}\right) \quad (6.136)$$

其中, E_b 和前面一样,是每比特的信号能量,而 $N_0/2$ 为信号功率谱密度。因子 α 是一个由时间 – 带宽积 WT_b 决定的常数。比较 GMSK 对应的式(6.136)和普通 MSK 对应的式(6.127)可知, GMSK(对于指定的 WT_b)与普通 MSK 相比,有 $10 \log_{10}(\alpha/2)$ dB 的性能下降。图 6.34 显示了

$10 \log_{10}(\alpha/2)$ 的计算值随 WT_b 的变化规律。对于普通 MSK, 有 $WT_b = \infty$, 在此情况下将 $\alpha = 2$ 代入式(6.136), 即得到与式(6.127)完全相同的形式, 此时没有任何的性能下降, 这在图 6.34 中可以得到证实。而 $WT_b = 0.3$ 的 GMSK 中则存在大约 0.46 dB 的性能下降, 如图 6.34 所示, 此时对应的 $\alpha/2 = 0.9$ 。为了获得我们所期望的具有较高收敛性的 GMSK 信号, 出现这样的性能下降是允许的。

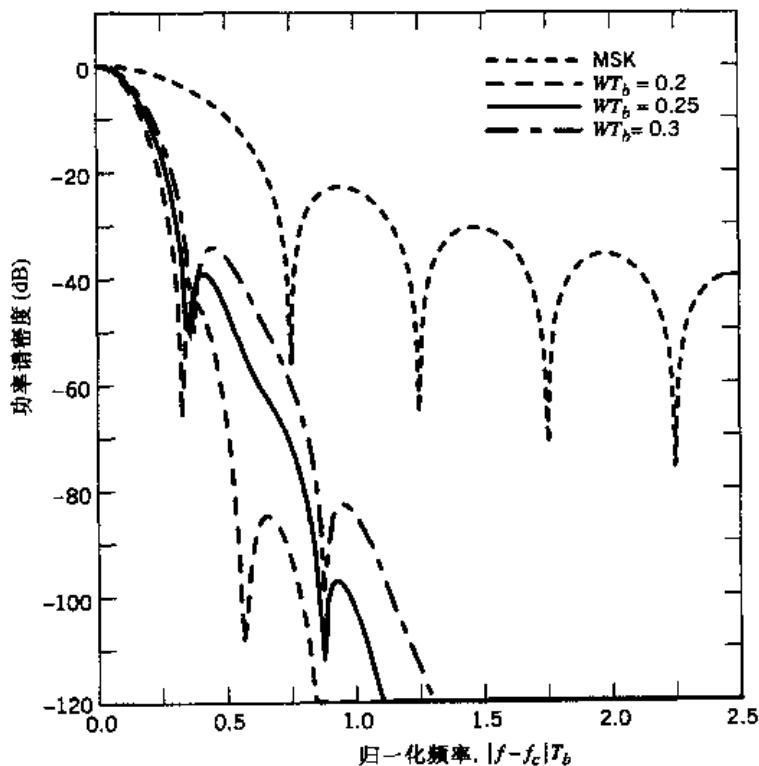


图 6.33 不同时间 - 带宽积下, MSK 和 GMSK 信号的功率谱密度(经 Georgia Tech 的 Gordon Stüber 博士允许而复制)

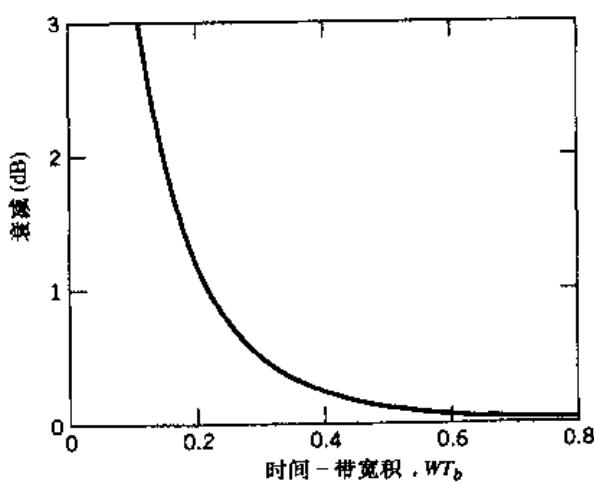


图 6.34 不同时间 - 带宽积下, GMSK 的理论 E_b/N_0 衰减值(经 IEEE 允许摘自 Murata 和 Hirade, 1981)

例 6.6 GMSK 用于 GSM 无线通信

GMSK 的一个重要应用就是在众所周知的标准无线通信系统 GSM 中, 即第 8 章将要讨论的

TDM 系统。对于这种应用, CMSK 的时间 - 频率积 WT_b 为标准值 0.3, 这样可以在提高带宽利用率和减小共信道干扰之间获得最好的折中。GMSK 信号 99% 的射频功率被限制在 250 kHz 的带宽内。也就是说, 对于所有的实际应用, 在此频带外的旁瓣实际上是 0。

在 GSM 系统中, 可用频谱被分成宽为 200 kHz 的子信道, 每个子信道都以 271 Kbps 的速率发送数据。图 6.35 显示了一个子信道与两个相邻子信道之间的功率谱关系。该图是对应于图 6.33 中 $WT_b = 0.3$ 的基带功率谱曲线的通带形式。由图 6.35 可得以下重要结论: 阴影所示的子信道的射频功率谱在两个相邻信道内的载波频率处都下降了至少 40 dB, 这意味着公共信道干扰实际上是可以忽略的。

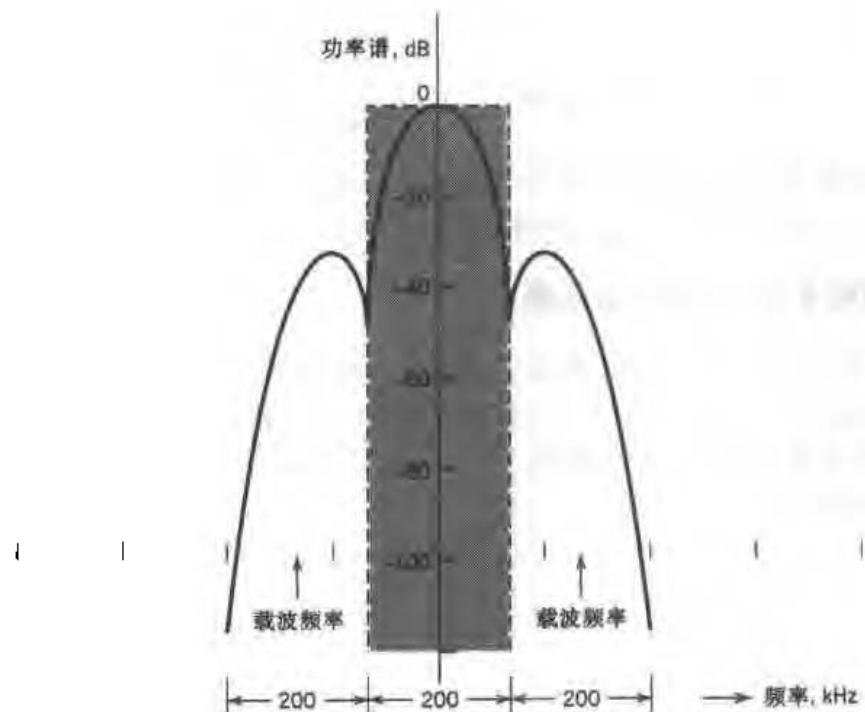


图 6.35 GSM 无线通信系统中的 GMSK 信号功率谱

6.5.12 M 进制 FSK

下面考虑 M 进制 FSK, 其发射信号的定义为

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos\left[\frac{\pi}{T}(n_c + i)t\right], \quad 0 \leq t \leq T \quad (6.137)$$

其中 $i = 1, 2, \dots, M$, 载波频率 $f_c = n_c/2T$, n_c 为固定整数。发送符号的持续时间为 T , 并具有相同的能量 E 。由于单个信号频率相隔 $1/2T$ 赫, 因此式(6.137)中的信号是正交的, 即

$$\int_0^T s_i(t)s_j(t) dt = 0, \quad i \neq j \quad (6.138)$$

M 进制 FSK 的这一特性表明, 可将发射信号 $s_i(t)$ 按能量进行归一化, 从而构造出基本函数的完全正交集, 即

$$\phi_i(t) = \frac{1}{\sqrt{E}} s_i(t), \quad 0 \leq t \leq T \quad i = 1, 2, \dots, M \quad (6.139)$$

因此, M 进制 FSK 由 M 维信号空间图描述。

对于相干 M 进制 FSK, 最佳接收机由 M 个相关器或者匹配滤波器组成, 由式(6.139)中的 $\phi_i(t)$ 提供相干参考信号。在抽样时刻 $t = kT$, 接收机根据最大匹配滤波器的输出, 进行符合最大似然译码准则的判决。对于相干 M 进制 FSK 系统, 很难推导出确切的符号差错概率计算公式。但可利用第 5 章式(5.96)中的联合边界, 给出 M 进制 FSK 的平均符号差错概率的上界。特别地, 由于 M 进制 FSK 的最小距离 d_{\min} 为: $\sqrt{2E}$, 利用式(5.96)可得(假设符号等概地发送):

$$P_e \leq \frac{1}{2} (M - 1) \operatorname{erfc}\left(\sqrt{\frac{E}{2N_0}}\right) \quad (6.140)$$

对于固定的 M , 随着 E/N_0 的增大, 该边界迅速减小。当 $P_e \leq 10^{-3}$ 时, 该边界已经和 P_e 十分接近了。更进一步, 对于 $M = 2$ (二进制 FSK)的情况, 式(6.140)中的边界是一个等式。

6.5.13 M 进制 FSK 信号的功率谱

M 进制 FSK 信号^[10]的谱分析比 M 进制 PSK 信号要复杂得多。我们特别关心的是当分配的频率为多级, 从而使得频率间隔均匀, 且频率偏移为 $k = 0.5$ 时的情况。也就是说, M 个信号频率相隔 $1/2T$, 其中 T 是符号持续时间。 $k = 0.5, M = 2, 4, 8$ 时 M 进制 FSK 信号的基带功率谱密度如图 6.36 所示。

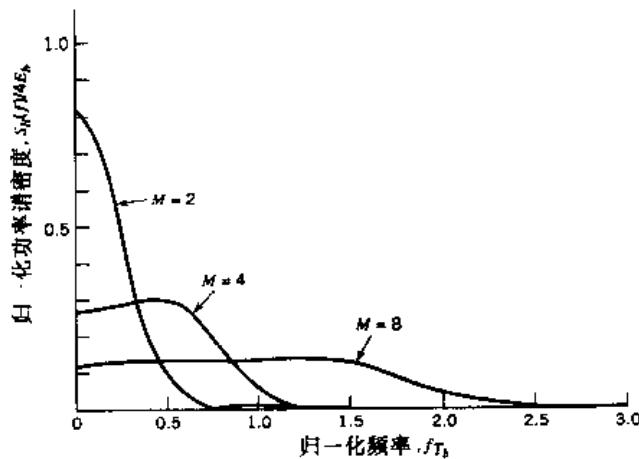


图 6.36 M 进制 FSK 信号在 $M = 2, 4, 8$ 时的功率谱

6.5.14 M 进制 FSK 信号的带宽效率

当对 M 进制 FSK 信号的正交信号进行相干检测时, 为保持其正交性, 就要确保相邻信号的频率彼此相隔 $1/2T$ 。因此, 将传输 M 进制 FSK 信号所需的信道带宽定义为

$$B = \frac{M}{2T} \quad (6.141)$$

对于频率分配的各级,式(6.141)的带宽 B 中包含了信号功率中的很大一部分。这从图 6.36 的基带功率谱曲线就可以得到证实。由式(6.48)可知,符号持续时间 T 等于 $T_b \log_2 M$ 。利用 $R_b = 1/T_b$,可将 M 进制 FSK 信号的信道带宽 B 重新定义为

$$B = \frac{R_b M}{2 \log_2 M} \quad (6.142)$$

因此, M 进制信号的带宽效率为

$$\rho = \frac{R_b}{B} = \frac{2 \log_2 M}{M} \quad (6.143)$$

表 6.6 给出了不同的 M 下,利用式(6.143)计算出的 ρ 值。

表 6.6 M 进制 FSK 信号的带宽效率

M	2	4	8	16	32	64
$\rho(\text{bs}^{-1}\text{Hz}^{-1})$	1	1	0.75	0.5	0.3125	0.1875

比较表 6.4 和表 6.6 可以看出,增加级数 M ,将带来 M 进制 PSK 信号带宽效率的提高,但同时也会带来 M 进制 FSK 信号带宽效率的下降。换言之, M 进制 PSK 信号的频谱效率高,而 M 进制 FSK 信号的频谱效率低。

6.6 未知相位信号的检测

在前面的讨论中,均假设接收机与发射机进行了很好的同步,且仅有的信道干扰是噪声。然而实际上,除了信道噪声引起的不确定因素外,还存在由于信号参数的随机性引起的不确定因素。产生这种不确定因素的原因是发送媒质里的失真。最普遍的随机信号参数可能就是载波相位,尤其是在窄带信号中。例如,通过长度不同且变化着的多个信道进行传输,或者从发射机到接收机的传播媒质里存在迅速变化的时延,都会使接收信号的相位以接收机无法跟踪的方式发生变化。此时按照发射载波的相位进行同步的代价将变得很大。设计者可简单地选择忽略接收信号的相位信息,其代价是噪声性能的少量下降。不提供载波相位恢复的数字通信接收机称为是非相干的。

6.6.1 最佳二次接收机

考虑具有如下发射信号的二进制数字通信系统:

$$s_i(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_i t), \quad 0 \leq t \leq T, \quad i = 1, 2 \quad (6.144)$$

其中, E 为信号能量, T 为信号持续时间,载波频率 f_i 取决于 i ,且为 $1/2T$ 的整数倍。假设系统是非相干的。此时,AWGN 信道的接收信号可写为如下形式:

$$x(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_i t + \theta) + w(t), \quad 0 \leq t \leq T, \quad i = 1, 2, \quad (6.145)$$

其中, θ 为未知的载波相位, $w(t)$ 为具有零均值且功率谱密度为 $N_0/2$ 的高斯白噪声样本函数。在现实环境下, 可假设完全没有关于先前 θ 的信息, 并将 θ 看成是具有如下均匀分布的随机变量:

$$f_\theta(\theta) = \begin{cases} \frac{1}{2\pi}, & -\pi < \theta \leq \pi \\ 0, & \text{其他} \end{cases} \quad (6.146)$$

现将要解决的二进制检测问题列举如下:

假设接收信号 $x(t)$ 具有未知的载波相位 θ , 设计最佳接收机, 以检测出 $x(t)$ 中包含的由信号分量 $\sqrt{E/2T} \cos(2\pi f_i t + \theta)$ 代表的符号 s_i 。

按照与 5.3 节 ~ 5.6 节相类似的方法, 可将给定载波相位 θ 时, 符号 s_i 的条件概率函数表述为

$$L(s_i(\theta)) = \exp\left(\sqrt{\frac{E}{N_0 T}} \int_0^T x(t) \cos(2\pi f_i t + \theta) dt\right) \quad (6.147)$$

为了进行更加深入的研究, 需要去除 $L(s_i(\theta))$ 对相位 θ 的依赖性, 这通过将 $L(s_i(\theta))$ 对 θ 的所有可能值进行积分就能够实现。于是

$$\begin{aligned} L(s_i) &= \int_{-\pi}^{\pi} L(s_i(\theta)) f_\theta(\theta) d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left(\sqrt{\frac{E}{N_0 T}} \int_0^T x(t) \cos(2\pi f_i t + \theta) dt\right) d\theta \end{aligned} \quad (6.148)$$

也就是说, 对概率函数而不是对数概率函数进行积分, 就能够去除 $L(s_i(\theta))$ 对 θ 的依赖性。

利用三角恒等式, 可将 $\cos(2\pi f_i t + \theta)$ 展开为如下形式:

$$\cos(2\pi f_i t + \theta) = \cos(2\pi f_i t) \cos \theta - \sin(2\pi f_i t) \sin \theta$$

因此, 可将式(6.148)中的积分重新写为

$$\int_0^T x(t) \cos(2\pi f_i t + \theta) dt = \cos \theta \int_0^T x(t) \cos(2\pi f_i t) dt - \sin \theta \int_0^T x(t) \sin(2\pi f_i t) dt \quad (6.149)$$

定义

$$l_i = \left[\left(\int_0^T x(t) \cos(2\pi f_i t) dt \right)^2 + \left(\int_0^T x(t) \sin(2\pi f_i t) dt \right)^2 \right]^{1/2} \quad (6.150)$$

$$\beta_i = \tan^{-1} \left(\frac{\int_0^T x(t) \sin(2\pi f_i t) dt}{\int_0^T x(t) \cos(2\pi f_i t) dt} \right) \quad (6.151)$$

因此, 可进一步将函数(6.149)简化为

$$\begin{aligned} \int_0^T x(t) \cos(2\pi f_i t + \theta) dt &= l_i (\cos \theta \cos \beta_i - \sin \theta \sin \beta_i) \\ &= l_i \cos(\theta + \beta_i) \end{aligned} \quad (6.152)$$

将式(6.152)代入式(6.148)。可得

$$\begin{aligned} L(s_i) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left(\sqrt{\frac{E}{N_0T}} l_i \cos(\theta + \beta_i)\right) d\theta \\ &= \frac{1}{2\pi} \int_{-\pi+\beta_i}^{\pi+\beta_i} \exp\left(\sqrt{\frac{E}{N_0T}} l_i \cos \theta\right) d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left(\sqrt{\frac{E}{N_0T}} l_i \cos \theta\right) d\theta \end{aligned} \quad (6.153)$$

由附录3可知,式(6.153)中的积分就是经过修正的零阶贝塞尔函数:

$$I_0\left(\sqrt{\frac{E}{N_0T}} l_i\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left(\sqrt{\frac{E}{N_0T}} l_i \cos \theta\right) d\theta \quad (6.154)$$

因此,此时可将信号检测问题的概率函数简写为

$$L(s_i) = I_0\left(\sqrt{\frac{E}{N_0T}} l_i\right) \quad (6.155)$$

于是,假设发送的是信号 $s_1(t)$ 或信号 $s_2(t)$,二进制假设检验可写为

$$I_0\left(\sqrt{\frac{E}{N_0T}} l_1\right) \stackrel{H_1}{\geq} I_0\left(\sqrt{\frac{E}{N_0T}} l_2\right) \quad (6.156)$$

其中假设值 H_1 和 H_2 分别对应于信号 $s_1(t)$ 和 $s_2(t)$ 。由附录3可知,修正的贝塞尔函数 $I(\cdot)$ 是一个随幅角单调递增的函数。因此,假设检测可以仅利用 $I_0(\sqrt{E/N_0T}l_i)$ 或 l_i 来进行。为了实现的方便,通常利用 l_i^2 而不是 l_i 进行假设检测,即

$$l_1^2 \stackrel{H_1}{\geq} l_2^2 \quad (6.157)$$

基于式(6.157)的接收机称为二次接收机。根据式(6.150)中对 l_i 的定义,用于计算 l_i 的接收机结构如图6.37(a)所示。不难注意到,式(6.157)所描述的假设检测与信号能量 E 是相互独立的。正是由于这个原因,假设检验尤其适用于符号能量 E 。

6.6.2 二次接收机的两种等价形式

接下来推导图6.37(a)所示的二次接收机的两种等价形式。将图6.37(a)中的相关器都用等价的匹配滤波器来代替,就可以得到第一种形式,即图6.37(b)所示的正交接收机。在该接收机的一个支路里,有一个与信号 $\cos(2\pi f_i t)$ 相匹配的滤波器,而在另一个支路里,有一个与信号 $\sin(2\pi f_i t)$ 相匹配的滤波器。这两个匹配滤波器都定义在时间间隔 $0 \leq t \leq T$ 上。在时刻 $t = T$,滤波器输出被抽样、平方和相加。

为了得到正交接收机的第二种等价形式,假设有一个在时间间隔 $0 \leq t \leq T$ 上与 $s(t) = \cos(2\pi f_i t + \theta)$ 相匹配的滤波器。该匹配滤波器输出的包络显然不受相位 θ 的影响。因此,为了方便起见,可简单地选择冲激响应为 $\cos[2\pi f_i(T-t)]$ 的匹配滤波器(相当于 $\theta = 0$ 的情况)。该滤波器对应于接收信号 $x(t)$ 的输出为

$$\begin{aligned}
 y(t) &= \int_0^T x(\tau) \cos[2\pi f_i(T-t+\tau)] d\tau \\
 &= \cos[2\pi f_i(T-t)] \int_0^T x(\tau) \cos(2\pi f_i \tau) d\tau - \sin[2\pi f_i(T-t)] \int_0^T x(\tau) \sin(2\pi f_i \tau) d\tau
 \end{aligned} \tag{6.158}$$

匹配滤波器输出的包络与式(6.158)中积分的平方和的平方根成正比例。因此时刻 $t = T$ 的包络为

$$l_i = \left\{ \left[\int_0^T x(\tau) \cos(2\pi f_i \tau) d\tau \right]^2 + \left[\int_0^T x(\tau) \sin(2\pi f_i \tau) d\tau \right]^2 \right\}^{1/2} \tag{6.159}$$

但这只是正交接收机的输出。因此,对于任意相位 θ ,由包络检波器得到的与信号 $\cos(2\pi f_i t + \theta)$ 相匹配的滤波器输出(在时刻 T),与图 6.37(a)所示正交接收机的相应输出相一致。这种形式的接收机如图 6.37(c)。图 6.37(c)中匹配滤波器和包络检波器的组合称为非相干匹配滤波器。

在图 6.37(c)中匹配滤波器后加上包络检波器,是很有必要的。与方波射频相匹配的滤波器输出在抽样时刻 $t = T$ 达到正峰值。如果该滤波器的相位和信号不相匹配,峰值就很可能不会出现在抽样时刻。实际上,如果相位相差了 180° ,在抽样时刻将得到负峰值。图 6.38 表明了在两个限制条件($\theta = 0^\circ$ 和 $\theta = 180^\circ$)下匹配滤波器的输出。为了避免由于缺少关于相位 θ 的先前信息而导致的抽样值恶化,可仅保留匹配滤波器的输出包络,因为输出包络与相位失配 θ 是无关的。

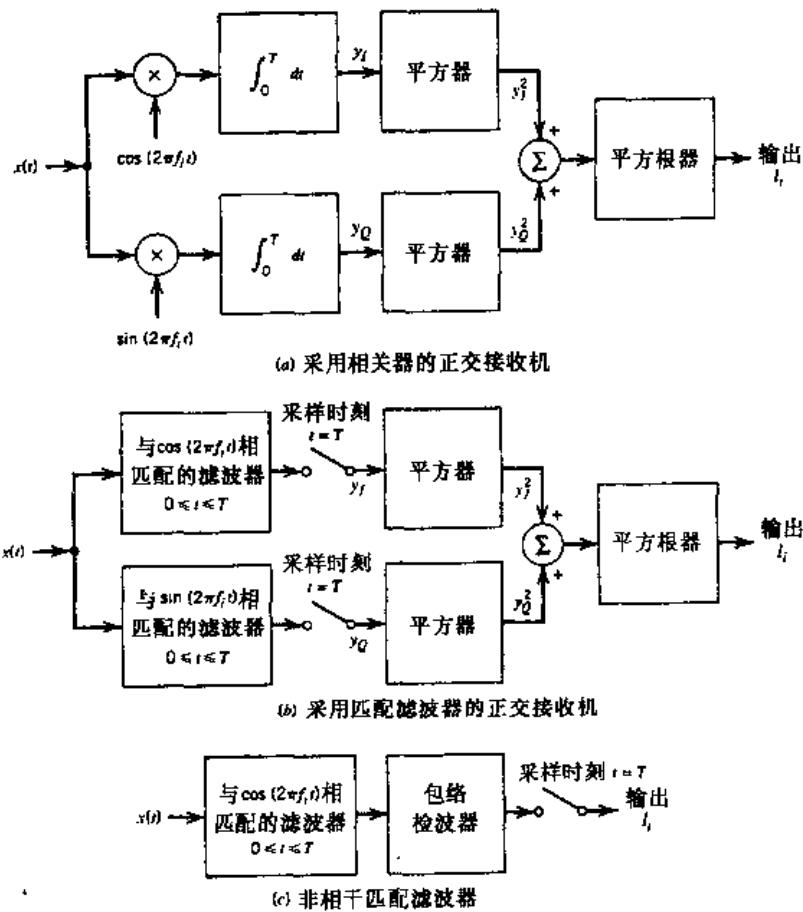


图 6.37 非相干接收机

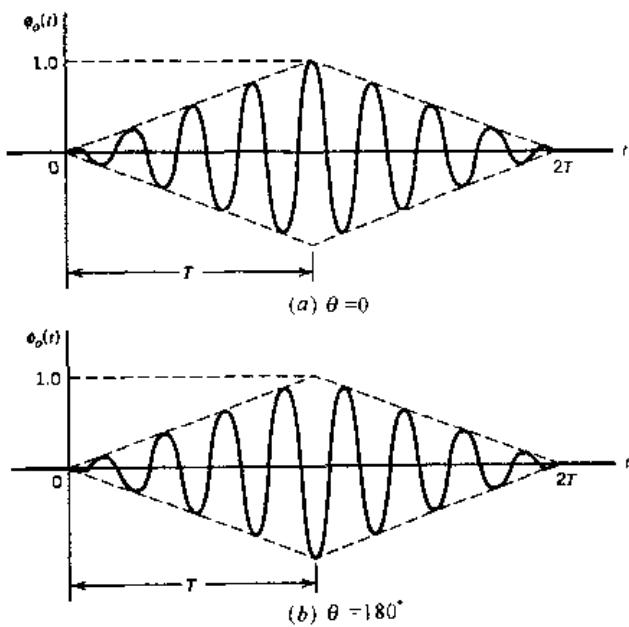


图 6.38 矩形射频信号的匹配滤波器输出

6.7 非相干正交调制

利用图 6.37 所示的非相干接收机结构, 可对非相干正交调制的噪声性能进行研究。考虑两种非相干接收机: 非相干二进制 FSK 和差分 PSK。

考虑一个采用两个等能量正交信号 $s_1(t)$ 和 $s_2(t)$ 的二进制信号方案。在时间间隔 $0 \leq t \leq T$ 内, 两个信号之一被送入非理想信道, 其载波相位将产生大小未知的相移。用 $g_1(t)$ 和 $g_2(t)$ 分别表示 $s_1(t)$ 和 $s_2(t)$ 的相移结果。不考虑未知的载波相位, $g_1(t)$ 和 $g_2(t)$ 都保持正交性并具有相等的能量 E 。这样的调制方式称为非相干正交调制。根据正交信号 $s_1(t)$ 和 $s_2(t)$ 的定义, 非相干二进制 FSK 和 DPSK 可以看成是这种调制方式的特例。

信道中还引入了具有零均值且功率谱密度为 $N_0/2$ 的加性高斯白噪声 $w(t)$, 因此, 可将接收信号 $x(t)$ 写为

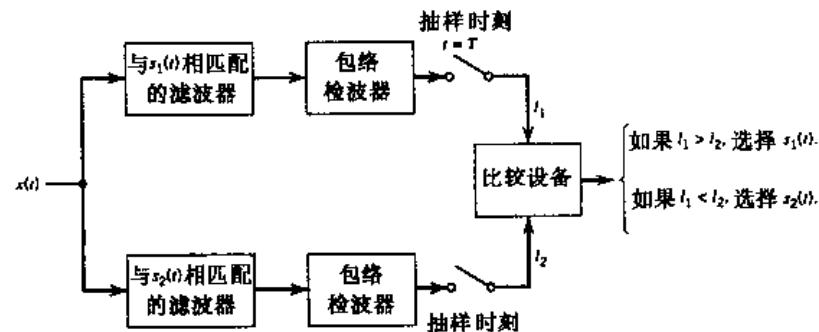
$$x(t) = \begin{cases} g_1(t) + w(t), & \text{发送 } s_1(t), 0 \leq t \leq T \\ g_2(t) + w(t), & \text{发送 } s_2(t), 0 \leq t \leq T \end{cases} \quad (6.160)$$

其中不考虑载波相位, 并利用 $x(t)$ 来区分 $s_1(t)$ 和 $s_2(t)$ 。

为此, 采用图 6.39(a) 所示的接收机。该接收机由一对分别与发射信号 $s_1(t)$ 及 $s_2(t)$ 相匹配的滤波器组成。由于载波相位未知, 接收机仅利用幅度作为判决标准。因此, 对匹配滤波器的输出进行包络检波、抽样, 然后相互比较。如果图 6.39(a) 中上支路输出的幅度 I_1 大于下支路输出的幅度 I_2 , 接收机判决为信号 $s_1(t)$, 反之判决为信号 $s_2(t)$ 。当两条支路的输出幅度相等时, 接收机进行随机判决。无论如何, 当抑制接收信号 $x(t)$ 信号分量的匹配滤波器输出幅度(仅由噪声引起)大于通过该分量的匹配滤波器的输出幅度时, 就会产生判决错误。

从 6.6 节的讨论可知, 构成图 6.39(a) 所示接收机的上、下支路的非相干匹配滤波器可看成是正交接收机的一种等效形式。正交接收机自身有两个信道。图 6.39(b) 为正交接收机的

一种形式。在上方信道即同相信道中,接收信号 $x(t)$ 与对应于零载波相位的发射信号 $s_1(t)$ 或 $s_2(t)$ 的比例形式的函数 $\psi_i(t)$ 进行相关运算。在下方信道即正交信道中, $x(t)$ 则与对应于载波相位 $\psi_i(t)$ 移位 -90° 的函数 $\hat{\psi}_i(t)$ 进行相关运算。自然, $\psi_i(t)$ 和 $\hat{\psi}_i(t)$ 彼此正交。



(a) 非相干正交调制的通用二进制接收机

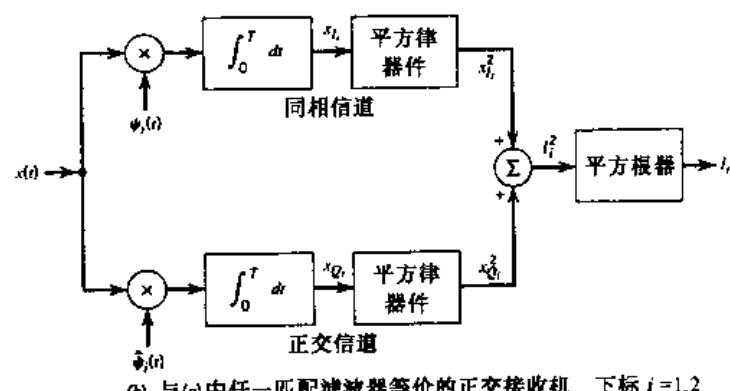
(b) 与(a)中任一匹配滤波器等价的正交接收机,下标 $i=1,2$

图 6.39 非相干正交调制接收机

信号 $\hat{\psi}_i(t)$ 实际上是 $\psi_i(t)$ 的希尔伯特变换。关于希尔伯特变换的介绍见附录 2。为了说明这一关系的实质,令

$$\psi_i(t) = m(t) \cos(2\pi f_i t) \quad (6.161)$$

其中, $m(t)$ 为带限信号。通常,载波频率 f_i 是大于 $m(t)$ 的最高频分量。因此,按照与 6.4 节对 CAP 相类似的讨论方法,可将 $\psi_i(t)$ 的希尔伯特变换定义为

$$\hat{\psi}_i(t) = m(t) \sin(2\pi f_i t) \quad (6.162)$$

由于

$$\cos\left(2\pi f_i t - \frac{\pi}{2}\right) = \sin(2\pi f_i t)$$

可见, $\hat{\psi}_i(t)$ 实际上是将 $\psi_i(t)$ 的载波相位 $\cos(2\pi f_i t)$ 移位 -90° 而得到的。希尔伯特变换的一个重要特性就是信号与其希尔伯特变换是彼此正交的。也就是说, $\psi_i(t)$ 与 $\hat{\psi}_i(t)$ 是彼此正交的。

图 6.39(a) 所示的非相干接收机的平均差错概率为

$$P_e = \frac{1}{2} \exp\left(-\frac{E}{2N_0}\right) \quad (6.163)$$

其中, E 为每个符号的信号能量, $N_0/2$ 为噪声功率谱密度。

6.7.1 式(6.163)的推导^①

为推导式(6.163), 需用到图 6.39 所示的等价形式。特别地, 由于载波相位未知, 图 6.39(a)中各个匹配滤波器输出端的噪声都有二维的自由度, 即同相和正交。因此, 在给定相位 θ 时, 图 6.39(a)所示的非相干接收机共有 4 个条件独立且均匀分布的噪声参数。这 4 个噪声参数的抽样值分别用 $x_{R1}, x_{Q1}, x_{R2}, x_{Q2}$ 表示。其中前两个代表了与图 6.39(a)中上方支路相对应的自由度, 后两个则代表了与下方支路相对应的自由度。

图 6.39(a)所示的接收机具有对称的结构。因此, 发送信号 $s_1(t)$ 时判决为 $s_2(t)$ 的概率, 与发送信号 $s_2(t)$ 时判决为 $s_1(t)$ 的概率相同。这就意味着当 $s_1(t)$ 和 $s_2(t)$ 等概地出现时, 只要在发送 $s_1(t)$ 时计算出判决为 $s_2(t)$ 的概率就可以得到平均差错概率, 反之亦然。

假设信号 $s_1(t)$ 在时间间隔 $0 \leq t \leq T$ 内发送。如果信道噪声 $w(t)$ 使得图 6.39(a)中下方支路的输出 l_2 大于上方支路的输出 l_1 , 接收机将错误地将 $s_1(t)$ 判决为 $s_2(t)$ 。为了计算由此产生的差错概率, 必须先得到随机变量 L_2 (用抽样值 l_2 代表)的概率密度函数。由于下方支路的滤波器与信号 $s_2(t)$ 相匹配, 而 $s_2(t)$ 和发射信号 $s_1(t)$ 正交, 因此, 该匹配滤波器的输出仅仅由噪声决定。用 x_{R2} 和 x_{Q2} 分别代表图 6.39(a)中下方支路匹配滤波器输出的同相分量和正交分量, 并根据图 6.39(b)所示等价形式, 可得(对于 $i=2$):

$$l_2 = \sqrt{x_{R2}^2 + x_{Q2}^2} \quad (6.164)$$

图 6.40(a)为上述关系的几何图示。信道噪声 $w(t)$ 为白色(功率谱密度为 $N_0/2$)高斯(均值为 0)噪声。相应地, 在给定相位 θ 时, 随机变量 X_{R2} 和 X_{Q2} (分别用抽样值 x_{R2} 和 x_{Q2} 代表)都是高斯分布的, 其均值为 0, 方差为 $N_0/2$ 。因此有

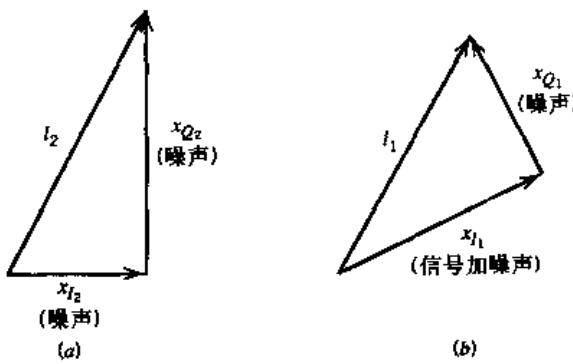


图 6.40 非相干接收机两路输出 l_1 和 l_2 的几何解释

^① 对式(6.163)的正式推导过程不感兴趣的读者, 可以跳过此部分, 开始学习非相干二进制 FSK(见 6.7 节)和 DPSK(见 6.8 节)这两个非相干正交调制的特例。

$$f_{X_{R2}}(x_{R2}) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{x_{R2}^2}{N_0}\right) \quad (6.165)$$

和

$$f_{X_{Q2}}(x_{Q2}) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{x_{Q2}^2}{N_0}\right) \quad (6.166)$$

下面,需用到概率论中的一个众所周知的结论,即高斯过程的包络是瑞利分布的,且与相位 θ 相互独立(见 1.12 节)。特别地,此时可认为随机变量 L_2 具有如下概率密度函数:

$$f_{L_2}(l_2) = \begin{cases} \frac{2l_2}{N_0} \exp\left(-\frac{l_2^2}{N_0}\right), & l_2 \geq 0 \\ 0, & \text{其他} \end{cases} \quad (6.167)$$

根据式(6.164), L_2 抽样值 l_2 与 x_{R2} 和 x_{Q2} 有关。图 6.41 为上式中概率密度函数的曲线。在给定抽样值 l_1 的条件下, $l_2 > l_1$ 的条件概率密度如图 6.41 中阴影部分所示。因此有

$$P(l_2 > l_1 | l_1) = \int_{l_1}^{\infty} f_{L_2}(l_2) dl_2 \quad (6.168)$$

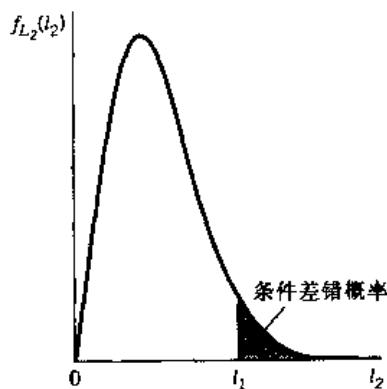


图 6.41 给定 l_1 时, $l_2 > l_1$ 的条件概率

将式(6.167)代入式(6.168)并积分,可得

$$P(l_2 > l_1 | l_1) = \exp\left(-\frac{l_1^2}{N_0}\right) \quad (6.169)$$

下面考虑图 6.39(a)中上方支路的输出幅度 l_1 。由于该支路的滤波器与信号 $s_1(t)$ 相匹配,假设发送的是 $s_1(t)$,因此, l_1 包括信号和噪声。用 x_{R1} 和 x_{Q1} 分别表示图 6.39(a)中上方支路的匹配滤波器输出端分别与接收信号同相和正交的分量。根据图 6.39(b)所示的等价形式,可得(对于 $i=1$):

$$l_1 = \sqrt{x_{R1}^2 + x_{Q1}^2} \quad (6.170)$$

图 6.40(b)为上述关系的几何图示。由于一个可进行傅里叶变换的信号与其希尔伯特变换构成了正交对,因此, x_{R1} 是由信号和噪声共同引起的,而 x_{Q1} 仅由噪声引起。这就意味着:(1)由抽样值 x_{R1} 代表的随机变量 X_{R1} 具有均值为 \sqrt{E} 、方差为 $N_0/2$ 的高斯分布,其中 E 为每个符号

的信号能量; (2)由抽样值 x_{Q1} 代表的随机变量 X_{Q1} 具有均值为零、方差为 $N_0/2$ 的高斯分布。因此, 可将这两个独立随机变量的概率密度函数写为

$$f_{X_H}(x_H) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{(x_H - \sqrt{E})^2}{N_0}\right) \quad (6.171)$$

和

$$f_{X_{Q1}}(x_{Q1}) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{x_{Q1}^2}{N_0}\right) \quad (6.172)$$

由于这两个随机变量 X_H 和 X_{Q1} 彼此独立, 其联合概率密度函数就是式(6.171)和式(6.172)各自概率密度函数的简单相乘。

为了得到平均差错概率, 需要按照式(6.169)计算出所有可能的 L_1 的取值对应的条件差错概率的平均值。因此, 首先需要知道由抽样值 L_1 代表的随机变量 L_1 的概率密度函数。标准方法是将式(6.171)和式(6.172)相结合, 从而得到信号和噪声共同引起的 L_1 的概率密度函数。这就带来了较为复杂的运算, 而且要用到贝塞尔函数。采用下述方法就可以避免上述的复杂计算。当给定 x_H 和 x_{Q1} 时, 如果图 6.39(a)中下方支路的输出幅度 L_2 超过了上方支路的输出幅度 L_1 , 就会发生差错。由式(6.170)可得

$$L_1^2 = x_H^2 + x_{Q1}^2 \quad (6.173)$$

将式(6.173)代入式(6.169), 得到所求的概率为

$$P(\text{差错} | x_H, x_{Q1}) = \exp\left(-\frac{x_H^2 + x_{Q1}^2}{N_0}\right) \quad (6.174)$$

这是一个条件差错概率, 条件是上方支路匹配滤波器在 X_H 和 X_{Q1} 上取值。给定 x_H 和 x_{Q1} , 将该条件概率密度和 X_H 与 X_{Q1} 的联合概率密度函数相乘, 就得到了差错概率密度。由于 X_H 和 X_{Q1} 是统计独立的, 因此其联合概率密度函数等于各自概率密度函数的乘积。所得的差错密度是一个由 x_H 和 x_{Q1} 表示的复数。我们所关注的平均差错概率可通过相对简单的方法获得。首先利用式(6.171), 式(6.172)和式(6.174)来计算期望差错密度, 即

$$\begin{aligned} & P(\text{差错} | x_H, x_{Q1}) f_{X_H}(x_H) f_{X_{Q1}}(x_{Q1}) \\ &= \frac{1}{\pi N_0} \exp\left\{-\frac{1}{N_0} [x_H^2 + x_{Q1}^2 + (x_H - \sqrt{E})^2 + x_{Q1}^2]\right\} \end{aligned} \quad (6.175)$$

计算出式(6.175)中指数里的平方项, 可将除 $-1/N_0$ 以外的指数写为如下形式:

$$x_H^2 + x_{Q1}^2 + (x_H - \sqrt{E})^2 + x_{Q1}^2 = 2\left(x_H - \frac{\sqrt{E}}{2}\right)^2 + 2x_{Q1}^2 + \frac{E}{2} \quad (6.176)$$

接着, 将式(6.176)代入式(6.175), 并将差错密度对所有的 x_H 和 x_{Q1} 进行积分, 这样就可以按下式计算出平均差错概率:

$$\begin{aligned}
 P_e &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(\text{error} | x_{I1}, x_{Q1}) f_{X_{I1}}(x_{I1}) f_{X_{Q1}}(x_{Q1}) dx_{I1} dx_{Q1} \\
 &= \frac{1}{\pi N_0} \exp\left(-\frac{E}{2N_0}\right) \int_{-\infty}^{\infty} \exp\left[-\frac{2}{N_0} \left(x_{I1} - \frac{\sqrt{E}}{2}\right)^2\right] dx_{I1} \\
 &\quad \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{2x_{Q1}^2}{N_0}\right) dx_{Q1}
 \end{aligned} \tag{6.177}$$

现在利用下面的恒等式：

$$\int_{-\infty}^{\infty} \exp\left[-\frac{2}{N_0} \left(x_{I1} - \frac{\sqrt{E}}{2}\right)^2\right] dx_{I1} = \sqrt{\frac{N_0 \pi}{2}} \tag{6.178}$$

和

$$\int_{-\infty}^{\infty} \exp\left(-\frac{2x_{Q1}^2}{N_0}\right) dx_{Q1} = \sqrt{\frac{N_0 \pi}{2}} \tag{6.179}$$

考慮均值为 $\sqrt{E/2}$ 、方差为 $N_0/4$ 的高斯分布变量，并注意到随机变量概率密度函数曲线下的总面积等于 1，就可以得到式(6.178)。式(6.179)则是式(6.178)的特例。根据这两个恒等式，可将式(6.177)简化为

$$P_e = \frac{1}{2} \exp\left(-\frac{E}{2N_0}\right)$$

这正是式(6.163)中所得到的结果。

有了这个公式，就可以在下面两节中分别讨论非相干二进制 FSK 和 DPSK 这两个特例了。^[11]

6.8 非相干二进制 FSK

在二进制 FSK 里，发射信号的定义如下：

$$s_i(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_i t), & 0 \leq t \leq T_b \\ 0, & \text{其他} \end{cases} \tag{6.180}$$

其中载波频率 f_i 为两个可能值 f_1 和 f_2 之一。为确保代表这两个频率的信号是正交的，选择 $f_i = n_i/T_b$ ，其中 n_i 为整数。频率 f_1 的发送对应着符号 1，而频率 f_2 的发送则对应着符号 0。为了对该频率调制波形进行非相干检测，采用的接收机是由匹配滤波器和包络检波器的组合构成，如图 6.42 所示。接收机上方支路的滤波器与 $\cos(2\pi f_1 t)$ 相匹配，下方支路的滤波器则与 $\cos(2\pi f_2 t)$ 相匹配，其中 $0 \leq t \leq T_b$ 。在 $t = T_b$ 时刻对包络检波器的输出进行抽样，并将上、下支路的抽样值进行比较。图 6.42 中上、下支路的包络抽样值分别记为 l_1 和 l_2 。如果 $l_1 > l_2$ ，接收机判决为符号 1。如果 $l_1 < l_2$ ，接收机则判决为符号 0。如果 $l_1 = l_2$ ，则简单地随机判决发送的是符号 1 还是 0。

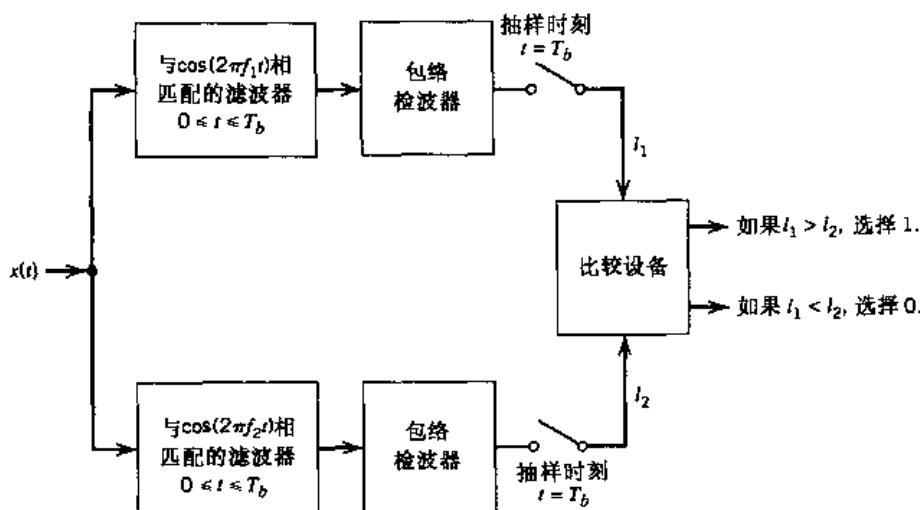


图 6.42 检测二进制 FSK 信号的非相干接收机

这里介绍的非相干二进制 FSK 为 $T = T_b$, $E = E_b$ 的非相干正交调制的一种特例。其中, T_b 为比特持续时间, 而 E_b 则为每比特的信号能量。因此, 利用式(6.163), 可得非相干二进制 FSK 的误比特率为

$$P_e = \frac{1}{2} \exp\left(-\frac{E_b}{2N_0}\right) \quad (6.181)$$

式(6.181)是非相干正交调制的特例。在习题 6.31 中, 通过分析瑞利分布和莱斯分布, 提到了同样的问题。这两种分布分别属于抽样值由式(6.164)和式(6.170)定义的随机变量 L_2 和 L_1 。

6.9 差分相移键控

正如在 6.1 节中所提到的, 我们可以把差分相移键控(DPSK)看成是一种非相干 PSK。在 DPSK 中, 通过在发射机端引入两个基本操作, 从而消除了接收机端对相干参考信号的需要。这两个基本操作是:(1)对输入二进制波形的差分编码;(2)PSK。因此, 这种调制方式称为差分相移键控。实际上, 为了发送符号 0, 可将当前信号波形向前移相 180°; 为了发送符号 1, 则应保持当前信号波形的相位不变。接收机具有存储能力, 因此能够比较两个相邻比特持续时间内接收波形的相对相位差。假设接收波形中的未知相位 θ 变化缓慢(也就是说, 慢得足以认为它在两个相继比特持续时间内是恒定的), 则在两个相邻比特持续时间内接收波形的相差将与 θ 无关。

当考察时间为比特持续时间的两倍时, DPSK 是另一种非相干正交调制的例子。假设在 $0 \leq t \leq T_b$ 内, 发送的 DPSK 信号为 $\sqrt{E_b/2T_b} \cos(2\pi f_c t)$, 其中, T_b 为比特持续时间, E_b 为每比特信号的能量。用 $s_1(t)$ 表示时间间隔 $0 \leq t \leq 2T_b$ 内发送的 DPSK 信号, 在该时间间隔的后半部分即 $T_b \leq t \leq 2T_b$ 内, 如果发射机输入信号为二进制符号 1, 则载波相位在时间间隔 $0 \leq t \leq 2T_b$ 内保持不变。因此, 可将 $s_1(t)$ 定义为

$$s_1(t) = \begin{cases} \sqrt{\frac{E_b}{2T_b}} \cos(2\pi f_c t), & 0 \leq t \leq T_b \\ \sqrt{\frac{E_b}{2T_b}} \cos(2\pi f_c t + \pi), & T_b \leq t \leq 2T_b \end{cases} \quad (6.182)$$

用 $s_2(t)$ 表示时间间隔 $0 \leq t \leq 2T_b$ 内发送的 DPSK 信号, 如果在时间间隔 $T_b \leq t \leq 2T_b$ 内, 发射机输入信号为二进制符号 0, 则载波相位提前 180° , 因此可将 $s_2(t)$ 定义为

$$s_2(t) = \begin{cases} \sqrt{\frac{E_b}{2T_b}} \cos(2\pi f_c t), & 0 \leq t \leq T_b \\ \sqrt{\frac{E_b}{2T_b}} \cos(2\pi f_c t + \pi), & T_b \leq t \leq 2T_b \end{cases} \quad (6.183)$$

由式(6.182)和式(6.183)可知, $s_1(t)$ 和 $s_2(t)$ 在两比特的持续时间 $0 \leq t \leq 2T_b$ 内实际上是正交的。换言之, DPSK 是 $T = 2T_b$, $E = 2E_b$ 的非相干正交调制的一种特例。因此, 利用式(6.163), 可得 DPSK 的误比特率如下:

$$P_e = \frac{1}{2} \exp\left(-\frac{E_b}{N_0}\right) \quad (6.184)$$

在相同的 E_b/N_0 下, DPSK 与非相干 FSK 相比有 3 dB 的增益。

6.9.1 DPSK 的产生和检测

下面分析 DPSK 信号的产生。在发射机输入端的差分编码过程中, 选取任意的一个比特作为起始(参考)。用 $\{d_k\}$ 表示加入了该参考比特的差分编码序列。下面介绍产生该序列过程中的一些约定:

- 如果输入二进制符号 b_k 是 1, 则保持符号 d_k 与前一比特一致。
- 如果输入二进制符号 b_k 是 0, 则将符号 d_k 改为与前一比特不同。

这样产生的差分编码序列 $\{d_k\}$ 用来对载波进行移相。其中, 符号 1 和 0 分别对应于 0 和 π 弧度的相角。差分相位编码过程如表 6.7 所示。注意, d_k 是 b_k 和 d_{k-1} 的模 2 加。

表 6.7 DPSK 信号的产生

$ b_k $	1	0	0	1	0	0	1	1
$ d_{k-1} $	1	1	0	1	1	0	1	1
差分编码序列 $\{d_k\}$	1	1	0	1	1	0	1	1
发送相位(弧度)	0	0	π	0	0	π	0	0

DPSK 发射机的框图如图 6.43(a)所示, 其中包括相互连接的一个逻辑网络和一个单比特延迟单元。从而, 可以将原始的二进制序列 $|b_k|$ 转换为差分编码序列 $|d_k|$ 。该序列是幅度编码的, 用来调制频率为 f_c 的载波, 从而产生期望的 DPSK 信号。

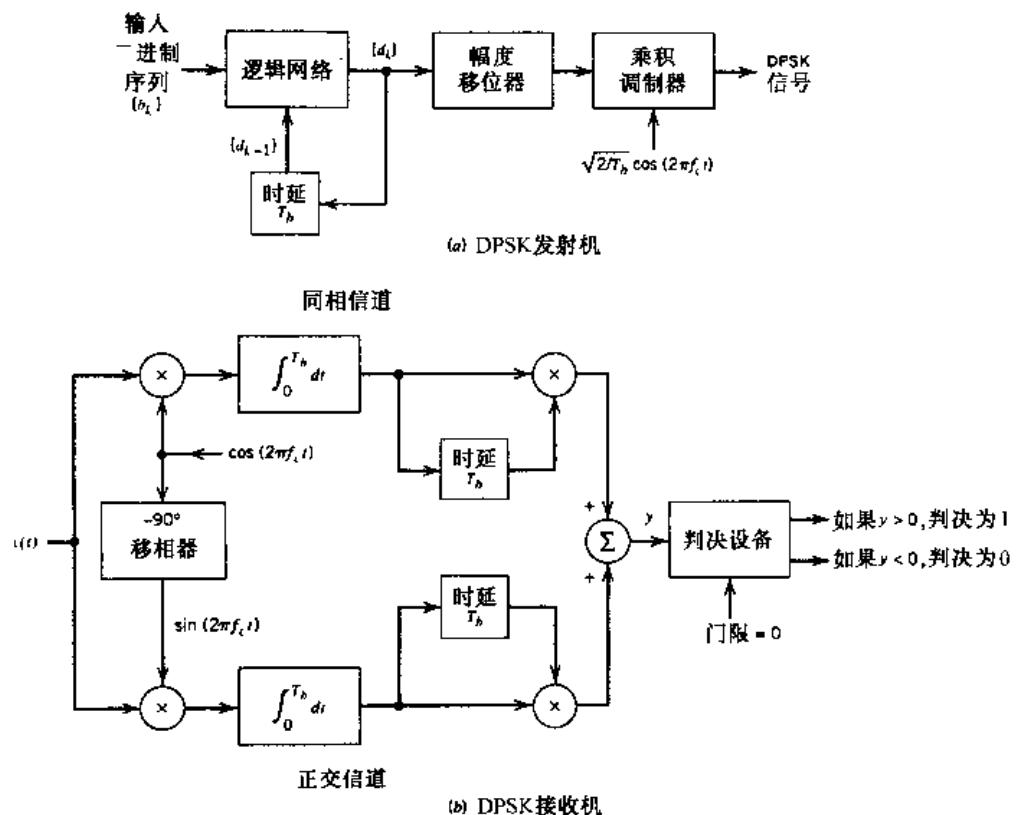


图 6.43 DPSK 发射机和接收机框图

下面,假设在对二进制 DPSK 进行差分相干检测时,载波相位是未知的。由于接收机包括同相和正交信道,所得到的信号空间图中的接收信号点分别位于 $(A \cos \theta, A \sin \theta)$ 和 $(-A \cos \theta, -A \sin \theta)$, θ 代表未知相位, A 代表幅度。图 6.44 是可能信号的几何图示。在时刻 $t = T_b$, 接收机测出坐标 (x_{I_0}, x_{Q_0}) ; 在时刻 $t = 2T_b$, 接收机测出坐标 (x_{I_1}, x_{Q_1}) 。要解决的问题是,这两点是否映射到相同的信号点。终点坐标分别为 (x_{I_0}, x_{Q_0}) 和 (x_{I_1}, x_{Q_1}) 的两个矢量 \mathbf{x}_0 和 \mathbf{x}_1 , 在内积为正值时的指向几乎完全相同,前面的假设表述就可表述为: 内积 $\mathbf{x}_0^T \mathbf{x}_1$ 是正值还是负值?

因此,有

$$x_{I_0}x_{I_1} + x_{Q_0}x_{Q_1} \stackrel{\text{判决为1}}{\geq 0} \quad \stackrel{\text{判决为0}}{\leq 0} \quad (6.185)$$

同时

$$x_{I_0}x_{I_1} + x_{Q_0}x_{Q_1} = \frac{1}{4} [(x_{I_0} + x_{I_1})^2 - (x_{I_0} - x_{I_1})^2 + (x_{Q_0} + x_{Q_1})^2 - (x_{Q_0} - x_{Q_1})^2]$$

将该恒等式代入式(6.185),并在两边同时乘以 4,得到等价式:

$$(x_{I_0} + x_{I_1})^2 + (x_{Q_0} + x_{Q_1})^2 - (x_{I_0} - x_{I_1})^2 - (x_{Q_0} - x_{Q_1})^2 \stackrel{\text{判决为1}}{\geq 0} \quad \stackrel{\text{判决为0}}{\leq 0} \quad (6.186)$$

因此,判决过程可等价地表述为: 测试点 (x_{I_0}, x_{Q_0}) 是与 (x_{I_1}, x_{Q_1}) 接近, 还是与镜像 $(-x_{I_1}, -x_{Q_1})$ 更接近。

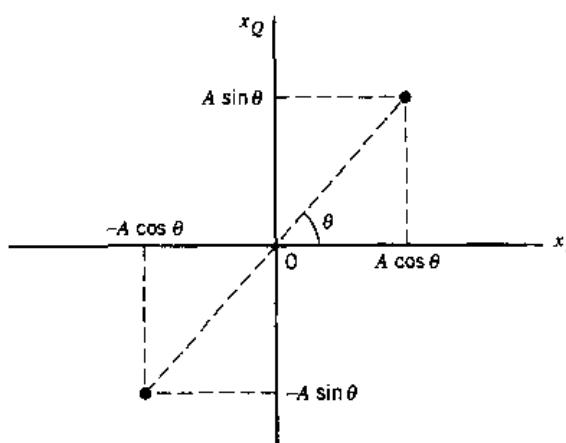


图 6.44 接收 DPSK 信号的信号空间图

因此,对二进制 DPSK 进行差分相干检测的最佳接收机^[12]如图 6.43(b)所示,这是由式(6.185)直接导出的。该最佳接收机的实现仅需要存储抽样值,从而避免了其他方式中可能需要的延迟器。按式(6.186)设计的对平方单元进行测试的接收机实现起来则较为复杂。但是,这种接收机使得要分析的两个信号在时间间隔 $(0, 2T_b)$ 内正交,从而使分析更容易进行。因此,最好的办法是用非相干正交解调分析。

6.10 单载波数字调制方式的比较

6.10.1 差错概率

表 6.8 总结了在 AWGN 信道中,相干二进制 PSK、单比特编码的传统相干二进制 FSK、DPSK、非相干二进制 FSK、相干 QPSK 和相干 MSK 的误比特率的表达式。图 6.45 为利用表 6.8 中总结的表达式绘制出的误比特率作为比特信噪比(E_b/N_0)的函数的曲线。

表 6.8 不同数字调制方式下误比特率的公式总结

调制方式	误比特率
(a) 相干二进制 PSK 相干 QPSK 相干 MSK	$\frac{1}{2} \operatorname{erfc}(\sqrt{E_b/N_0})$
(b) 相干二进制 PSK	$\frac{1}{2} \operatorname{erfc}(\sqrt{E_b/2N_0})$
(c) DPSK	$\frac{1}{2} \exp(-E_b/N_0)$
(d) 非相干二进制 FSK	$\frac{1}{2} \exp(-E_b/2N_0)$

根据图 6.45 所示的性能曲线、表 6.8 中的公式以及相关的调制方式的定义,可得以下结论:

- 所有系统的误比特率都随着 E_b/N_0 值的增大而单调下降。定义曲线都有着相似的抛物线形式。
- 对于任意的 E_b/N_0 ,相干二进制 PSK, QPSK 和 MSK 产生的误比特率,比其他几种调制方式产生的要小。

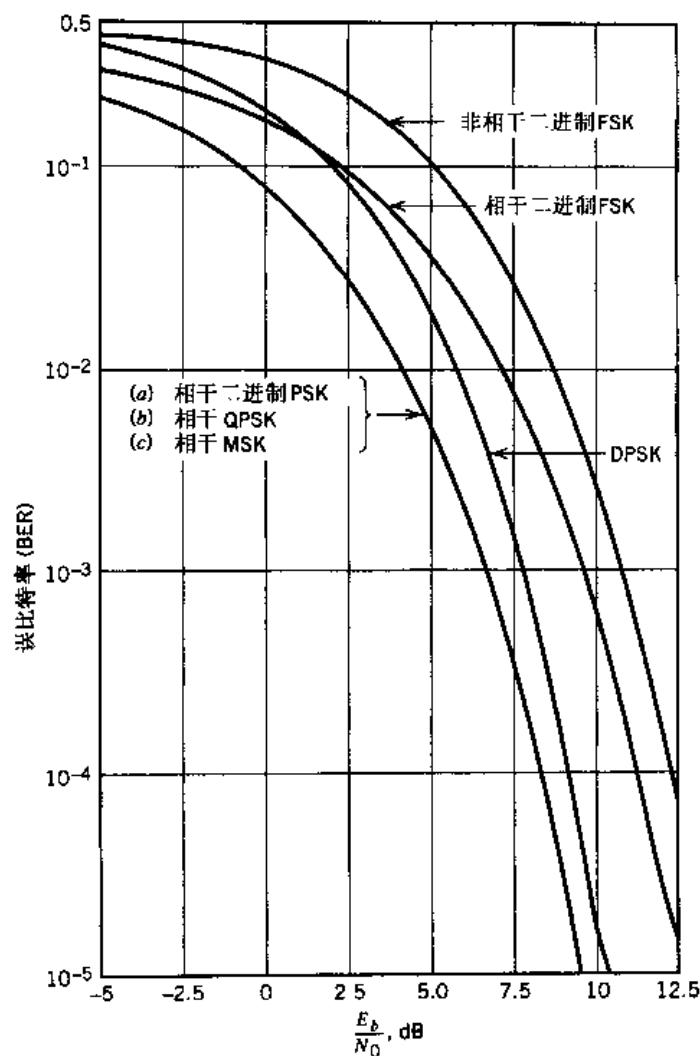


图 6.45 不同 PSK 和 FSK 方式的噪声性能比较

3. 为了得到相同的误比特率, 相干二进制 PSK 和 DPSK 需要的 E_b/N_0 值, 比传统的相干二进制 FSK 和非相干二进制 FSK 所需的 E_b/N_0 值要小 3 dB。
4. 在 E_b/N_0 取较大值时, 在相同比特速率和每比特信号能量下, DPSK 和非相干二进制 FSK 分别具有与相干二进制 PSK 和传统的相干二进制 FSK 几乎相同(差别小于 1 dB)的性能。
5. 在相干 QPSK 中, 使用两个正交载波 $\sqrt{2/T} \cos(2\pi f_c t)$ 和 $\sqrt{2/T} \sin(2\pi f_c t)$, 其中载波频率 f_c 为符号速率 $1/T$ 的整数倍。从而使得两个独立的比特流可以同时发送并在接收机中相继检测出来。
6. 在相干 MSK 中, 使用两个正交载波 $\sqrt{2/T_b} \cos(2\pi f_c t)$ 和 $\sqrt{2/T_b} \sin(2\pi f_c t)$, 这两个正交载波是在 $2T_b$ 的时间间隔内, 分别由两个相反的符号整形脉冲 $\cos(\pi t/2T_b)$ 和 $\sin(\pi t/2T_b)$ 调制而得, 其中 T_b 是比特持续时间。因此, 接收机在两个连续的比特持续时间内使用相干相位检测来恢复原始的比特流。
7. MSK 方式和 QPSK 方式不同在于, 其接收机是有存储器的。特别地, MSK 接收机根

据在两个连续比特持续时间内的观测做出判决。因此,尽管通过发送两个不同的频率来表示发射信号具有二进制形式,接收机中存储器的存在,仍使得接收机假设存在二维信号空间图。因此,包括发送的二进制符号(0或1),以及前面FSK信号的相位在内,共有四个信息点。

6.10.2 M 进制数字调制技术的带宽效率

表6.9总结了相干二进制和M进制PSK方式功率谱宽度需求的典型值,其中假设平均符号差错概率为 10^{-4} ,且系统在同样的噪声环境下工作。由此表可见,在M进制PSK信号集中,QPSK(相当于 $M=4$)在功率和带宽需求之间有较好的折中,因此QPSK在实际中得到了广泛应用。当 $M > 8$ 时,所需功率太大,因此, $M > 8$ 的M进制PSK方式并没有得到广泛应用。此外,就信号的产生和检测来说,相干M进制PSK方式比相干二进制PSK方式需要更复杂的设备,特别是当 $M > 8$ 时。(相干八相PSK被用于数字卫星通信。)

表6.9 符号差错概率为 10^{-4} 的M进制PSK与二进制FSK的功率谱带宽要求比较

M值	(带宽) _{M进制}	(平均功率) _{M进制}
	(带宽) _{二进制}	(平均功率) _{二进制}
4	0.5	0.34 dB
8	0.333	3.91 dB
16	0.25	8.52 dB
32	0.2	13.52 dB

摘自 Shanmugan(1979, p.424)。

基本上,M进制PSK和M进制QAM具有相似的频谱和带宽特性。但当 $M > 4$ 时,这两种方式具有不同的信号星座图。M进制PSK的信号星座图是圆形的,而M进制QAM的信号星座图则是矩形的。更进一步,对这两个星座图进行比较可见,M进制PSK信息点之间的距离比M进制QAM信息点之间的距离要小。图6.46显示了 $M=16$ 时两种方式之间的这一根本差别。因此,在AWGN信道中,当 $M > 4$ 时,M进制QAM的差错性能优于M进制PSK。但是,只有当信道不具有非线性时,M进制QAM才体现出较高的性能。

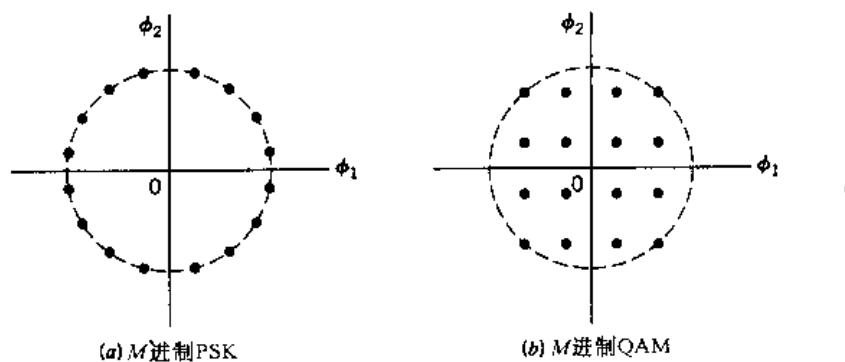


图6.46 M进制PSK和M进制QAM的信号星座图, $M=16$

对于M进制FSK,在固定的差错概率下,M的增大将带来功率需求的减小。但是,这种发射功率的减小是以增大信道带宽为代价的。换言之,M进制FSK和M进制PSK的特性相反。第9章的信息理论中将再次提到这个问题,从而使读者更加透彻地理解M进制PSK和M进制FSK之间的差异。

6.11 语音频带调制解调器

“调制解调器”是对术语调制器 - 解调器的缩写,该设备用于 PSTN 中数据的发送和接收。¹³这里所说的数据主要是由计算机或服务提供商的设备产生的数字信号。在这类应用中,调制解调器的调制器部分将输入的数字数据转换为适于通过 PSTN 的电话信道传输的标准形式;其解调器部分则接收信道输出,并将其恢复为原始的数字信号形式。而在另一类应用,即传真调制解调器,或具有传真能力的更高性能的调制解调器中,数据可以是文本、图形、照片或兼而有之。在这类应用里,需传输的文件被编码为一系列压缩图形元素——像素。这些像素将按预先规定的调制标准进行调制,然后利用电话信道进行传输。当传真调制解调器工作在接收模式下时,调制解调器的接收机部分对接收到的模拟信号进行处理,并将代表解调信号的二进制数据解压缩为与原始发送图像相似或完全相同的形式。在下文中,我们将着重讨论通过 PSTN 在用户和因特网服务商(ISP)之间提供通信的调制解调器。

PSTN 在传统上被看成是模拟网络。但实际上,我们所知道的 PSTN 已经几乎是完全的数字网了。在几乎所有的应用中,PSTN 中惟一保持了模拟性(并将继续保持许多年)的模块就是本地环路。本地环路就是从用户到中心局(CO)的较短连接。因此,根据 PSTN 的应用方式的不同,我们可以定义两种不同的调制解调器配置,也就是下面将要介绍的对称配置和非对称配置。

6.11.1 调制解调器的对称配置

设计调制解调器的最简单方法就是把整个的 PSTN 看成一个如图 6.47(a)所示的线性模拟网络。由第 3 章可知,由于为传送语音信号而采用了 PCM,PSTN 几乎是全数字的。在这样的背景下,调制解调器在通过 PSTN 发送和接收信号时需要用到模 - 数或数 - 模转换器。图 6.47(a)所示调制解调器配置的“对称性”就体现在两边的调制解调器是完全一样的,并且下行数据速率(从 ISP 到用户端)与上行数据速率(从用户端到 ISP)相同。

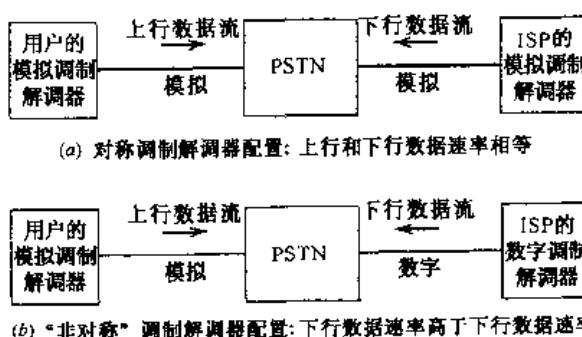


图 6.47 调制解调器的两种配置

图 6.47(a)所示的调制解调器的对称配置中可采用的调制解调器类型很多,数据速率从 300 bps 直到 36 600 bps,具体见表 A6.7 对标准调制解调器的总结。调制解调器最初采用的是 FSK,适用于较低的数据速率。随着电话信道传输数据的增加,越来越复杂的调制解调技术被采用,以便更好地利用电话信道的信息容量。

例如普遍采用的 V.32 调制解调器标准。该标准具有如下特点:

载波频率 = 1800 Hz

调制速率 = 2400 波特

数据速率 = 9600 bps

之所以采用 9600 bps 的信号数据速率是为了适应高信噪比的情况。V.32 标准详细说明了两种可选的调制方式：

非冗余编码。在该方式下，输入数据流被分割为四位组（四个连续比特构成的组），然后以 16-QAM 的方式通过电话信道传输。在每个四位组里，高位的输入二位组采用 PM，低位的输入二位组则采用 AM。首先讨论 PM。为了使接收机仅需对相位差进行检测，实用中通常选用差分 PM。实现方法是采用图 6.48(a) 所示的差分编码器。这种编码器由一个只读存储器和一对延迟单元组成。用 $Q_{1,n} Q_{2,n}$ 代表高位的输入二位组的当前值，而用 $I_{1,n-1} I_{2,n-1}$ 代表编码器输出的相应二位组的先前值。根据二位组 $Q_{1,n} Q_{2,n}$ 和 $I_{1,n-1} I_{2,n-1}$ ，差分编码器产生二位组 $I_{1,n} I_{2,n}$ ，而这又将引起调制信号的相位变化。按逆时针方向的相位变化符合表 6.10 中的格雷编码方式。不难注意到，相位变化完全由输入二位组 $Q_{1,n} Q_{2,n}$ 决定。在讨论差分 PM 时，还会遇到的一个问题是：对二维信号空间的四个象限需要用一个码来识别。通过采用图 6.49 所示的格雷码就能够解决这个问题。

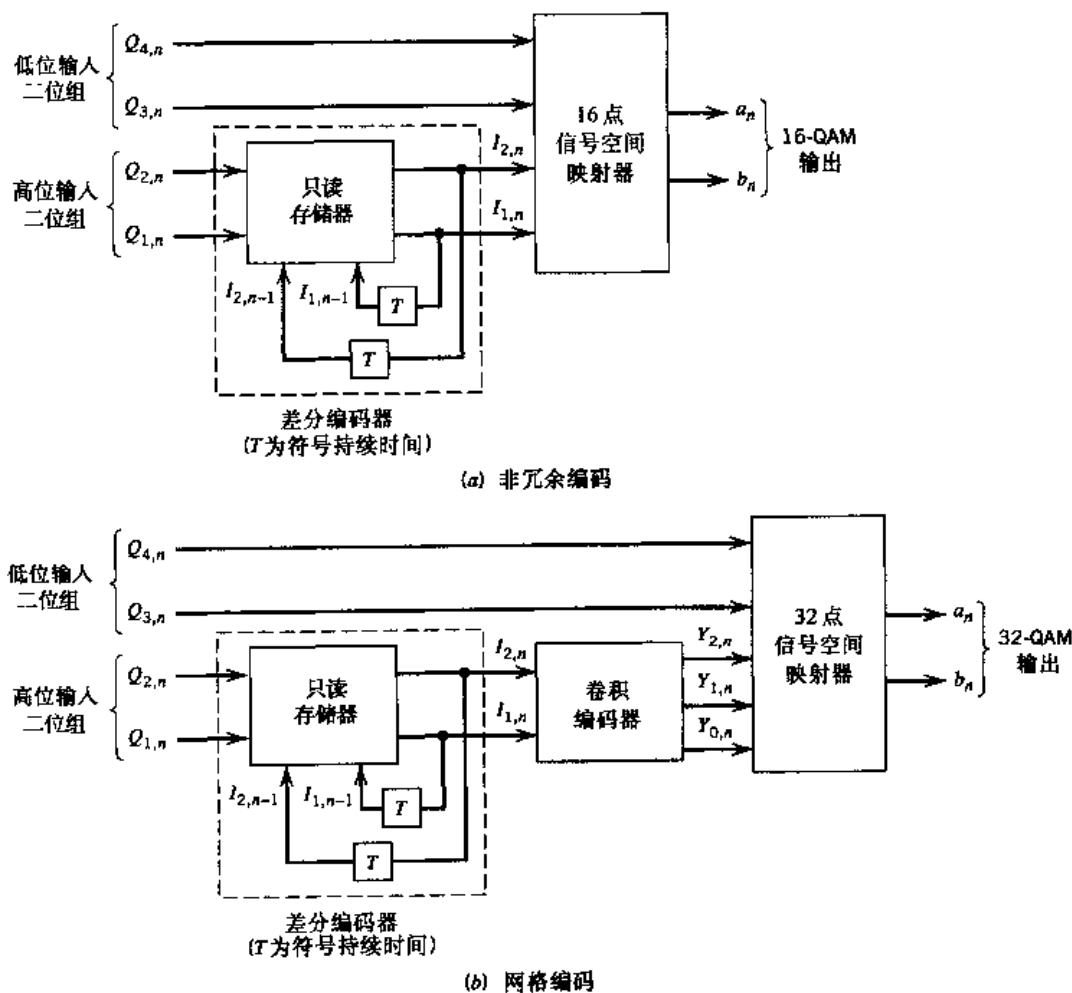


图 6.48 V.32 调制解调器框图

表 6.10 V.32 调制解调器对应于不同输入二进制数据的差分编码相位变化

当前输入二位组 $Q_{1,n}$ $Q_{2,n}$		相位变化
0	0	90°
0	1	0°
1	0	180°
1	1	270°

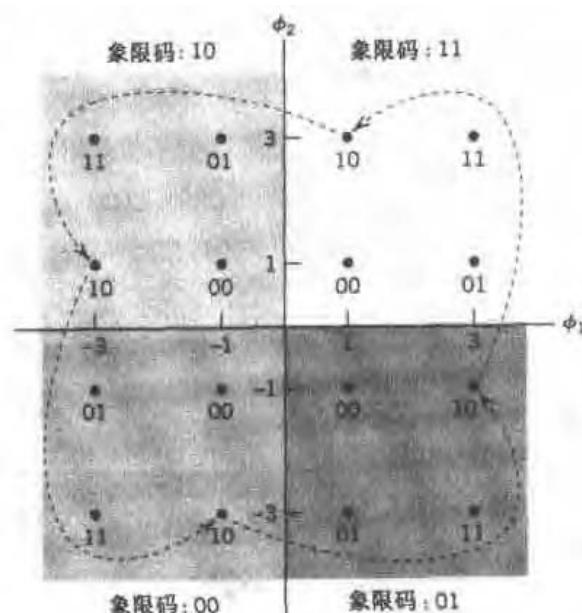


图 6.49 V.32 调制解调器四个象限的格雷码及各象限的二位组。虚箭头描述了 90°旋转不变性

下面考虑 AM。假设在第一象限内,需要对 $Q_{3,n}Q_{4,n}$ 代表低位输入二位组的四个可能取值进行编码。通过对图 6.49 阴影显示的第一象限中的四个信息点进行格雷编码就可以解决这个问题。

最后,需要解决的问题是采用差分编码所要满足的 90°旋转不变性。这种恒定性意味着 M 进制 QAM 星座图无论是否进行了编码,在旋转了 90°的整数倍后依然保持不变。因此,当本地振荡器相位与载波相位相差 90°的整数倍时,接收机仍然可以正确地对发送的信息序列进行译码。通过对其余三个象限的信号点按照图 6.49 所示的方法进行格雷编码就可以满足上述旋转恒定性要求。图 6.49 中的虚线箭头就表明了 90°旋转不变性。

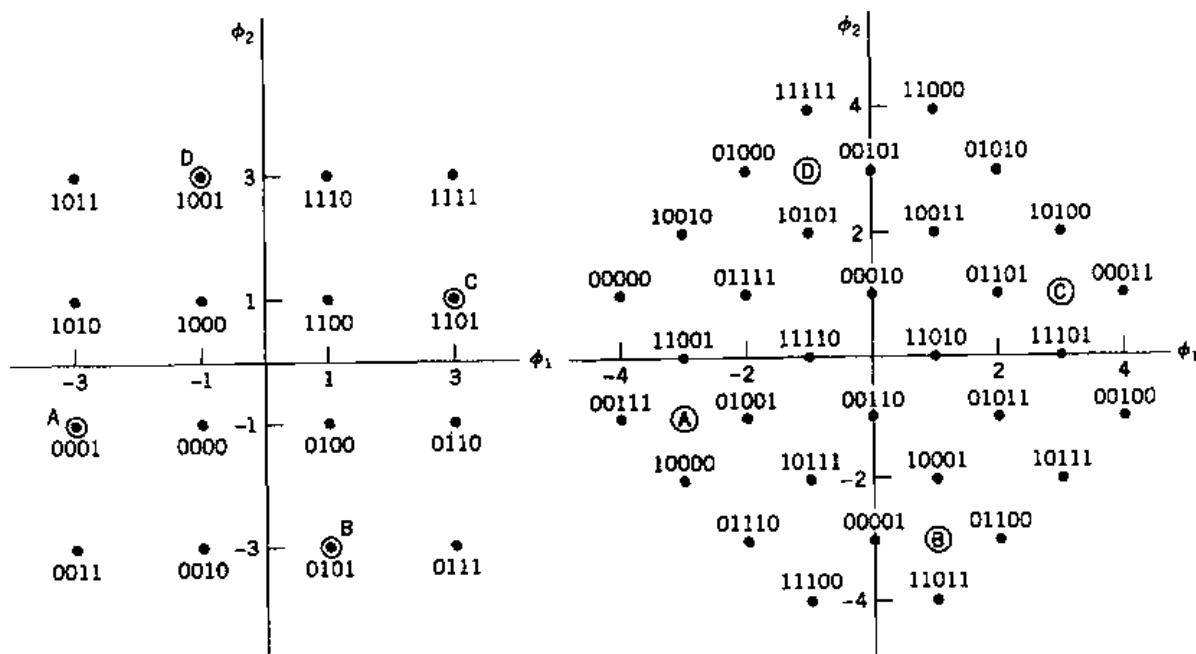
将上述两种方法相结合,即采用幅度相位联合调制,就可以得到图 6.17(a)所示的 16-QAM 星座图。图 6.50(a)与图 6.17(a)相同。因此,编码系统由一个差分编码器和一个 16 点信号空间映射器连接而成,如图 6.48(a)所示。这样配置的 V.32 调制解调器的星座图里共有 16 个点,因此发送的 4 比特码字没有冗余位,所以这种配置称为是非冗余的。

下面举例说明这种独特的 V.32 调制解调器是如何工作的。假设当前 4 位输入组为 1001, 调制解调器先前输出的二位组为 11。因此可得

$$\begin{aligned}Q_{1,n}Q_{2,n} &= 10 \\Q_{3,n}Q_{4,n} &= 01 \\I_{1,n-1}I_{2,n-1} &= 11\end{aligned}$$

根据图 6.49 中指定的对四个象限的编码方式,先前的输出二位组 11 意味着调制器先前处于第一象限。由表 6.10 可知,由于相应的输入二位组是 10, 调制器将沿逆时针方向发生 180° 的相位变化,从而转入二位组 00 对应的第三象限。最后,利用低位的二位组 $Q_{3,n}Q_{4,n}$ 的当前值 01, 调制器输出坐标为 $a_n = -3$ (沿 ϕ_1 轴)和 $b_n = -1$ (沿 ϕ_2 轴)的 QAM 信号。该输出对应于码字 0001。

当信噪比不够大时,V.32 调制解调器转为使用 QPSK 模式,以较低的速率(4800 bps)工作。在此工作模式下,调制解调器的四个状态如图 6.50(a)中的点 A, B, C, D 所示。



(a) V.32 调制解调器采用非冗余编码时的信号星座图

(b) V.32 调制解调器采用网格编码时的信号星座图

图 6.50 V.32 调制解调器采用不同编码方式时的信号星座图

网格编码。网格编码是一种前向纠错方式,其中将编码和调制看成是结合在一起的整体,而不是两个分离的过程。图 6.48(b)为采用网格编码的 V.32 调制解调器的编码系统。输入数据流被分割为四位组,但与非冗余编码不同的是,这些四位组以 32-QAM 信号的形式通过信道传输。

如图 6.48(b)所示,网格编码器中采用卷积编码器对差分编码器的输出进行处理(卷积编码器将在第 10 章介绍)。但为了与差分编码(如 90° 旋转不变性)相适应,V.32 调制解调器中对卷积编码方式的选择是有限制的。实际上,使用线性卷积编码器是不合适的。因此这里所用的卷积编码器都必须是非线性的,^[14]见习题 10.30。

采用网格编码的 V.32 调制解调器的数据编码过程包括三个步骤:

1. 图 6.48(b)所示的差分编码器根据当前输入二位组 $Q_{1,n}Q_{2,n}$ 和先前的差分编码二位组 $I_{1,n-1}I_{2,n-1}$ 产生二位组 $I_{1,n}I_{2,n}$ 。

2. 将差分编码器当前二位组 $I_{1,n}, I_{2,n}$ 输入到图 6.48(b) 所示的卷积编码器, 可得到 3 位的输出。其中一位是奇偶校检位, 用 $Y_{0,n}$ 表示。 $Y_{0,n}$ 的值由卷积编码器根据另两位 $I_{1,n}$ 和 $I_{2,n}$ 计算而得。
3. 卷积编码器的输出 $Y_{0,n}, Y_{1,n}$ 和 $Y_{2,n}$, 以及低位的输入二位组 $Q_{3,n} Q_{4,n}$ 都被送入图 6.48(b) 所示的信号空间映射器。信号空间映射器将选择图 6.50(b) 所示 32 点星座图的一个状态作为调制解调器的输出。

奇偶校检位 $Y_{0,n}$ 使得采用网格编码的调制解调器比采用非冗余编码的调制解调器更不易受到信道干扰的影响, 并且没有增加对信道带宽的需求。定量地说, 网格编码相对于 16-QAM 产生了 4 dB 的编码增益。编码增益说明, 为获得同样的噪声性能, 未采用网络编码的调制解调器每比特数据需要增加的信号能量。

但是, 为了实现网格编码的上述优点, 信噪比必须足够大。否则, V.32 调制解调器将切换到 QPSK 工作模式, 如图 6.50(b) 所示(用标注有 A, B, C, D 的 4 个状态表示)。在该工作模式下, 调制解调器的数据速率将降低为 4800 bps。

6.11.2 调制解调器的非对称配置

为了更有效地利用 PSTN, 我们可以把它看成是一个非线性的几乎全数字的网络。特别地, 由于 ISP 是数字实现的, 在 ISP 调制解调器中不需要进行模 - 数转换。这就意味着 ISP 和 PSTN 之间的通信是全数字的, 如图 6.47(b) 所示。但用户的调制解调器必须是模拟的, 因为本地环路是模拟的。这就要求用户的调制解调器每次向 PSTN 发送和接收信号时要分别使用模 - 数和数 - 模转换器。图 6.47(b) 中调制解调器的配置为“非对称”的, 因为下行信号数据速率可比上行信号数据速率高出很多, 后面将对此进行解释。

如前所述, 数字 PSTN 是采用 PCM 来传输语音信号的。和当前讨论相关的 PCM 系统的特性如下(见第 3 章):

- 数据信号速率为 64 Kbps。其中, 抽样速率为 8 kHz, 每个语音抽样由一个 8 位码字表示。
- 使用 15 折线压扩律(例如 $\mu = 255$ 时的 μ 律算法)在发射端对语音信号进行压缩, 并在接收机进行解压。

由第 3 章对 PCM 的讨论可知, 量化只影响模 - 数转换而不影响数 - 模转换。这一发现对非对称调制解调器的优化设计有深刻的影响。

假设在 ISP 端的数字调制解调器与 PSTN 的数字部分之间没有模 - 数转换器, 同时该调制解调器中的发射机经过特别设计以便最有效地利用数字 PSTN 非均匀分布的 256 个(离散的)门限值。由于数 - 模转换根本不受量化噪声的影响, 所以 ISP 的数字调制解调器发送的信息总是可以无损地到达用户端的模拟调制解调器。根据上述结论, 理论上能以 64 Kbps 的数据速率通过数字 PSTN 从 ISP 向用户传输数据。但是, 由于 PSTN 系统内在的限制, 使得能够达到的数据速率降低到 56 Kbps, 接下来我们将对此进行介绍。

6.11.3 数字调制解调器

从第 3 章对 PCM 音频信道的介绍可知, 数字调制解调器的设计会受到三个无法控制的因

素限制,即:

1. 抽样速率 $f_s = 8 \text{ kHz}$ 。
2. 一组 $M = 256$ 的可用门限值,用于压缩器中(即压扩器的发射机部分)。
3. 带宽约为 3.5 kHz 的基带(抗混叠)滤波器,位于 PCM 发射机的前端。

根据这些限制,可将设计数字调制解调器的基本思路总结如下:

在数字调制解调器的输入端设计一个信号 $s(t)$,并使 $s(t)$ 以速率 $f_s = 8 \text{ kHz}$ 抽样而得到的每个抽样值分别对应于压缩器的 $M = 256$ 个门限值之一,且发射信号满足无 ISI 的奈奎斯特准则。

无 ISI 奈奎斯特准则在第 4 章中介绍过。

6.11.4 数字调制解调器的一种实现

由于 PCM 发射滤波器的带宽约为 3.5 kHz 而不是 4 kHz (抽样速率 f_s 的一半),信号的设计变得非常困难。这个带宽限制的直接含义是,按照无符号间干扰的奈奎斯特标准,可在每秒仅产生 $2 \times 3500 = 7000$ 个独立抽样点,而不是 8000 个。那么,如何把每秒的 7000 个独立抽样点放入前述的每秒 8000 个抽样点的结构中呢?

为了回答这一基本问题,需要用到抽样定理的循环不均匀等效形式。特别地,考虑图 6.51 所示的情况,其中抽样点被分为组,每组包括 N 个均匀分布的抽样点,且各组的循环周期都是 $(N+1)T_s$ 秒,其中 $T_s = 1/f_s$ 。图 6.51 对应于 $T_s = 125 \mu\text{s}$ 和 $N = 7$ 的情况。图 6.51 所示的非均匀分布的瞬时抽样值可写为

$$\begin{aligned} t_{k,l} &= t_k + (N+1)lT_s \\ &= (k-1)T_s + (N+1)lT_s, \quad k = 1, 2, \dots, N \\ l &= 0, \pm 1, \pm 2, \dots \end{aligned} \quad (6.187)$$

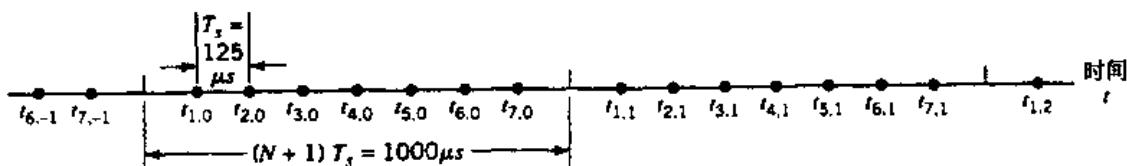


图 6.51 一组 N 个均匀分布的抽样点。每 $(N+1)T_s$ 秒重复一次

将带限信号 $s(t)$ 定义如下:^[15]

$$s(t) = \sum_{l=-\infty}^{\infty} \sum_{k=1}^N s(t_{k,l}) \psi_k(t - (N+1)lT_s) \quad (6.188)$$

其中,插值函数 $\psi_k(t)$ 的定义为

$$\psi_k(t) = \operatorname{sinc}\left(\frac{t - t_k}{(N+1)T_s}\right) \prod_{q=1, q \neq k}^N \frac{\sin\left(\frac{\pi}{(N+1)T_s}(t - t_q)\right)}{\sin\left(\frac{\pi}{(N+1)T_s}(t_k - t_q)\right)} \quad (6.189)$$

对于 $N=7$ 的情况, 根据式(6.189), 得到图 6.52 所示的 7 个标准脉冲。其中将时间按照抽样时间间隔 T_s 进行了归一化。这 7 个脉冲具有如下特性:

- 每个标准脉冲都进行了归一化, 从而可得

$$\psi_k\left(\frac{t_k}{T_s}\right) = \psi_k(k-1) = 1 \quad k = 1, 2, \dots, 7$$

需注意第 k 个脉冲的峰值并不出现在 $t_k = (k-1)T_s$ 时刻。

- 对于 $k = 1, 2, \dots, 7$, 脉冲 $\psi_k(t/T_s)$ 除了在 $t = (-1)[\text{模}(N+1)]$ 的时刻外, 在 $t \neq (k-1)T_s[\text{模}(N+1)]$ 的时刻都经过 0。

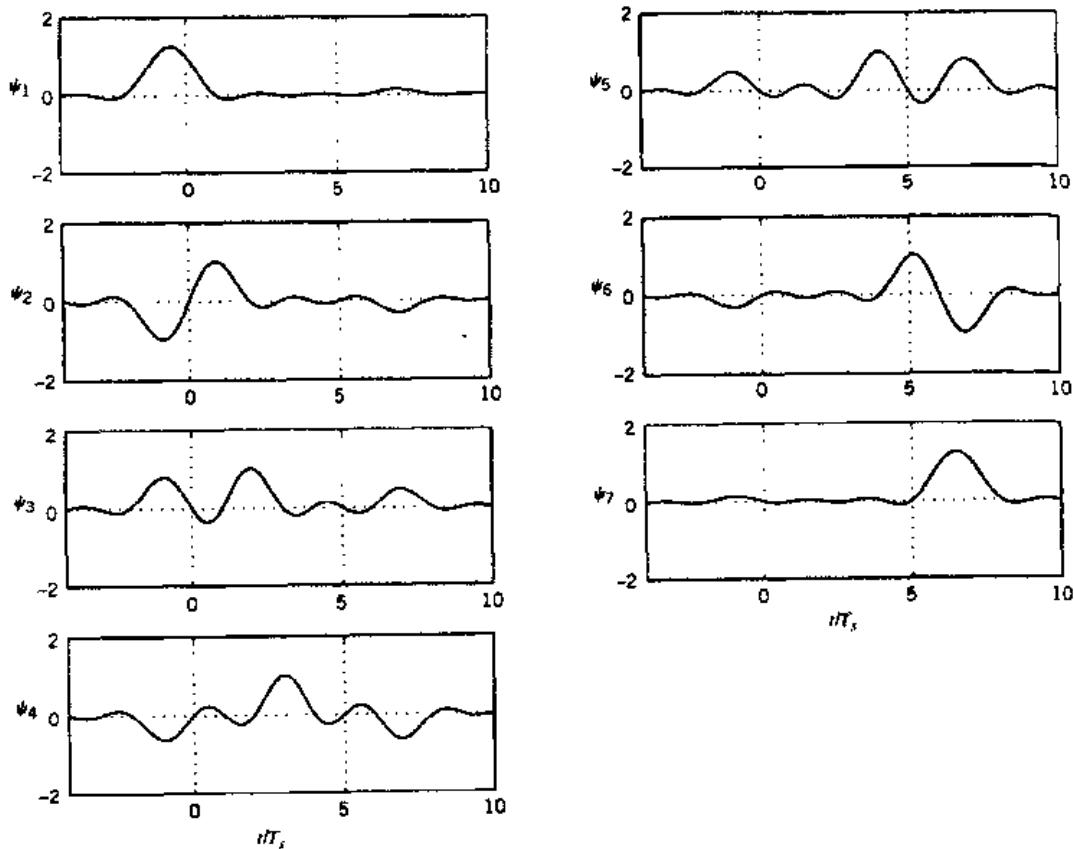


图 6.52 标准脉冲 $\psi_k(t)$ 的数字调制解调器波形, $k = 1, 2, \dots, 7$

因此, 数字调制解调器的信号由循环非均匀脉冲 AM 组成。每组 8 个抽样点中的 7 个均匀分布的抽样点的幅度由输入数据流决定, 并与 PCM 发射机中压缩器的门限值相对应。实际上, 这 7 个抽样值就是在每 $1000 \mu\text{s}$ 内, 对通过 PSTN 传输的输入数据流进行抽样所得的独立抽样值。更进一步, 这 7 个抽样值将无 ISI 的数据传送给接收机。剩下的第 8 个抽样值完全由独立抽样值和系统预设决定。它们并不携带信息, 因此在接收机端被舍弃。因此, 数字调制解调器能够以 56 Kbps 的速率通过 PSTN 几乎无差错地传输数字数据:

$$7 \times 1000 \times \log_2 256 = 56000 \text{ bps}$$

现在介绍最后一点。可将标准脉冲 $\psi_k(t)$ 构造成在速率大于 $1/t$ 时开始衰减。为此, 简单

地将式(6.189)中的 sinc 函数替换为具有与第 4 章所介绍的滚降相似的奈奎斯特脉冲。

6.11.5 数字调制解调器的另一种实现

上述数字调制解调器是双向的,且假设数据链路的两端都是模拟的。当数据链路的一端是数字的,且允许采用非对称的数据速率时,数字调制解调器的设计可以变得非常简单。^[16]

考虑由 8 位组(8 比特编码字)组成的数据序列通过 PSTN 进行传输的情况。根据 PSTN 所处地域的不同,这些 8 位组将被看成分别依照 μ 律或 A 律编码的语音编码码字。因此,通过采用模拟调制解调器的数 - 模转换器,可产生定义如下的连续时间信号:

$$s(t) = \sum_k a(c_k)g(t - kT_s) \quad (6.190)$$

其中, c_k 为数据序列中的第 k 个 8 位组, $a(c_k)$ 为由压扩律规定的代表电平, T_s 为抽样时间间隔(等于 $125 \mu\text{s}$), $g(t)$ 为插值函数,带限为仅允许 $1/2T_s$ 以下(或大约 4 kHz 以下)的频率值,以满足抽样定理重构部分的要求(见 3.2 节)。

在 PSTN 的标准过程中,信号 $s(t)$ 代表了重构的语音信号。但对于输入数据, $s(t)$ 看上去却很像噪声。从通信理论的观点来看,式(6.190)中的信号 $s(t)$ 可以看成是 PAM 信号。这就是设计数字调制解调器的理论基础。特别地,设计中可利用与模拟调制解调器相似的星座图,只是该星座图是由一维 PCM 符号,而不是二维 QAM 符号构造而得的。

通常,由于以下因素,数字调制解调器能够达到的数据速率限制在 56 Kbps :

1. PSTN 的压扩器内部各级非常接近(如表 3.4 所示),因此容易受残留的 ISI 和调制解调器的均衡器噪声的影响。
2. 由于 PSTN 的内部原因,数据流的最低位将被舍弃。这种“位舍弃”可达(但通常是小于) 8 Kbps ,并且总是周期性的。

6.11.6 模拟调制解调器

和数字调制解调器不同的是,模拟调制解调器的噪声性能主要受 PCM 压扩器依照 μ 律或 A 律操作过程中的量化噪声影响。一个通信质量较好的典型 PCM 音频信道的信噪比在 34 dB 到 38 dB 之间。影响模拟调制解调器工作的其他信道干扰,包括由抗混叠滤波器和插值滤波器引起的带宽限制,其典型值为 3.5 kHz 。

模拟调制解调器的一种改进选择是标准 V.34 调制解调器,其速率可以高达 33.6 Kbps 。该调制解调器的基本设计原理包含五个与众不同的特点:^[17]

1. 960-QAM 超星座图

由于信号星座图由四个星座图组成:图 6.53 所示中有 240 个信息点的 QAM 星座图,及其分别旋转 90° , 180° 和 270° 度后得到的星座图。因此该星座图称为超星座图或者嵌套星座图。

2. 自适应的带宽

发射机通过发送一组音频信号来探测信道,在信道输出端测量出以频率为自变量的信噪比函数。因此,调制解调器可以根据探测结果以及可用码元速率来选择合适的载波频率和带宽。

3. 自适应的比特率

当误比特率较低(在 10^{-6} 到 10^{-5} 之间)时, 在接收机训练期间, 调制解调器根据接收机对最大比特率的估计来选择比特率。

4. 网格编码

这种差错控制编码技术用来提供约为 3.6 dB 的有效编码增益。其中一种网格编码可以提供约为 4.7 dB 的有效编码增益。

5. 判决反馈均衡

为了充分利用可用的电话信道带宽, 包括接近通带边缘的(衰减可达 10 dB 到 20 dB 的)频率值, 可采用判决反馈均衡器(判决反馈均衡器在第 4 章中讨论过)。但并不是将编码和判决反馈均衡器简单结合在一起就可以了, 因为判决反馈要求即时判决, 而编码包括译码延迟。为解决这个问题, 可将判决反馈均衡器的反馈部分移至发射机端。这可以通过采用 Tomlinson-Harashima 预编码来实现。(这种预编码均衡方式在第 4 章注释[12]中曾简单介绍过。)

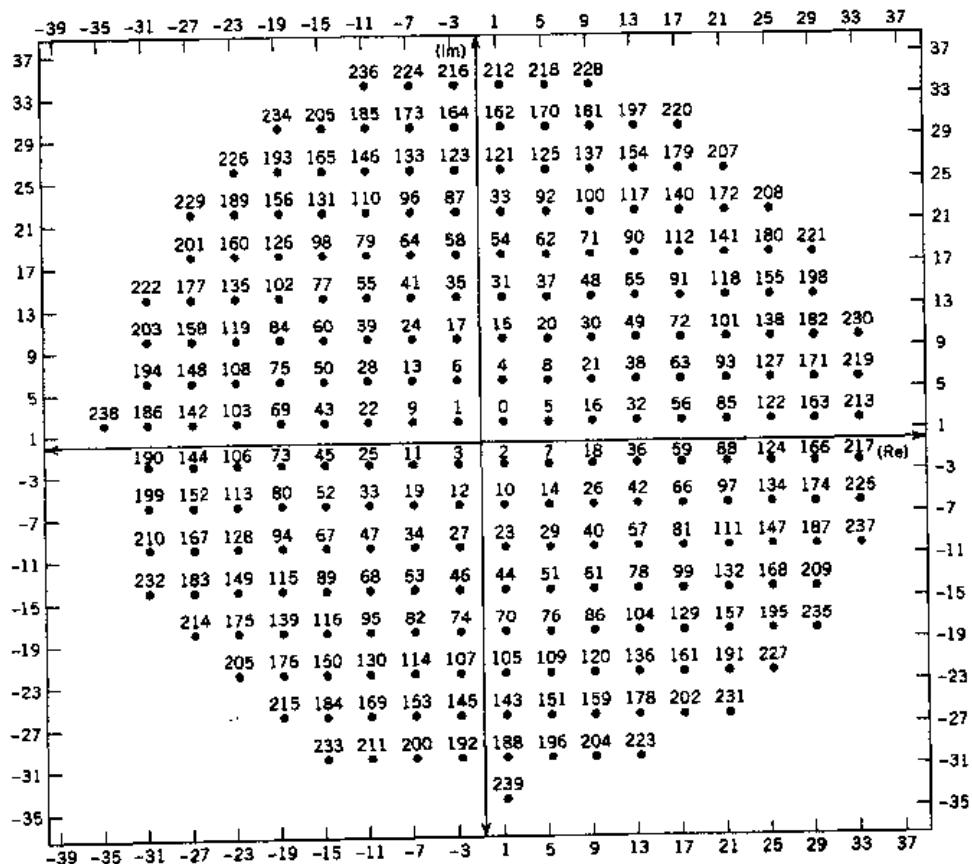


图 6.53 240 信号点的 V.34 调制解调器的四分之一超星座图。将这些点分别旋转 $0^\circ, 90^\circ, 180^\circ, 270^\circ$ 后, 再结合在一起, 就能够得到完整的超星座图(经 IEEE 允许摘自 Forney 等人的著作, 1996)

6.11.7 V.90 调制解调器

V.90 调制解调器标准包括数字和模拟两部分。ISP 端的数字调制解调器是以前面讨论的第二种实现为基础的, 其中发送数据的下行速率为 56 Kbps。用户端的模拟调制解调器采用

V.34 调制解调器标准,其中发送数据的上行速率为 33.6 Kbps。这两个差别较大的数据速率体现了 V.90 调制解调器的非对称特性。

由于 V.90 调制解调器 56 Kbps 下行数据速率的显著特点,因此适用于在因特网上以接近 ISDN 的速度,提供清晰的 Web 图片、音频和视频下载服务。

6.12 多信道调制

4.8 节中介绍的 ADSL,是能够在现有的双绞线电话线上实现百万数量级速率的数据传输系统。特别地,ADSL 的下行速率可以高达 9 Mbps,而上行速率可达到 1 Mbps。这样数据速率能够很好地满足访问因特网的需求。正如在 4.8 节提到的,IP 协议要求上行比特率应约为下行比特率的十分之一。设计 ADSL 的难点在于找到信道利用率较高的线路码。6.4 节中讨论的 CAP 就提供了这样一种方法。另一个方法是采用性能与 CAP 相同的,被称为离散多音的调制技术。离散多音是一种多信道调制技术^[18],其中将调制器特性作为测得的信道特性的函数。要讨论离散多音,就必须首先了解多信道调制。

多信道调制的基本思想为分割和解决,这也是一个普遍应用的工程学原理。根据这一原理,一个复杂问题被分割为多个较为简单的问题,通过将这些简单问题的解决方法相结合,从而得到该复杂问题的解决方法。现在需要解决的复杂问题是存在严重 ISI 的宽带信道中的数据传输问题。根据上述原理,可先将该复杂问题转化为较为简单的问题,如 AWGN 信道中的数据传输。因此,可将多信道调制的本质总结如下:

通过使用先进的信号处理技术,将复杂信道中的数据传输问题进行转化,变为已知数据流通过大量子信道并行传输的问题。其中每个子信道都可以看成是 AWGN 信道。

显然,总的数据速率是并行工作的所有子信道数据速率之和。

6.12.1 AWGN 信道的容量

根据香农的信息容量理论,AWGN 信道(不存在 ISI)容量的定义为

$$C = B \log_2(1 + SNR) \text{ b/s} \quad (6.191)$$

其中,B 为信道带宽,SNR 为信道输出端测得的信噪比。在第 9 章中将对该式进行证明。对于给定的信噪比,假设采用复杂度足够高的编码系统,就能够通过带宽为 B 的 AWGN 信道,以不超过 C bps 的速率和任意小的差错概率进行数据传输。可等价地将信道容量 C 表达为如下的比特/传输(或比特/信道)的形式:

$$C = \frac{1}{2} \log_2(1 + SNR) \quad \text{比特/传输} \quad (6.192)$$

实际上的编码系统只能以低于最大可能速率 C 的速率 R 来传输数据。因而,对于一个在足够小的符号差错概率下工作的系统,需要引入用 Γ 表示的信噪比间隙或间隙。间隙是可允许的符号差错概率 P_e 与编码系统的函数。间隙用来衡量式(6.192)所示理想传输系统的有效性。用 C 代表理想编码系统的容量,R 代表实际编码系统的容量,则间隙的定义如下:

$$\begin{aligned}\Gamma &= \frac{2^{2C} - 1}{2^{2R} - 1} \\ &= \frac{\text{SNR}}{2^{2R} - 1}\end{aligned}\quad (6.193)$$

等价地,有

$$R = \frac{1}{2} \log_2 \left(1 + \frac{\text{SNR}}{\Gamma} \right) \quad \text{比特/传输} \quad (6.194)$$

例如,对于 $P_e = 10^{-6}$ 的编码 PAM 或 QAM, 间隙 Γ 恒等于 8.8 dB。通过采用某些编码技术(如第 10 章将讨论的网格编码), 间隙 Γ 可减小至 1 dB。

用 P 代表发射信号功率, σ^2 代表带宽 B 中测得的信道噪声方差, 则信噪比为

$$\text{SNR} = \frac{P}{\sigma^2}$$

其中

$$\sigma^2 = N_0 B$$

因此,可将可获得的数据速率定义如下:

$$R = \frac{1}{2} \log_2 \left(1 + \frac{P}{\Gamma \sigma^2} \right) \quad \text{比特/传输} \quad (6.195)$$

有了这个公式,就可以定量地介绍多信道调制了。

6.12.2 连续时间信道分割

考虑一个具有任意频率响应 $H(f)$ 的线性宽带信道(例如双绞线)。假设用如图 6.54 所示的阶梯函数来逼近幅度响应 $|H(f)|$, 图中的 Δf 代表了每阶的宽度。在极限情况下, 随着频率增量 Δf 逼近 0, 信道的阶梯近似也逼近实际的 $H(f)$ 。在进行每一步逼近时, 都可以假设信道是不存在 ISI 的 AWGN 信道。这样, 发送单个宽带信号的问题就转化为发送一组窄带正交信号的问题。每个窄带正交信号都有对应的载波, 这些载波是通过 M 进制 QAM 之类的(频谱有效的)调制技术产生的, 本质上主要的干扰源只有加性高斯白噪声。这也就意味着, 根据香农的信息容量理论, 可对每个带宽为 Δf 的子信道中进行的数据传输进行优化, 且与各子信道的优化与所有其他子信道都是相互独立的。因此, 实际信号处理过程中的宽带信道均衡问题, 就转化为在许多相连但不相交的窄带子信道上的复用和解复用问题。尽管当子信道数量较多时, 所得的多载波系统复杂度其实很高, 但通过采用 VLSI 技术, 就能够以较低的成本来解决整个系统的实现问题。

图 6.55 为最基本的多信道数据传输系统框图。其中采用 QAM, 目的是为了利用其频谱效率高的特点。输入的二进制数据流首先被送入分接器(图中未画出), 从而产生一组 N 个子流。每个子流代表一个二元的子符号序列, 在符号时间间隔 $0 \leq t \leq T$ 内, 该二元的子符号可表示为

$$(a_n, b_n), \quad n = 1, 2, \dots, N$$

其中 a_n 和 b_n 为子信道 n 对应的两个并列通道的元素值。

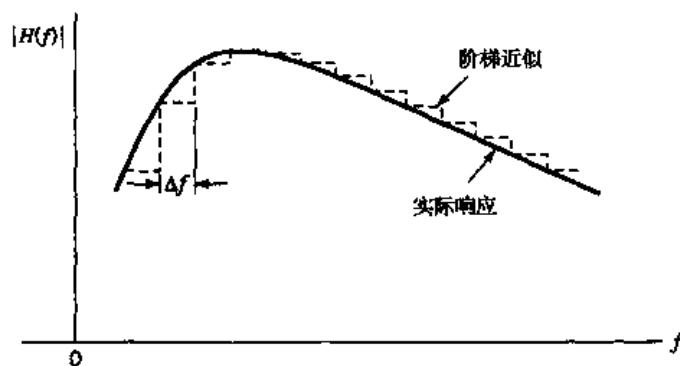
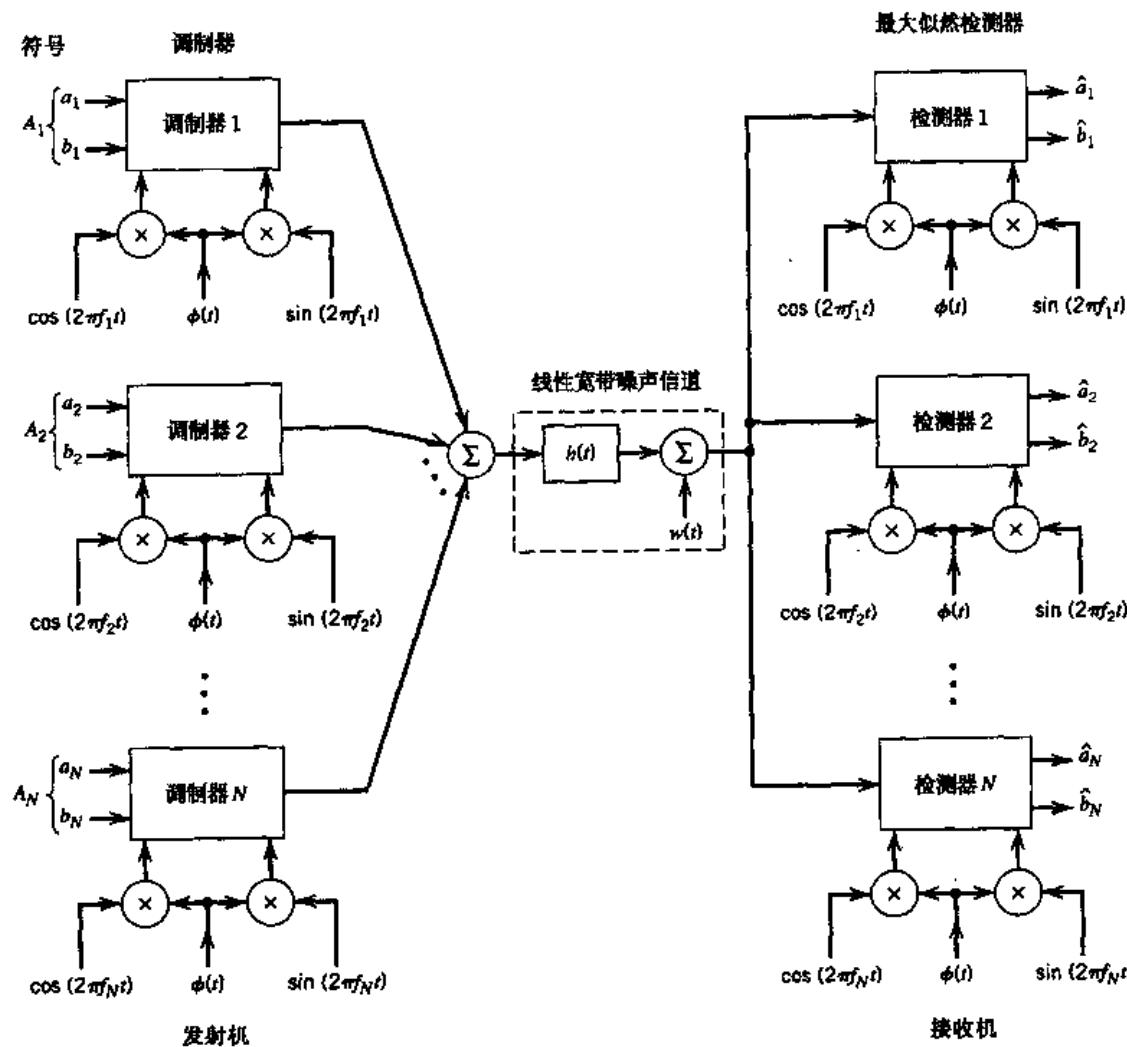
图 6.54 任意幅度响应 $|H(f)|$ 的阶梯近似;仅显示了正频率部分

图 6.55 多信道数据传输系统框图

因此,正交幅度调制器的通带基本函数由下列函数对定义:

$$\{\phi(t) \cos(2\pi f_n t), \phi(t) \sin(2\pi f_n t)\}, \quad n = 1, 2, \dots, N \quad (6.196)$$

其中第 n 个调制解调器的载波频率 f_n 为符号速率 $1/T$ 的整数倍,即

$$f_n = \frac{n}{T}, \quad n = 1, 2, \dots, N$$

而低通函数 $\phi(t)$ 为 sinc 函数:

$$\phi(t) = \sqrt{\frac{2}{T}} \operatorname{sinc}\left(\frac{t}{T}\right), \quad -\infty < t < \infty \quad (6.197)$$

此处定义的通带基本函数有如下我们所期望的性质(证明过程见习题 6.41):

性质 1

对于每个 n , 两个正交调制的 sinc 函数构成正交对, 即

$$\int_{-\infty}^{\infty} (\phi(t) \cos(2\pi f_n t))(\phi(t) \sin(2\pi f_n t)) dt = 0 \quad \text{对于所有 } n \quad (6.198)$$

根据上式所示的正交关系, N 个调制器的信号星座图都可以表达为平方网格的形式。

性质 2

由于

$$\exp(j2\pi f_n t) = \cos(2\pi f_n t) + j \sin(2\pi f_n t)$$

因此可将通带基本函数重新定义为复数形式

$$\left\{ \frac{1}{\sqrt{2}} \phi(t) \exp(j2\pi f_n t) \right\}, \quad n = 1, 2, \dots, N \quad (6.199)$$

其中引入了因子 $1/\sqrt{2}$, 以确保比例函数 $\phi(t)/\sqrt{2}$ 具有单位能量。因此, 通带基本函数构成了正交组, 即

$$\int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2}} \phi(t) \exp(j2\pi f_n t) \right) \left(\frac{1}{\sqrt{2}} \phi(t) \exp(j2\pi f_k t) \right)^* dt = \begin{cases} 1, & k = n \\ 0, & k \neq n \end{cases} \quad (6.200)$$

其中 * 代表复共轭。

式(6.200)确保了 N 个调制解调器对能够彼此独立地工作。

性质 3

对于具有任意冲激响应 $b(t)$ 的线性信道, 得到的一组信道输出函数 $|b(t) \star \phi(t)|$ 彼此正交, 其中 \star 代表卷积。

因此, 信道可分割为一组在连续时间内相互独立的子信道。

图 6.55 也包括了接收机框图。接收机由 N 个相干检测器组成, 信道输出同时送至所有检测器的输入端。每个检测器都对应于一对本地产生的正交调制 sinc 函数, 这些正交 sinc 函数对都与发射机中相应的通带基本函数进行了同步。

每个子信道中都可能存在残余的 ISI。但是, 随着子信道数 N 接近无穷, 理论上 ISI 最终将消失。因此, 当 N 足够大时, 图 6.55 所示的相干检测器组将以最大似然检测器的方式工作, 各检测器彼此独立, 一个子符号接着一个子符号地进行处理。

根据输入子符号定义检测器的输出时, 采用复数表示形式更为方便。用 A_n 代表在符号持续时间 $0 \leq t \leq T$ 内送入第 n 个调制器的子符号:

$$A_n = a_n + jb_n, \quad n = 1, 2, \dots, N \quad (6.201)$$

因此,检测器输出为

$$Y_n = H_n A_n + W_n, \quad n = 1, 2, \dots, N \quad (6.202)$$

其中 H_n 为在子信道载波频率 $f = f_n$ 时计算出的信道复值频率响应:

$$H_n = H(f_n), \quad n = 1, 2, \dots, N \quad (6.203)$$

W_n 为由信道噪声 $w(t)$ 引起的复值随机变量。 W_n 的实部和虚部的均值均为零,方差均为 $N_0/2$ 。当标准频率响应 $H(f)$ 已知时,可利用式(6.202)计算出发送子符号 A_n 的最大似然估计值。最后将得到的估计值 $\hat{A}_1, \hat{A}_2, \dots, \hat{A}_N$ 进行复用,从而得到在时间间隔 $0 \leq t \leq T$ 内发送的原始二进制数据的相应估计值。

当 N 足够大时,接收机可采用最优化的最大似然检测器,并以 N 个(子符号到子符号的)检测器的方式工作。之所以能够用如此简单的方式来构造最大似然接收机,其原因就在于,通带基本函数构成了正交组,且其正交性对于任意的信道冲激响应 $b(t)$ 都保持不变。

6.12.3 几何信噪比

在图 6.55 所示的多信道传输系统中,每个子信道都用其自身的信噪比代表。实际上,这种用信噪比代表子信道的度量方法并不实用。下面推导图 6.55 所示系统性能的一种简单度量方法。

为了简化推导,假设图 6.55 中的所有子信道都由一维星座图代表。则用比特/传输表示的整个系统的信道容量为

$$\begin{aligned} R &= \frac{1}{N} \sum_{n=1}^N R_n = \frac{1}{2N} \sum_{n=1}^N \log_2 \left(1 + \frac{P_n}{\Gamma \sigma_n^2} \right) = \frac{1}{2N} \log_2 \prod_{n=1}^N \left(1 + \frac{P_n}{\Gamma \sigma_n^2} \right) \\ &= \frac{1}{2} \log_2 \left[\prod_{n=1}^N \left(1 + \frac{P_n}{\Gamma \sigma_n^2} \right)^{1/N} \right] \end{aligned} \quad (6.204)$$

用 $(\text{SNR})_{\text{overall}}$ 代表整个系统的整体信噪比,因此可将 R 用比特/传输的形式表示为

$$R = \frac{1}{2} \log_2 \left(1 + \frac{(\text{SNR})_{\text{overall}}}{\Gamma} \right) \quad (6.205)$$

比较式(6.205)和式(6.204),可得

$$(\text{SNR})_{\text{overall}} = \Gamma \left(\prod_{n=1}^N \left(1 + \frac{P_n}{\Gamma \sigma_n^2} \right)^{1/N} - 1 \right) \quad (6.206)$$

假设 $P_n/\Gamma \sigma_n^2$ 很高,足以忽略式(6.206)中的两个单位项,则可将整体信噪比 SNR 近似表示为

$$(\text{SNR}) = \prod_{n=1}^N \left(\frac{P_n}{\sigma_n^2} \right)^{1/N} \quad (6.207)$$

因此,可用等于各单个子信道信噪比的几何平均信噪比来代表整个系统的信噪比。

通过将可用传输功率非均匀地分配到 N 个子信道中,可以显著提高式(6.207)所示的几何平均信噪比。下面将要讨论的加载就能够实现这一点。

6.12.4 多信道传输系统的加载

描述整个多信道系统比特率的式(6.204)忽略了信道对系统性能的影响。下面就探讨这种影响,定义

$$g_n = |H(f_n)|, \quad n = 1, 2, \dots, N \quad (6.208)$$

假设子信道数 N 足够大,因此,可假设对于所有的 n , g_n 在分配给子信道 n 的整个带宽 Δf 内都是恒定的。在这种情况下,可将式(6.204)的第二行改写为

$$R = \frac{1}{2N} \sum_{n=1}^N \log_2 \left(1 + \frac{g_n^2 P_n}{\Gamma \sigma_n^2} \right) \quad (6.209)$$

g_n^2 和 Γ 通常是固定的。对于所有的 n , 噪声方差 σ_n^2 都等于 $\Delta f N_0$, 其中 Δf 为每个子信道的带宽, $N_0/2$ 为噪声功率谱密度。因此,可通过在不同信道间合理分配全部传输功率来优化整体比特率 R 。为了使这种优化具有实用价值,必须将总的传输功率维持在恒定值 P , 即

$$\sum_{n=1}^N P_n = P = \text{常数} \quad (6.210)$$

现在面临的问题变为:

通过在 N 个子信道间优化地分享全部传输功率 P , 来最大化整个多信道传输系统的比特率 R , 约束条件是 P 保持恒定。

为了解决优化问题,我们首先使用拉格朗日因子法(method of Lagrange multipliers)¹¹⁹ 来构造包含式(6.210)中约束条件的目标函数,即

$$\begin{aligned} J &= \frac{1}{2N} \sum_{n=1}^N \log_2 \left(1 + \frac{g_n^2 P_n}{\Gamma \sigma_n^2} \right) + \lambda \left(P - \sum_{n=1}^N P_n \right) \\ &= \frac{1}{2N} \log_2 e \sum_{n=1}^N \log_e \left(1 + \frac{g_n^2 P_n}{\Gamma \sigma_n^2} \right) + \lambda \left(P - \sum_{n=1}^N P_n \right) \end{aligned} \quad (6.211)$$

其中 λ 为拉格朗日因子。将 J 对 P_n 求导,并令该导数值为 0 并进行移项,可得

$$\frac{\frac{1}{2N} \log_2 e}{P_n + \frac{\Gamma \sigma_n^2}{g_n^2}} = \lambda \quad (6.212)$$

这样,上述约束条件就变为

$$P_n + \frac{\Gamma \sigma_n^2}{g_n^2} = K \quad n = 1, 2, \dots, N \quad (6.213)$$

其中 K 为设计者能够控制的预设常数。也就是说,每个子信道的传输功率与噪声方差(功率)的 Γ/g_n^2 倍之和必须保持恒定。将传输功率 P 分配到各个子信道中,从而使整个多信道传输系统的比特率最大化的过程,就称为加载。

6.12.5 优化问题的注水解释

要解决刚才提到的带约束条件的优化问题，必须满足两个条件，即式(6.210)和式(6.213)。图6.56为 $N=6$ 时加载方案的一个有趣解释，其中假设间隙 Γ 在所有的子信道中都是恒定的。为了简化图6.56，令 $\sigma_n^2 = N_0 \Delta f = 1$ ，即 N 个子信道的平均噪声功率均为1。由图可知：

- 当指定常数 K 的值时，对于其中四个子信道，功率 P_n 与噪声功率的 Γ/g_n^2 倍之和满足式(6.213)的约束。
- 分配到这四个子信道的功率之和即为全部可用传输功率 P ， P 为恒定值。
- 剩下的两个子信道不考虑，因为对于指定的常数 K ，为满足式(6.213)，它们都需要负功率，这显然是不可行的。

图6.56解释了式(6.213)所示的优化方案与式(6.210)所示的约束条件之间的关系，称为注水方案。之所以使用这样的术语，是因为此处讨论的优化问题类似于将固定量的水（代表传输功率）倒入一个分成了若干具有不同深度（代表噪声功率）的相连区域的容器的问题。水的分配应当使得整个容器具有恒定的水平面高度。第9章将更详细地介绍信息容量的注水解释。

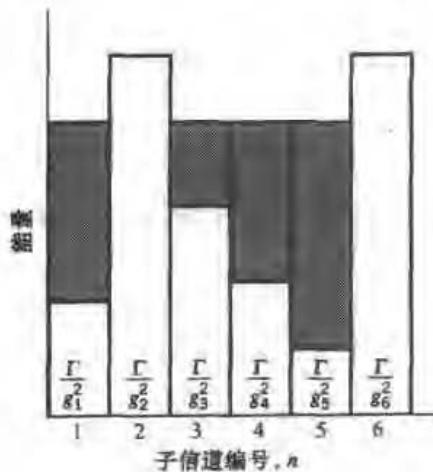


图6.56 加载问题的注水解释

下面继续研究如何将固定的传输功率 P 分配到多信道传输系统的各个子信道中，从而优化整个系统的比特率，可按如下过程进行分配。如式(6.210)所示，将总传输功率固定为常数 P 。用 K 代表式(6.213)所示的所有 n 的情况下 $P_n + \Gamma\sigma_n^2/g_n^2$ 对应的指定常数值。因此，可利用这对公式，构造出如下一组联立方程：

$$\begin{aligned}
 P_1 + P_2 + \cdots + P_N &= P \\
 P_1 - K &= -\Gamma\sigma^2/g_1^2 \\
 P_2 - K &= -\Gamma\sigma^2/g_2^2 \\
 &\vdots \\
 P_N - K &= -\Gamma\sigma^2/g_N^2
 \end{aligned} \tag{6.214}$$

其中，共有 $(N+1)$ 个未知数和 $(N+1)$ 个用于求解这些未知数的等式。可将这组联立方程写为如下

的矩阵形式:

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 0 \\ 1 & 0 & \cdots & 0 & -1 \\ 0 & 1 & \cdots & 0 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_N \\ K \end{bmatrix} = \begin{bmatrix} P \\ -\Gamma\sigma^2/g_1^2 \\ -\Gamma\sigma^2/g_2^2 \\ \vdots \\ -\Gamma\sigma^2/g_N^2 \end{bmatrix} \quad (6.215)$$

将式(6.215)的两边同时乘以其左边 $(N+1) \times (N+1)$ 维矩阵的逆矩阵,就可以解出未知数 P_1, P_2, \dots, P_N ,以及 K 。不难发现, K 总是正数,但一些 P 却可能是负数。由于功率不可能为负,所以这些 P 值将被舍弃。

例 6.7

考虑一个线性信道,其平方幅度响应 $|H(f)|^2$ 具有如图 6.57 所示的分段线性形式。为简化起见,令间隙 $\Gamma=1$,噪声方差 $\sigma^2=1$ 。应用式(6.214),可得

$$P_1 + P_2 = P$$

$$P_1 - K = -1$$

$$P_2 - K = -1/l$$

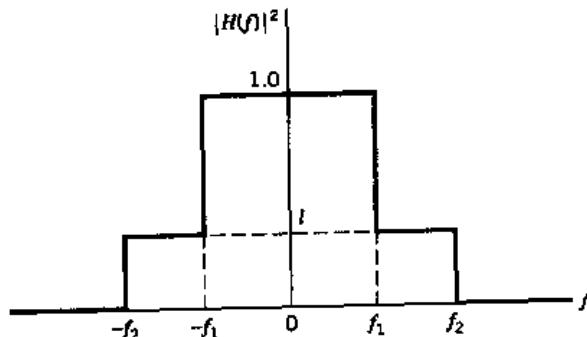


图 6.57 例 6.7 对应的平方幅度响应

其中全部传输功率 P 按照噪声方差进行了归一化。由这三个等式可解出 P_1, P_2 和 K ,即

$$P_1 = \frac{1}{2} \left(P - 1 + \frac{1}{l} \right)$$

$$P_2 = \frac{1}{2} \left(P + 1 - \frac{1}{l} \right)$$

$$K = \frac{1}{2} \left(P + 1 + \frac{1}{l} \right)$$

由于 $0 < l < 1$,因此 $P_1 > 0$,但 P_2 可以是负数。出现这种情况的条件是

$$l < \frac{1}{P+1}$$

但是,此时 P_1 超过了传输功率 P 的指定值。因此,在这个例子里惟一可行的是满足 $1/(P+1) < l < 1$ 的方案。假设有 $P=10$ 和 $l=0.1$,因此

$$K = 10.5$$

$$P_1 = 9.5$$

$$P_2 = 0.5$$

注水图如图 6.58 所示。

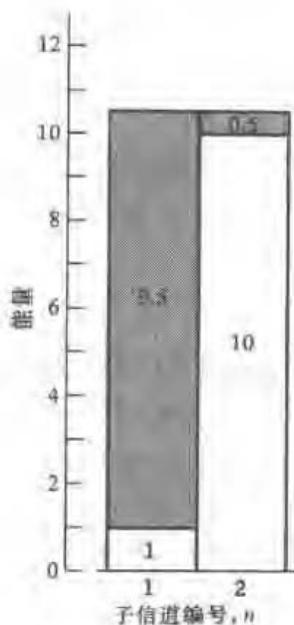


图 6.58 例 6.7 对应的注水剖面

6.13 离散多音

6.12 节对多信道调制的概念进行了深入的介绍。特别地,由式(6.196)或式(6.199)所示的通带基本函数引起的连续时间信道分割,表现出了很好的性质:即使与信道冲激响应进行卷积后,基本函数(以及对应的信道分割)的正交性仍然存在。但该系统仍然存在两个缺点:

1. 通带基本函数使用的 sinc 函数在无限长的时间间隔内是非零的。但在实际应用中,观察时间通常是有有限长度的。
2. 对于有限的子信道数 N ,系统并不是最优的。只有当 N 接近无穷大时,系统才能获得最优性能。

可通过采用离散多音(DMT)来克服上述缺点,其基本思想是将宽带信道转换为一组 N 个并行工作的子信道。DMT 与众不同的原因就在于这种转换是针对离散的时间和频率进行的。因此,整个通信系统中发射机的输入 - 输出特性可以用一个线性矩阵代表,因此可用离散傅里叶变换来实现。

下面就开始介绍 DMT。在实际情况下,信道有其自身的非零冲激响应 $b(t), b(t)$ 本质上限制在有限时间间隔 $[0, T_b]$ 内。用序列 b_0, b_1, \dots, b_n 代表以速率 $1/T_b$ 进行抽样的信道基带冲激响应,于是有

$$T_b = (1 + \nu)T_s \quad (6.216)$$

根据抽样定理,选择的抽样率 $1/T_s$ 应大于对应的高频分量频率的两倍。令 $s[n] = s(nT_s)$ 代表发送符号 $s(t)$ 的一个抽样值, $w[n] = w(nT_s)$ 代表信道噪声 $w(t)$ 的一个抽样值, $x[n] = x(nT_s)$ 代表信道输出(接收信号)相应的抽样值。信道对长度为 N 的输入符号序列 $|s[n]|$ 进行线性卷积,产生长度为 $N+v$ 的信道输出序列 $|x[n]|$ 。信道输出序列与信道输入序列相比,扩展了 v 个抽样点,这是由信道中的 ISI 引起的。

为克服 ISI 的影响,引入循环扩展保护时隙,从而使每个符号序列的前端都是其自身的周期性扩展。特别地,符号序列的最后 v 个抽样点在发送序列的前端进行重复,即

$$s[-k] = s[N-k] \quad k = 1, 2, \dots, v \quad (6.217)$$

这一条件称为循环前缀。由于引入了循环前缀而产生的额外带宽因子为 v/N ,其中 N 为保护时隙之后发送的符号数。

有了循环前缀,信道的矩阵描述形式变为

$$\begin{bmatrix} x[N-1] \\ x[N-2] \\ \vdots \\ x[N-v-1] \\ x[N-v-2] \\ \vdots \\ x[0] \end{bmatrix} = \begin{bmatrix} b_0 & b_1 & b_2 & \cdots & b_{v-1} & b_v & 0 & \cdots & 0 \\ 0 & b_0 & b_1 & \cdots & b_{v-2} & b_{v-1} & b_v & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & b_0 & b_1 & \cdots & b_v \\ b_v & 0 & 0 & \cdots & 0 & 0 & b_0 & \cdots & b_{v-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ b_1 & b_2 & b_3 & \cdots & b_v & 0 & 0 & \cdots & b_0 \end{bmatrix} \begin{bmatrix} s(N-1) \\ s(N-2) \\ \vdots \\ s[N-v-1] \\ s[N-v-2] \\ \vdots \\ s[0] \end{bmatrix} + \begin{bmatrix} w[N-1] \\ w[N-2] \\ \vdots \\ w[N-v-1] \\ w[N-v-2] \\ \vdots \\ w[0] \end{bmatrix} \quad (6.218)$$

同样,可将信道的离散时间形式写为复数矩阵的形式:

$$\mathbf{x} = \mathbf{H}\mathbf{s} + \mathbf{w} \quad (6.219)$$

其中发送符号矢量 \mathbf{s} ,信道噪声矢量 \mathbf{w} ,以及接收信号矢量 \mathbf{x} 均为 $N \times 1$ 维的矢量,其定义分别为

$$\mathbf{s} = [s[N-1], s[N-2], \dots, s[0]]^T \quad (6.220)$$

$$\mathbf{w} = [w[N-1], w[N-2], \dots, w[0]]^T \quad (6.221)$$

和

$$\mathbf{x} = [x[N-1], x[N-2], \dots, x[0]]^T \quad (6.222)$$

这样就得到了图 6.59 所示的信道离散时间形式。 $N \times N$ 维的信道矩阵 \mathbf{H} 的定义为

$$\mathbf{H} = \begin{bmatrix} b_0 & b_1 & b_2 & \cdots & b_{v-1} & b_v & 0 & \cdots & 0 \\ 0 & b_0 & b_1 & \cdots & b_{v-2} & b_{v-1} & b_v & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & b_0 & b_1 & \cdots & b_v \\ b_v & 0 & 0 & \cdots & 0 & 0 & b_0 & \cdots & b_{v-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ b_1 & b_2 & b_3 & \cdots & b_v & 0 & 0 & \cdots & b_0 \end{bmatrix} \quad (6.223)$$

由定义可知,矩阵 \mathbf{H} 具有如下结构成分:矩阵的每一行都是通过将前一行右移一位而得到的,且前一行的最右端元素在移位过程中溢出,并循环回到新的一行的最左端。因此,矩阵 \mathbf{H} 被称为循环矩阵。

在继续深入研究之前,有必要简单回顾一下离散傅里叶变换及其在循环矩阵 \mathbf{H} 的频谱分解中的作用。

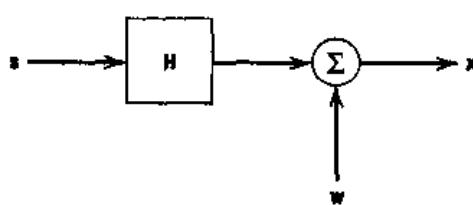


图 6.59 多信道数据传输系统的离散时间形式

6.13.1 离散傅里叶变换

考虑式(6.222)所示的 $N \times 1$ 维矢量 \mathbf{x} 。矢量 \mathbf{x} 的离散傅里叶变换(DFT)定义为一个 $N \times 1$ 维矢量

$$\mathbf{X} = [X[N-1], X[N-2], \dots, X[0]]^T \quad (6.224)$$

其中

$$X[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \exp\left(-j \frac{2\pi}{N} kn\right), \quad k = 0, 1, \dots, N-1 \quad (6.225)$$

指数项 $\exp(-j2\pi kn/N)$ 称为 DFT 核。因此, $N \times 1$ 维矢量 \mathbf{X} 的离散傅里叶反变换(IDFT, Inverse DFT)的定义为

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X[k] \exp\left(j \frac{2\pi}{N} kn\right), \quad n = 0, 1, \dots, N-1 \quad (6.226)$$

虽然式(6.225)和式(6.226)看上去很相似,但却有着不同的解释。给定信号矢量 \mathbf{x} , 式(6.225)提供了在一组离散频率值 $f_k = k/N$ 下计算出的该信号的频域形式, 该频域形式已经按抽样速率进行了归一化。给定变换后的矢量 \mathbf{X} , 式(6.226)则恢复出原始信号矢量 \mathbf{x} 。因此可将式(6.225)看成分解公式, 将式(6.226)看成合成公式。

以式(6.223)所示的信道矩阵 \mathbf{H} 为例, 循环矩阵的一个重要性质就是能够进行如下形式的频谱分解

$$\mathbf{H} = \mathbf{Q}^H \mathbf{A} \mathbf{Q} \quad (6.227)$$

其中上标 H 代表埃尔米特(Hermitian)共轭变换(复共轭和普通矩阵变换的结合)。接下来依次对矩阵 \mathbf{Q} 和 \mathbf{A} 进行介绍。

矩阵 \mathbf{Q} 是如下的一个由 N 点 DFT 核定义的方阵:

$$\mathbf{Q} = \frac{1}{\sqrt{N}} \begin{bmatrix} \exp\left(-j \frac{2\pi}{N} (N-1)(N-1)\right) & \cdots & \exp\left(-j \frac{2\pi}{N} 2(N-1)\right) & \exp\left(-j \frac{2\pi}{N} (N-1)\right) & 1 \\ \exp\left(-j \frac{2\pi}{N} (N-1)(N-2)\right) & \cdots & \exp\left(-j \frac{2\pi}{N} 2(N-2)\right) & \exp\left(-j \frac{2\pi}{N} (N-2)\right) & 1 \\ \vdots & & \vdots & \vdots & \vdots \\ \exp\left(-j \frac{2\pi}{N} (N-1)\right) & \cdots & \exp\left(-j \frac{2\pi}{N} 2\right) & \exp\left(-j \frac{2\pi}{N}\right) & 1 \\ 1 & \cdots & 1 & 1 & 1 \end{bmatrix} \quad (6.228)$$

由定义可见,从右下方 $k=0, l=0$ 处起,一步一步向上计算,可得 $N \times N$ 维矩阵 \mathbf{Q} 的第 k 个元素是

$$q_{kl} = \frac{1}{\sqrt{N}} \exp\left(-j \frac{2\pi}{N} kl\right), \quad (k, l) = 0, 1, \dots, N-1 \quad (6.229)$$

矩阵 \mathbf{Q} 是满足如下条件的标准正交矩阵或酉矩阵:

$$\mathbf{Q}^\dagger \mathbf{Q} = \mathbf{I} \quad (6.230)$$

其中 \mathbf{I} 为单位矩阵。也就是说, \mathbf{Q} 的逆矩阵等于 \mathbf{Q} 的埃尔米特变换。

矩阵 \mathbf{A} 是一个对角阵,其中包含了序列 b_0, b_1, \dots, b_N 的 N 个离散傅里叶变换值,代表了信道特性。用 $\lambda_{N-1}, \dots, \lambda_1, \lambda_0$ 来代表这些变换值,可将 \mathbf{A} 表示为

$$\mathbf{A} = \begin{bmatrix} \lambda_{N-1} & 0 & \cdots & 0 \\ 0 & \lambda_{N-2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_0 \end{bmatrix} \quad (6.231)$$

(注意:不要把这里的 λ 和 6.12 节的拉格朗日因子混淆。)

由于采用快速傅里叶变换(FFT)算法可进行高效的运算,DFT 已经成为数字信号处理的一个主要工具。^[20]特别地,FFT 算法只需进行 $N \log_2 N$ 次运算,而直接采用 DFT 则需要 N^2 次运算。为有效地实现 FFT 算法,选择的块长 N 必须恰好等于整数的平方。利用式(6.225)中定义的 DFT 的特殊结构,就可以使用 FFT 算法使计算行得到简化。更进一步,随着 N 的增大,这种简化变得越来越重要。

6.13.2 信道的频域描述

在对 DFT 进行了简要介绍后,就可以继续讨论 DMT 了。首先定义

$$\mathbf{s} = \mathbf{Q}^\dagger \mathbf{S} \quad (6.232)$$

其中 \mathbf{S} 为发射机输入的频域矢量形式。 $N \times 1$ 维矢量 \mathbf{S} 中的每个元素都可以看成是二维 QAM 信号星座图中的一个复值点。给定信道输出矢量 \mathbf{x} ,可将其相应的频域形式定义为

$$\mathbf{x} = \mathbf{Q} \mathbf{s} \quad (6.233)$$

利用式(6.227),式(6.232)和式(6.233),可将式(6.219)写为如下等价形式:

$$\mathbf{x} = \mathbf{Q}(\mathbf{Q}^\dagger \mathbf{A} \mathbf{Q} \mathbf{Q}^\dagger \mathbf{S} + \mathbf{W}) \quad (6.234)$$

利用式(6.230)所示的关系,容易得到

$$\mathbf{x} = \mathbf{A} \mathbf{s} + \mathbf{W} \quad (6.235)$$

其中

$$\mathbf{W} = \mathbf{Q} \mathbf{w} \quad (6.236)$$

式(6.235)的展开形式记为

$$x_k = \lambda_k s_k + w_k, \quad k = 0, 1, \dots, N-1 \quad (6.237)$$

其中,频域值 $\{\lambda_k\}_{k=0}^{N-1}$ 对于给定信道是已知的。

对于具有加性白噪声的信道,式(6.237)表明,接收机是由一组并行操作的独立处理器组成的。当 λ_k 均为已知时,就可以用频域值为 $\{X_k\}_{k=0}^{N-1}$ 的块来计算频域值为 $\{S_k\}_{k=0}^{N-1}$ 的相应发送块的估计值。

6.13.3 基于 DFT 的 DMT 系统

式(6.235),式(6.225),式(6.226)和式(6.237)提供了用 DFT 实现 DMT 的数学基础。图 6.60 揭示了由这些公式及其实际含义推导出的系统框图。

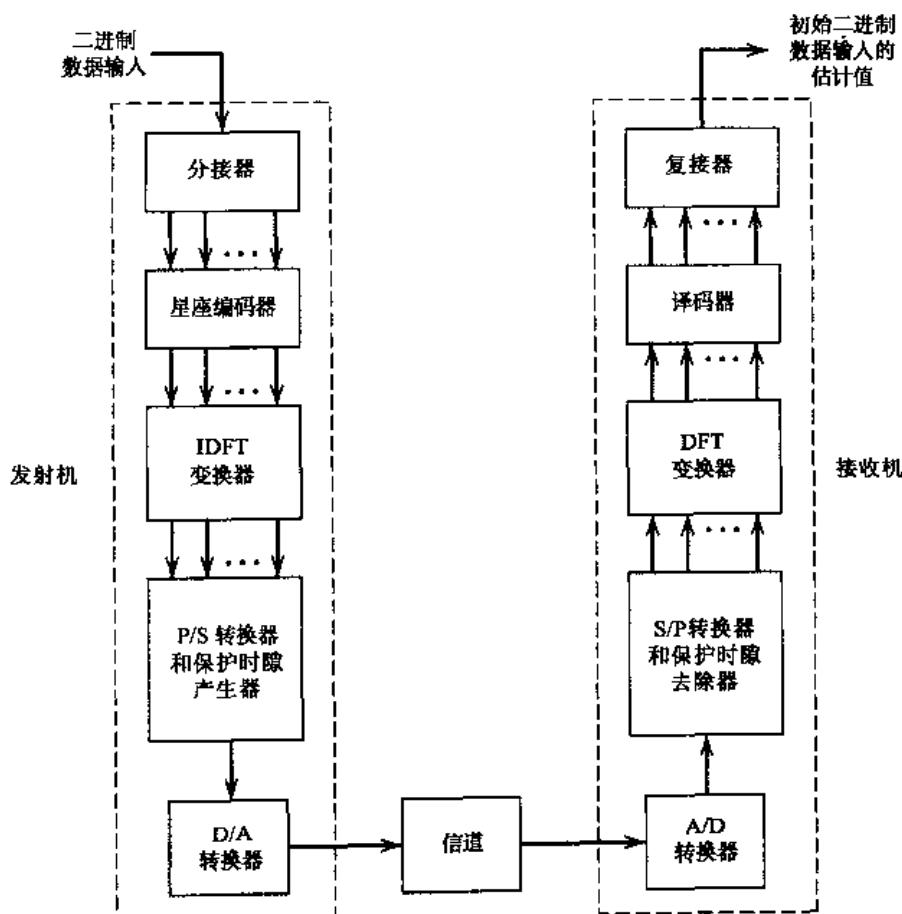


图 6.60 DMT 数据传输系统框图

发射机由下列功能块组成:

- 分接器,将输入的串行数据流转换为并行的形式。
- 星座编码器,将并行数据映射到 $N/2$ 个多比特子信道,每个子信道都用一个 QAM 信号星座代表。此处子信道间的比特分配根据加载算法进行。
- IDFT 变换器,将星座编码器输出端的频域并行数据转换为并行的时域数据。为了用 FFT 算法有效地实现 IDFT,需选择 $N = 2^k$,其中 k 为正整数。
- P/S(并 - 串)转换器,将并行的时域数据转换为串行的形式。在转换为模拟形式之前,用来填充循环前缀的保护时隙被周期性地插入串行数据。

- D/A(数-模)转换器,将数字数据转换为模拟形式,从而为信道传输做好准备。

特别地,D/A转换器包含一个发射滤波器。因此,时间函数 $b(t)$ 应重新定义为发射滤波器与信道级联的联合冲激响应。

接收机完成与发射机相反的操作,即:

- A/D(模-数)转换器,将模拟信道输出转换为数字形式。
- S/P(串-并)转换器,将收到的比特流转换为并行的形式。在转换之前,要先去除保护时隙(循环前缀)。
- DFT变换器,将时域并行数据转换为频域并行数据。与IDFT类似,DFT也用FFT算法来实现。
- 译码器,利用DFT的输出,计算出原始多比特子信道数据的估计值。
- 复接器,结合上述估计值,重构发送的串行数据流。

6.13.4 DMT的应用

DMT的一个重要应用就是在双向信道的数据传输中。实际上,DMT在使用双绞线的ADSL中的应用已经进行了标准化。ADSL在第4章中已经介绍过。举个例子,DMT提供的下行速率(例如从ISP到用户)为1.544 Mbps(即DS1速率),上行速率(例如从用户到ISP)为160 Kbps。这样的数据传输能力非常适合于视频点播之类的突发数据应用。

DMT也是实现非对称超高速数字用户线^[21](VDSL)的核心技术。VDSL与所有其他DSL传输技术的区别就在于它能够提供非常高的数据速率。例如,在从光纤网络单元到大约1km以内的用户的双绞线上,VDSL能够提供13~26 Mbps的下行速率,以及2~3 Mbps的上行速率。如此高的数据速率,使VDSL适合于数字电视、超高速网上冲浪、文件传输以及家中虚拟办公等高速率应用。

将DMT用于ADSL和VDSL,可带来一系列好处:

- 能够最大化传输比特率,根据信道衰减和噪声条件进行信道间的信号分配,就能够实现这一点。
- 对信道条件变化的自适应性,这是通过将信道分割为一组子信道来实现的。
- 对突发噪声的低敏感性,这是通过将噪声能量分布到接收机的多个子信道来实现的。顾名思义,突发噪声的特点是较长的平静的时间间隔后出现的幅度随机变化的窄脉冲。在ADSL或VDSL环境中,突发噪声的形成原因包括中心局(CO)开关瞬时耦合于金属线对和用户端各种电子设备的影响。

6.13.5 DSL与音频调制解调器之间的比较

6.11节中所讨论的音频调制解调器已经能工作于接近理论极限的33.6 Kbps的上行速率和56 Kbps的下行速率。本节讨论DMT在VDSL中的应用。DMT工作在2~3 Mbps的上行速率和13~26 Mbps的下行速率下。这两个截然不同的上行/下行数据速率范围带来了以下问题:VDSL如何在同样的双绞线(如电话线)上提供高于音频调制解调器近3倍的数据速率?引起音频调制解调器和VDSL工作数据速率巨大差别的不是双绞线,而是PSTN中的数字交换。数字交换阻碍了宽带数据通过音频调制解调器传送到用户的过程。简单地说,数字交换将数

字信号都看成是待处理的语音信号。

图 6.61 特别给出了音频调制解调器和 xDSL 的工作环境, 其中 x 代表 ADSL 中的 A 和 VDSL 中的 V。在图 6.61 所示的对于音频调制解调器的模型中, ISP 与用户之间的传输路径相对较长。更重要的是, 该传输路径穿过一个窄带 PSTN, 而 PSTN 将可用信道带宽限制在约 3.5 kHz。相反, 在图 6.61 所示的对于 xDSL 的模型中, 传输路径可以适应于在 ISP 和用户之间通过 ISDN(宽带综合服务数字网)和一个相对较短的双绞线构成的本地环路传输宽带数据。系统允许 POTS 和 xDSL 信号共存于同一个本地环路中, 这通过采用图 6.61(b)所示的一对分接器来实现。分接器由双向低通和高通滤波器组成, 在 4.8 节已有介绍。

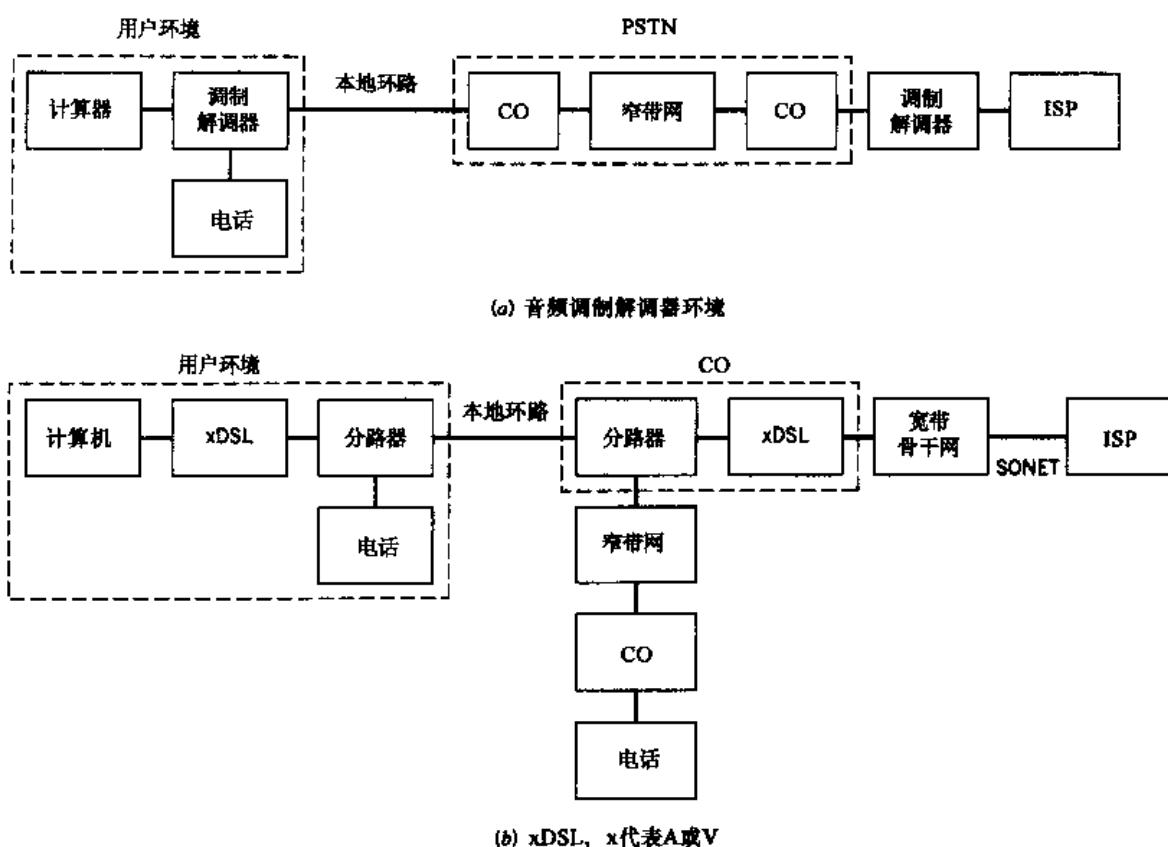


图 6.61 音频调制解调器与 xDSL

6.13.6 OFDM 技术^[22]

DMT 是多信道调制的一种特殊的离散形式。该方法的另一个密切相关的形式就是 OFDM(正交频分复用)。OFDM 与 DMT 在应用领域和设计方面有所差别。

OFDM 适用于通过无线电广播信道和无线通信信道进行数据传输。这就要求对 OFDM 系统设计做一些变化。和使用加载技术进行比特分配的 DMT 不同, OFDM 的每个子信道使用固定数目的比特数。由于广播信道为单向传输, 而无线通信环境中的信道却是双向且变换很快的, 因此对系统设计进行相应改变是有必要的。两种情况下, 发射机都不可能知道信道情况以及如何“加载”它。

图 6.60 所示框图也适用于 OFDM, 但在信号星座编码器中不包括用子比特分配的加载算

法。此外,在进行系统设计时还有两个不同:

- 在发射机里,D/A转换器后接有一个上行转换器,用来转换发射频率,这是为了便于无线电信道中发射信号的传播。
- 在接收机里,A/D转换器后接有一个下行转换器,用来对发射机里上行转换器的频率转换进行反操作。

OFDM的应用包括:

1. 无线通信

OFDM、编码以及交织技术的结合,提供了解决无线通信信道特有的多径衰落问题的有效技术。

2. 数字音频广播

在欧洲,OFDM已经被采纳为数字音频广播的标准。这种应用中也结合了编码和交织技术。

差错控制编码以及相关的问题将在第10章进行讨论。

6.14 同步

不管对任何形式的数字调制信号进行相干接收,都要求接收机与发射机保持同步。当一个序列中的事件和另一个序列中的相应事件同时发生时,我们就称这两个事件序列(代表发射机和接收机)是彼此同步的。产生并维持同步的过程就叫同步化(以下简称同步)。^[23]

在数字调制技术中,需要两种基本的同步模式:

1. 当使用相干检测时,需要知道载波的频率和相位。对载波频率和相位的估计称为载波恢复或载波同步。
2. 为了进行解调,接收机需要知道调制器改变状态的瞬时时刻。也就是说,接收机必须知道单个符号的起始和结束时刻,从而确定何时抽样以及何时结束乘积积分。对这些时间的估计称为时钟恢复或者符号同步。

这两种同步模式可以同时进行,也可以依次进行。在非相干系统中,不考虑载波同步的问题。

同步可以通过两种不同的方法来实现:

1. 数据辅助同步

在数据辅助同步系统中,周期性地以时分复用技术发送探测信号和携带数据的信号。探测信号包含了载波和符号的定时信息,只要在接收机中对信道输出进行适当处理,就可以提取出这些信息。这种方法通常用于数字卫星和无线通信,其目的是使接收机与发射机建立同步所需要的时间最小化。这种方法有两个局限性:(1)由于每个发送帧都有一部分被分配给探测信号,因此,数据吞吐效率有所降低;(2)由于发射功率的一部分被用于探测信号的发送,因此,功率效率有所降低。

2. 无数据辅助同步

在这种方法里,不需要使用探测信号,接收机的任务是通过从调制信号里提取出必要信息来建立起同步。建立同步所需时间比第一种方法长,但可换取吞吐效率和功率效

率的提高。

同步本质上是一个统计参数估计的问题。解决这一问题的一个原则性方法就是最大似然估计(见 5.5 节)。这种方法从构造所关心的参数在给定接收信号时的对数似然函数来开始进行。将信道噪声看成高斯过程,可使上述函数构造过程相对简单。更为重要的是,这样就不需要关于调制信号的先验信息。

在这一节中,我们将注意力集中在无数据辅助的载波和定时同步系统上。假设为了节省功率,采用抑制载波的调制信号。可定义两种解决该同步问题的方法:

1. 经典方式

经典同步方式的一个关键模块就是锁相环(锁相环在第 2 章中进行了介绍)。特别地,为了进行载波恢复,接收机需要使用抑制载波跟踪环来提供相干二次载波(副载波)参考。例如,可将科斯塔环或 M 次方环的变换形式用于 M 进制 PSK。用于 DSB-SC 调制的标准科斯塔环在第 2 章中进行了介绍。 M 次方环由 M 次方运算器、带通滤波器、锁相环以及 M 分频器级联而成。这里要用到的是锁相环的捕获和跟踪特性。为了进一步了解 M 次方环,可参考习题 6.47。

2. 现代(算法)方式

在现代方式中,通过采用离散时间信号处理,将最大似然函数的解表达为算法形式。同步器的实现中用到的算法是,基于迭代对载波相位或符号定时进行估计。该过程是在基带域进行的,从而为离散时间(数字)信号处理提供了基础。

本节将介绍对 M 进制 PSK 系统进行载波恢复和符号定时恢复的同步算法。

由于定时恢复在相位恢复之前进行,因此我们先介绍定时恢复。这样做的原因是:如果知道了信号通过信道造成的群时延,则在接收机匹配滤波器输出端的每个符号,只要取一个抽样点就足以估计未知的载波相位了。更进一步,通过使用以符号速率 $1/T$ 工作的同步算法,可最小化接收机的计算复杂度。

6.14.1 相位恢复的直接判决递归算法

前面已经提到,解决同步问题的首要步骤是,在已知包含有高斯噪声的接收信号时,构造出载波相位 θ 的对数似然函数。用 $I(\theta)$ 代表该对数似然函数,也就是对 θ 进行估计的目标函数。下一步就是求 $I(\theta)$ 对 θ 的导数。最后一步是构造出递归算法并一步一步地计算出未知相位 θ 的最大似然估计值。

$\partial I(\theta)/\partial \theta$ ^① 的计算

用 $s_k(t)$ 代表符号 $k = 0, 1, \dots, M - 1$ 对应的发射信号:

$$s_k(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t + \alpha_k), \quad 0 \leq t \leq T \quad (6.238)$$

其中, E 为符号能量, T 为符号持续时间,且

① 对 $\partial I(\theta)/\partial \theta$ 的推导不感兴趣的读者可以跳过这一部分,直接阅读后面的内容。

$$\alpha_k = 0, \frac{2\pi}{M}, \dots, (M-1) \frac{2\pi}{M} \quad (6.239)$$

等效地,有

$$s_k(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t + \alpha_k) g(t) \quad (6.240)$$

其中, $g(t)$ 为波形脉冲, 即具有单位幅度, 且持续时间为 T 的方波脉冲。用 τ_c 代表载波(相位)时延, τ_g 代表包络(群)时延, 它们都是由信道引起的。根据定义, τ_c 影响载波而 τ_g 影响包络。则接收信号为

$$\begin{aligned} x(t) &= \sqrt{\frac{2E}{T}} \cos(2\pi f_c(t - \tau_c) + \alpha_k) g(t - \tau_g) + w(t) \\ &= \sqrt{\frac{2E}{T}} \cos(2\pi f_c t + \theta + \alpha_k) g(t - \tau_g) + w(t) \end{aligned} \quad (6.241)$$

其中, $w(t)$ 为信道噪声, θ 等于 $-2\pi f_c \tau_c$, 以便与 6.6 节中一致。载波相位 θ 和群时延 τ_g 都是未知的, 假设它们在观察时间间隔 $0 \leq t \leq T_0$ 内(或者说在 $L_0 = T_0/T$ 个符号的传输过程中)保持恒定。等价地, 有(用 τ 代替 τ_g):

$$x(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_c t + \theta + \alpha_k) + w(t), \quad \tau \leq t \leq T + \tau \quad (6.242)$$

接收机端的基本函数为

$$\phi_1(t) = \sqrt{\frac{2}{T}} \cos(2\pi f_c t), \quad \tau \leq t \leq T + \tau \quad (6.243)$$

$$\phi_2(t) = \sqrt{\frac{2}{T}} \sin(2\pi f_c t), \quad \tau \leq t \leq T + \tau \quad (6.244)$$

其中, 假设接收机对载波频率 f_c 有充分了解, 否则就需要考虑载波频率的偏移, 这将使分析变得更为复杂。因此, 可用下面的矢量来表示接收信号 $x(t)$:

$$\mathbf{x}(\tau) = \begin{bmatrix} x_1(\tau) \\ x_2(\tau) \end{bmatrix} \quad (6.245)$$

其中

$$x_i(\tau) = \int_{\tau}^{T+\tau} x(t) \phi_i(t) dt, \quad i = 1, 2 \quad (6.246)$$

相应地, 可用下面的矢量来表示 $\mathbf{x}(\tau)$ 的信号组成:

$$\mathbf{s}(a_k, \theta, \tau) = \begin{bmatrix} s_1(a_k, \theta, \tau) \\ s_2(a_k, \theta, \tau) \end{bmatrix} \quad (6.247)$$

其中, a_k 为发送符号, 且

$$s_i(a_k, \theta, \tau) = \int_{\tau}^{T+\tau} \sqrt{\frac{2E}{T}} \cos(2\pi f_c t + \theta + \alpha_k) \phi_i(t) dt \quad i = 1, 2 \quad (6.248)$$

假设 f_c 是符号速率 $1/T$ 的整数倍, 于是有

$$s_1(a_k, \theta, \tau) = \sqrt{E} \cos(\theta + \alpha_k) \quad (6.249)$$

$$s_2(a_k, \theta, \tau) = -\sqrt{E} \sin(\theta + \alpha_k) \quad (6.250)$$

因此有

$$\mathbf{x}_k(\tau) = \mathbf{s}(a_k, \theta, \tau) + \mathbf{w} \quad (6.251)$$

其中, \mathbf{w} 为噪声矢量

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad (6.252)$$

且

$$w_i = \int_{\tau}^{\tau+\tau} w(t) \phi_i(t) dt, \quad i = 1, 2 \quad (6.253)$$

其中, w_i 为具有零均值且方差为 $N_0/2$ 的高斯随机变量 W 的抽样值, 其中, $N_0/2$ 为信道噪声 $w(t)$ 的(双边)功率谱密度。

给定传输符号 a_k , 载波相位 θ 和群时延 τ 时, 随机矢量 \mathbf{X} 的条件概率密度函数为

$$f_{\mathbf{X}}(\mathbf{x} | a_k, \theta, \tau) = \frac{1}{\pi N_0} \exp\left(-\frac{1}{N_0} \|\mathbf{x}_k(\tau) - \mathbf{s}(a_k, \theta, \tau)\|^2\right) \quad (6.254)$$

当 $a_k = 0$ 时, 接收信号 $x(t)$ 等于信道噪声 $w(t)$, 因此

$$f_{\mathbf{X}}(\mathbf{x} | a_k = 0) = \frac{1}{\pi N_0} \exp\left(-\frac{1}{N_0} \|\mathbf{x}_k(\tau)\|^2\right) \quad (6.255)$$

于是, 可将接收机端 M 进制 PSK 的似然函数定义为

$$\begin{aligned} L(a_k, \theta, \tau) &= \frac{f_{\mathbf{X}}(\mathbf{x} | a_k, \theta, \tau)}{f_{\mathbf{X}}(\mathbf{x} | a_k = 0)} \\ &= \exp\left(\frac{2}{N_0} \mathbf{x}_k^T(\tau) \mathbf{s}(a_k, \theta, \tau) - \frac{1}{N_0} \|\mathbf{s}(a_k, \theta, \tau)\|^2\right) \end{aligned} \quad (6.256)$$

在 M 进制 PSK 中, 有

$$\|\mathbf{s}(a_k, \theta, \tau)\| = \text{常数}$$

信息点位于半径为 \sqrt{E} 的圆上。忽略指数里的第二项, 可将似然函数简化为

$$L(a_k, \theta, \tau) = \exp\left(\frac{2}{N_0} \mathbf{x}_k^T(\tau) \mathbf{s}(a_k, \theta, \tau)\right) \quad (6.257)$$

假设发送 L_0 个统计独立的符号, 即

$$\mathbf{a} = [a_0, a_1, \dots, a_{L_0-1}]^T \quad (6.258)$$

则得到的似然函数为

$$L(\mathbf{a}, \theta, \tau) = \prod_{k=0}^{L_0-1} \exp\left(\frac{2}{N_0} \mathbf{x}_k^T(\tau) \mathbf{s}(a_k, \theta, \tau)\right) \quad (6.259)$$

因此,对数似然函数为

$$\begin{aligned} l(a, \theta, \tau) &= \log L(a, \theta, \tau) \\ &= \frac{2}{N_0} \sum_{k=0}^{L_0-1} \mathbf{x}_k^T(\tau) s(a_k, \theta, \tau) \end{aligned} \quad (6.260)$$

由式(6.249)和式(6.250)可推导出

$$\begin{aligned} \hat{s}_k(\theta) &= s(\hat{a}_k, \theta, \tau) \\ &= \sqrt{E} \begin{bmatrix} \cos(\hat{\alpha}_k + \theta) \\ -\sin(\hat{\alpha}_k + \theta) \end{bmatrix}, \quad k = 0, 1, \dots, L_0 - 1 \end{aligned} \quad (6.261)$$

其中 $\hat{\alpha}_k$ 为检测器输出端对符号 a_k 的实际值 α_k 的一个估计。因此,可将匹配滤波器的输出写为

$$\mathbf{x}_k = \begin{bmatrix} x_{1,k} \\ -x_{2,k} \end{bmatrix}$$

将此定义和式(6.261)代入式(6.260),可得

$$\begin{aligned} l(\theta) &= \frac{2\sqrt{E}}{N_0} \sum_{k=0}^{L_0-1} [x_{1,k} \cos(\hat{\alpha}_k + \theta) + x_{2,k} \sin(\hat{\alpha}_k + \theta)] \\ &= \frac{2\sqrt{E}}{N_0} \sum_{k=0}^{L_0-1} [(x_{1,k} \cos \hat{\alpha}_k + x_{2,k} \sin \hat{\alpha}_k) \cos \theta \\ &\quad - (x_{1,k} \sin \hat{\alpha}_k - x_{2,k} \cos \hat{\alpha}_k) \sin \theta] \end{aligned} \quad (6.262)$$

将 $l(\theta)$ 对 θ 取微分,可得

$$\begin{aligned} \frac{\partial l(\theta)}{\partial \theta} &= -\frac{2\sqrt{E}}{N_0} \sum_{k=0}^{L_0-1} [(x_{1,k} \cos \hat{\alpha}_k + x_{2,k} \sin \hat{\alpha}_k) \sin \theta \\ &\quad + (x_{1,k} \sin \hat{\alpha}_k - x_{2,k} \cos \hat{\alpha}_k) \cos \theta] \end{aligned} \quad (6.263)$$

通过引入如下符号,可简化式(6.263):

$$\tilde{x}_k = x_{1,k} + jx_{2,k} \quad (6.264)$$

及

$$\begin{aligned} a_k &= e^{j\alpha_k} \\ &= \cos \alpha_k + j \sin \alpha_k \end{aligned} \quad (6.265)$$

其中 x_k 为匹配滤波器(对应于第 k 个发送符号的)输出的复数包络(基带值),而 a_k 为 M 进制 PSK 信息点星座图的一个符号指示。因此有

$$\begin{aligned} \operatorname{Re}[\hat{a}_k^* \tilde{x}_k] &= \operatorname{Re}[(\cos \hat{\alpha}_k - j \sin \hat{\alpha}_k)(x_{1,k} + jx_{2,k})] \\ &= x_{1,k} \cos \hat{\alpha}_k + x_{2,k} \sin \hat{\alpha}_k \end{aligned} \quad (6.266)$$

$$\begin{aligned} \operatorname{Im}[\hat{a}_k^* \tilde{x}_k] &= \operatorname{Im}[(\cos \hat{\alpha}_k - j \sin \hat{\alpha}_k)(x_{1,k} + jx_{2,k})] \\ &= -x_{1,k} \sin \hat{\alpha}_k + x_{2,k} \cos \hat{\alpha}_k \end{aligned} \quad (6.267)$$

由欧拉公式

$$e^{-j\theta} = \cos \theta - j \sin \theta \quad (6.268)$$

可将式(6.263)写为复数形式

$$\begin{aligned}\frac{\partial l(\theta)}{\partial \theta} &= \frac{2\sqrt{E}}{N_0} \sum_{k=0}^{L_0-1} \{ (\text{Re}[\hat{a}_k^* \tilde{x}_k]) (\text{Im}[e^{-j\theta}]) + (\text{Im}[\hat{a}_k^* \tilde{x}_k]) (\text{Re}[e^{-j\theta}]) \} \\ &= \frac{2\sqrt{E}}{N_0} \sum_{k=0}^{L_0-1} \text{Im}[\hat{a}_k^* \tilde{x}_k e^{-j\theta}]\end{aligned}\quad (6.269)$$

其中 \hat{a}_k 为 a_k 的一个估计值, 而 * 代表复共轭。

6.14.2 对载波相位进行最大似然估计的递归算法

利用式(6.269)的对数似然函数 $l(\theta)$ 对载波相位 θ 的导数, 可推导出最大化 $l(\theta)$ 的算法。如果最大化算法能够迭代进行, 接收机就可以一个符号接着一个符号地对接收信号进行处理。为此, 可利用由自适应滤波器借鉴而来的算法思想(见第3和第4章对LMS算法的讨论):

$$(\text{修正估计}) = (\text{旧估计}) + (\text{步长参数})(\text{误差信号}) \quad (6.270)$$

其中, 误差信号(更准确地说是调节信号)是对数似然函数 $l(\theta)$ 关于 θ 的梯度的瞬时值。此处的目标是进行梯度上升, 因此, 式(6.270)中旧的估计值对应的参数调节值应为正数。由式(6.269)可见, 误差信号(由于发送单个符号引起的 $\partial l(\theta)/\partial \theta$ 的瞬时值)为

$$e[n] = \text{Im}[\hat{a}_n^* \tilde{x}_n e^{-j\theta}] \quad (6.271)$$

其中比例因子 $2\sqrt{E}/N_0$ 将在下文中进行说明。也可用 n 代替 k 来代表算法的时间步长或者迭代次数。由式(6.270)可得

$$\hat{\theta}[n+1] = \hat{\theta}[n] + \gamma e[n] \quad (6.272)$$

其中 $\hat{\theta}[n]$ 为载波相位 θ 的旧估计值, $\hat{\theta}[n+1]$ 为 θ 的新估计值, 而 γ 为步长参数(比例因子 $2\sqrt{E}/N_0$ 包含在 γ 内)。

式(6.271)和式(6.272)定义了相位恢复的递归算法。这一算法可以利用图6.62所示的系统来实现, 该系统可以看成是科斯塔环的递归生成器, 因此可称为相位同步的递归科斯塔环。

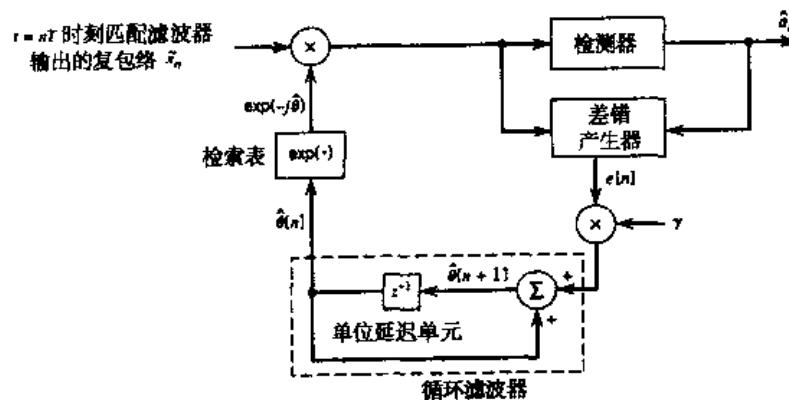


图 6.62 递归科斯塔环

在图6.62中应注意以下几点:

- 给定匹配滤波器的输出, 检测器就能给出对发送符号 a_n 的估计值。
- 对输入的 $(-j\hat{\theta}[n]) = \cos\hat{\theta}[n] - j\sin\hat{\theta}[n]$, 检索表提供对应的 $\hat{\theta}[n]$ 值。
- 差错生成器的输出为误差信号 $e[n]$ 。
- 标有 z^{-1} 字样的模块是时延等于符号持续时间 T 的单位延迟单元。

图 6.62 所示的递归科斯塔环中采用了一阶数字滤波器。为了提高该同步系统的跟踪性能, 可使用二阶数字滤波器。图 6.63 为由(包括两个一阶部分的)科斯塔环构成的二阶数字滤波器, 其中, ρ 为可调节的环参数。用于相位恢复的科斯塔环中的二阶滤波器有一个重要特性, 即当接收机和发射机之间的初始频差很小时, 该滤波器最终将无统计误差地锁定在输入载频上。

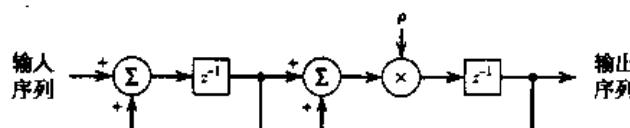


图 6.63 二阶数字滤波器

6.14.3 符号定时的无数据辅助递归算法

为进行定时同步, 所做的惟一假设是接收机已知载波频率 f_c 。目标是推导出对群时延 τ 进行递归估计的算法, 群时延 τ 是由于调制信号通过信道传输而引入的。

用 $L(a_k, \theta, \tau)$ 代表 τ 的似然函数。 $L(a_k, \theta, \tau)$ 也是发送符号 a_k 和载波相位 θ 的函数。似然函数由式(6.257)定义。为了进行进一步的研究, 应按如下方法去除 $L(a_k, \theta, \tau)$ 对发送数据序列 $|a_k|$ 及载波相位 θ 的依赖性。

在 $[0, 2\pi]$ 内的所有可能值上对似然函数 $L(a_k, \theta, \tau)$ 取平均, 而不是对数。假设在这个范围内 θ 是均匀分布的(这通常是可调整的), 因此有

$$\begin{aligned} L_{av}(a_k, \tau) &= \int_0^{2\pi} L(a_k, \theta, \tau) f_\theta(\theta) d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \exp\left(\frac{2}{N_0} \mathbf{x}_k^T(\tau) \mathbf{s}(a_k, \theta, \tau)\right) d\theta \end{aligned}$$

$L(a_k, \theta, \tau)$ 中的指数可表示为(见习题 6.49):

$$\begin{aligned} \frac{2}{N_0} \mathbf{x}_k^T(\tau) \mathbf{s}(a_k, \theta, \tau) &= \frac{2\sqrt{E}}{N_0} \operatorname{Re}[a_k^* \tilde{x}_k(\tau) e^{-j\theta}] \\ &= \frac{2\sqrt{E}}{N_0} \operatorname{Re}[|a_k \tilde{x}_k(\tau)| \exp(j(\arg[\tilde{x}_k(\tau)] - \arg[a_k] - \theta))] \quad (6.273) \\ &= \frac{2\sqrt{E}}{N_0} |a_k \tilde{x}_k(\tau)| \cos(\arg[\tilde{x}_k(\tau)] - \arg[a_k] - \theta) \end{aligned}$$

因此

$$\begin{aligned} L_{av}(a_k, \tau) &= \frac{1}{2\pi} \int_0^{2\pi} \exp\left(\frac{2\sqrt{E}}{N_0} |a_k \tilde{x}_k(\tau)| \cos(\arg[\tilde{x}_k(\tau)] - \arg[a_k] - \theta)\right) d\theta \\ &= \frac{1}{2\pi} \int_{-\arg[\tilde{x}_k(\tau)] + \arg[a_k]}^{2\pi - \arg[\tilde{x}_k(\tau)] + \arg[a_k]} \exp\left(\frac{2\sqrt{E}}{N_0} |a_k \tilde{x}_k(\tau)| \cos(\phi)\right) d\phi \quad (6.274) \end{aligned}$$

其中最后一行进行了如下替换:

$$\phi = \arg[\tilde{x}_k(\tau)] - \arg[a_k] - \theta$$

修正的零阶贝塞尔函数的定义为(见附录 3):

$$I_0(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{x \cos \varphi} d\varphi \quad (6.275)$$

因此,可将平均似然函数 $L_{av}(a_k, \tau)$ 表示为

$$L_{av}(a_k, \tau) = I_0\left(\frac{2\sqrt{E}}{N_0} |a_k \tilde{x}_k(\tau)|\right) \quad (6.276)$$

其中 $\tilde{x}_k(\tau)$ 为接收机中匹配滤波器对应于第 k 个发送符号 a_k 的输出的复包络。对于 M 进制 PSK, 有

$$|a_k| = 1 \quad \text{对于所有 } k$$

因此,式(6.276)可简化为

$$L_{av}(a_k, \tau) = I_0\left(\frac{2\sqrt{E}}{N_0} |\tilde{x}_k(\tau)|\right) \quad (6.277)$$

由此可见,在 M 进制 PSK 中,似然函数不但完成了对载波相位 θ 的平均,同时也去除了对发送符号 a_k 的依赖性。

最后,考虑发送 L_0 个独立符号 $a_0, a_1, \dots, a_{L_0-1}$ 的情况,此时可将关于 τ 的所有似然函数写为

$$\begin{aligned} L_{av}(\tau) &= \prod_{k=0}^{L_0-1} L_{av}(a_k, \tau) \\ &= \prod_{k=0}^{L_0-1} I_0\left(\frac{2\sqrt{E}}{N_0} |\tilde{x}_k(\tau)|\right) \end{aligned} \quad (6.278)$$

现在,就可以对 $L_{av}(\tau)$ 取对数,从而得到如下关于 τ 的对数似然函数:

$$\begin{aligned} l_{av}(\tau) &= \log L_{av}(\tau) \\ &= \sum_{k=0}^{L_0-1} \log I_0\left(\frac{2\sqrt{E}}{N_0} |\tilde{x}_k(\tau)|\right) \end{aligned} \quad (6.279)$$

为了进行进一步研究,需要估计出 $l_{av}(\tau)$ 。修正的零阶贝塞尔函数 $I_0(x)$ 可展开为如下的幂级数形式(见附录 3):

$$I_0(x) = \sum_{m=0}^{\infty} \frac{\left(\frac{1}{2}x\right)^{2m}}{(m!)^2}$$

对于较小的 x ,可将 $I_0(x)$ 近似地写为

$$I_0(x) \approx 1 + \frac{x^2}{4}$$

可利用下述近似式来进一步简化问题:

$$\begin{aligned}\log I_0(x) &\approx \log\left(1 + \frac{x^2}{4}\right) \\ &\approx \frac{x^2}{4} \quad \text{当 } x \text{ 很小时}\end{aligned}$$

在这个问题中,较小的 x 对应于小信噪比。在此情况下,可将式(6.279)近似地写为

$$I_{av}(\tau) \approx \frac{E}{N_0^2} \sum_{k=0}^{L_0-1} |\tilde{x}_k(\tau)|^2 \quad (6.280)$$

如前所述,这里的 $\tilde{x}_k(\tau)$ 为匹配滤波器对应于第 k 个发送符号的输出的复包络。

将 $I_{av}(\tau)$ 对群时延 τ 取微分可得

$$\begin{aligned}\frac{\partial I_{av}(\tau)}{\partial \tau} &= \frac{E}{N_0^2} \sum_{k=0}^{L_0-1} \frac{\partial}{\partial \tau} |\tilde{x}_k(\tau)|^2 \\ &= \frac{2E}{N_0^2} \sum_{k=0}^{L_0-1} \operatorname{Re}[\tilde{x}_k^*(\tau) \tilde{x}'_k(\tau)]\end{aligned} \quad (6.281)$$

其中 $\tilde{x}_k^*(\tau)$ 为 $\tilde{x}_k(\tau)$ 的复共轭, $\tilde{x}'_k(\tau)$ 为 $\tilde{x}_k(\tau)$ 对 τ 的导数。因此,可将定时恢复的误差信号定义为(下文中将对比例因子 $2E/N_0^2$ 进行说明):

$$e[n] = \operatorname{Re}[\tilde{x}_n^*(\tau) \tilde{x}'_n(\tau)]$$

其中,用 n 代替了 k ,以便与图 6.62 中的符号一致。用 $\hat{\tau}_n$ 代表在时刻 $t = nT$ 对未知时延 τ 的估计。接着引入定义

$$\tilde{x}_n(\tau) = \tilde{x}(nT + \hat{\tau}_n)$$

和

$$\tilde{x}'_n(\tau) = \tilde{x}'(nT + \hat{\tau}_n)$$

可将误差信号 $e(n)$ 写为

$$e[n] = \operatorname{Re}[\tilde{x}^*(nT + \hat{\tau}_n) \tilde{x}'(nT + \hat{\tau}_n)] \quad (6.282)$$

误差信号 $e[n]$ 的计算中需要用到两个滤波器:

1. 用于产生 $\tilde{x}_n(\tau)$ 的复匹配滤波器。
2. 用于产生 $\tilde{x}'_n(\tau)$ 的导数匹配滤波器。

接收机里已经配置有第一种滤波器,第二种则是新引入的。实际应用中,由导数匹配滤波器引入的附加计算复杂度是我们所不期望的。可通过有限差分来逼近导数 $\tilde{x}'_n(\tau)$,从而避免计算复杂度的增加:

$$\tilde{x}'(nT + \hat{\tau}_n) \approx \frac{1}{T} \left[\tilde{x}\left(nT + \frac{T}{2} + \hat{\tau}_{n+1/2}\right) - \tilde{x}\left(nT - \frac{T}{2} + \hat{\tau}_{n-1/2}\right) \right] \quad (6.283)$$

其中, $\hat{\tau}_{n \pm 1/2}$ 为在时刻 $nT \pm T/2$ 计算出的定时估计值。对此,需要进行进一步修正,以说明定时估计值是在符号持续时间 T 的整数倍时得到的更新值,并且可用的量化值只有 $\hat{\tau}_n$ 。因此,用 $\hat{\tau}_{n+1/2}$ 代替 $\hat{\tau}_n$ (τ 的最新估计值),并用 $\hat{\tau}_{n-1/2}$ 代替 $\hat{\tau}_{n-1}$ (τ 的前一个估计值)。因此式(6.283)可写为

$$\tilde{x}'(nT + \hat{\tau}_n) \approx \frac{1}{T} \left[\tilde{x}\left(nT + \frac{T}{2} + \hat{\tau}_n\right) - \tilde{x}\left(nT - \frac{T}{2} + \hat{\tau}_{n-1}\right) \right] \quad (6.284)$$

并最终将误差信号重新定义为

$$e[n] = \operatorname{Re} \left\{ \tilde{x}^*(nT + \hat{\tau}_n) \left[\tilde{x}\left(nT + \frac{T}{2} + \hat{\tau}_n\right) - \tilde{x}\left(nT - \frac{T}{2} + \hat{\tau}_{n-1}\right) \right] \right\} \quad (6.285)$$

其中比例因子 $1/T$ 也将在下文中进行说明。

现在可以明确表达出用于定时恢复的递归算法：

$$c[n+1] = c[n] + \gamma e[n] \quad (6.286)$$

其中, γ 为步长参数, 其中包括了 $2E/N_0^2$ 和 $1/T$ 。误差信号 $e[n]$ 由式(6.285)定义。 $c[n]$ 为一实数, 用于控制数控振荡器(NCO)的振荡器频率。由式(6.285)和式(6.286)得到的定时恢复算法的一种方法如图 6.64 所示。该方法类似于广泛用于定时恢复的早 - 晚门同步器的连续时间形式。因此, 该方法被称为无数据辅助早 - 晚延迟(NDA-ELD)同步器。在每次迭代中, 该同步器对匹配滤波器输出的三个连续抽样值 $\tilde{x}\left(nT + \frac{T}{2} + \hat{\tau}_n\right)$, $\tilde{x}(nT + \hat{\tau}_n)$ 和 $\tilde{x}\left(nT + \frac{T}{2} - \hat{\tau}_{n-1}\right)$ 进行操作。相对于第二个抽样值来说, 第一个抽样值是较早的, 而最后一个则是较晚的。

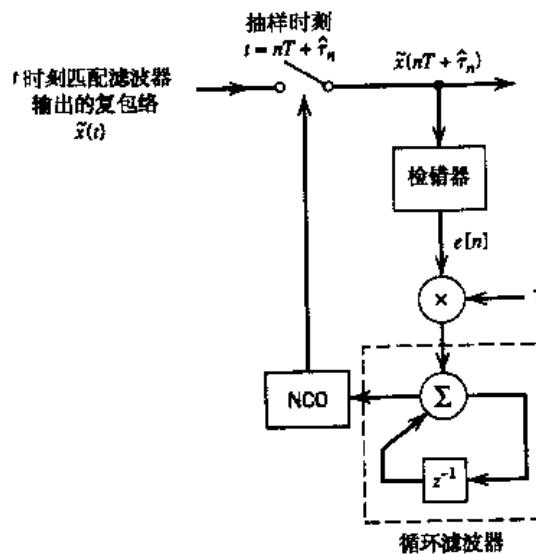


图 6.64 无数据辅助早 - 晚延迟同步器

采用附录 2 中介绍的带通到低通的变换, 可简化这一节的推导。但是, 为了便于理解, 我们并没有这样做。

6.15 计算机实验: 载波恢复与符号定时

本节通过一个具有下述特点的相干 QPSK 系统, 来阐明递归科斯塔环和无数据辅助早 - 晚延迟同步器的工作过程:

- (i) 信道响应: 滚降因子 $\alpha = 0.5$ 的升余弦线(奈奎斯特)。
(ii) 环路滤波器: 具有如下转移函数的一阶数字滤波器

$$H(z) = \frac{1}{z - (1 - \gamma A)} \quad (6.287)$$

其中, γ 为步长参数, 而 A 则为需要定义的参数。

- (iii) 环路带宽符号速率 $1/T$ 的 $B_L = 2\%$ 。也就是说, $B_L T = 0.02$ 。

6.15.1 实验 1: 载波相位恢复

为了研究递归科斯塔环的相位捕获特性, 需要用到相位误差发生器的所谓 S 曲线。 S 曲线的定义为相位误差固定条件下调节信号 $e[n]$ 的期望:

$$\varphi = \theta - \hat{\theta}$$

其中, θ 为载波相位的实际值, 而 $\hat{\theta}$ 为其估计值。也就是说

$$S(\varphi) = E[e[n] | \varphi] \quad (6.288)$$

通过实验方法, 在图 6.62 所示的递归科斯塔环的开环状态下, 测量调节信号 $e[n]$ 的平均值, 就能够得到 $S(\varphi)$, 如图 6.65 所示。

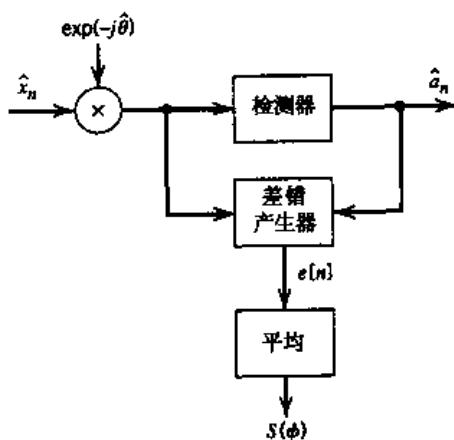


图 6.65 载波相位恢复中 S 曲线的测量方案

实现过程包括以下步骤。首先, 生成接收信号的复数包络

$$\tilde{x}_k(t) = \sqrt{\frac{2E}{T}} \exp(-j(2\pi f_c \tau_c - \alpha_k)) g(t - \tau_g) + \tilde{w}(t) \quad (6.289)$$

其中, $\alpha_k = 0, \pi/2, \pi, 3\pi/4$; τ_c 为载波时延, τ_g 为群时延, $\tilde{w}(t)$ 为复值信道噪声。整体信道响应 $g(t)$ 用奈奎斯特脉冲(见 4.5 节)表示:

$$g(t) = \frac{\sin(\pi t/T)}{(\pi t/T)} \cdot \frac{\cos(\pi \alpha t/T)}{1 - 4\alpha^2 t^2/T^2} \quad (6.290)$$

其中, $\alpha = 0.5$ 。如前所述, 假设符号定时(群时延 τ_g)已知, 现在的问题就是估计载波相位

$\theta = -2\pi f_c \tau_c$ 。 θ 的作用是按图 6.66 中的方式将信号星座图移位一个单元。

利用图 6.65 所示的实验过程, 就可以测量出 QPSK 系统的 S 曲线。图 6.67(a) 为无限大信噪比条件下的理想 S 曲线。由于在检测发送符号 a_k 过程中存在的不确定性, 该曲线在 $\varphi = \pm m\pi/4$ 时刻表现出不连续性, 其中, $m = 0, 1, 3, \dots$ 。信道噪声的存在则趋向于消除不连续性, 如图 6.67(b) 所示的实验测得的 S 曲线所示。图 6.67(b) 的结果是在 $E/N_0 = 10$ dB 条件下得到的。窄带噪声的同相和正交分量都具有零均值, 且具有与初始窄带噪声一样的高斯分布的方差。这两个分量定义了 $w(t)$ 。

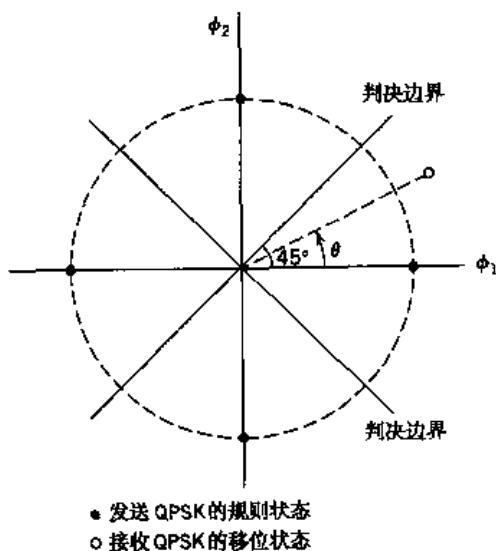


图 6.66 载波相位 θ 对 QPSK 信号状态的影响

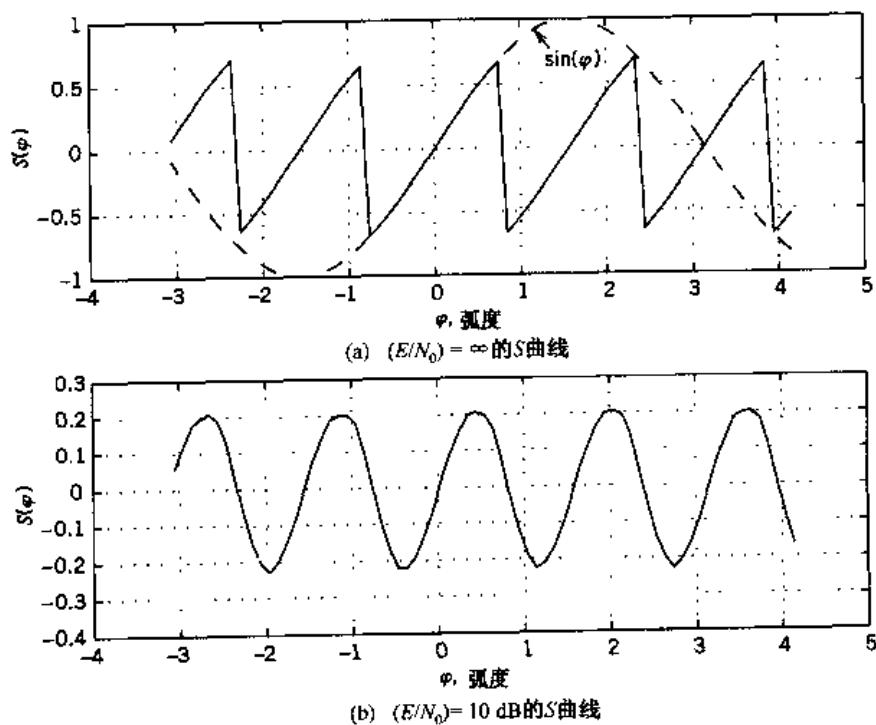


图 6.67 递归科斯塔环的性能

当建立起稳态条件时,估计相位值 $\hat{\theta}$ 将随真实值 θ 而波动。波动范围取决于步长参数 γ 和接收信噪比:

- (i) 图 6.68 所示为步长参数 γ 取两个不同值 0.1 和 0.5 时,相位误差 φ 相对于归一化时间 t/T 的变化,其中 $E/N_0 = 20 \text{ dB}$ 。由这幅图可清楚地看出, γ 越小, 相位误差 φ 的稳态波动也将越小。这种改进的代价是算法收敛速率的降低。为达到稳态, 该算法需要的迭代次数约为

$$L_0 \approx \frac{1}{2B_L T} \quad (6.291)$$

归一化带宽 $B_L T$ 可近似地表示为

$$B_L T \approx \frac{\gamma A}{4} \quad (6.292)$$

其中, A 为在原点处测得的 S 曲线的斜率。对于 $\gamma = 0.1$, $B_L T = 0.02$, 式(6.291)可得 $L_0 = 25$, 这与图 6.68 的实线曲线是一致的。更进一步, 由式(6.291)和式(6.292)可知, L_0 和 γ 成反比例, 这与图 6.68 中的结果也是一致的。

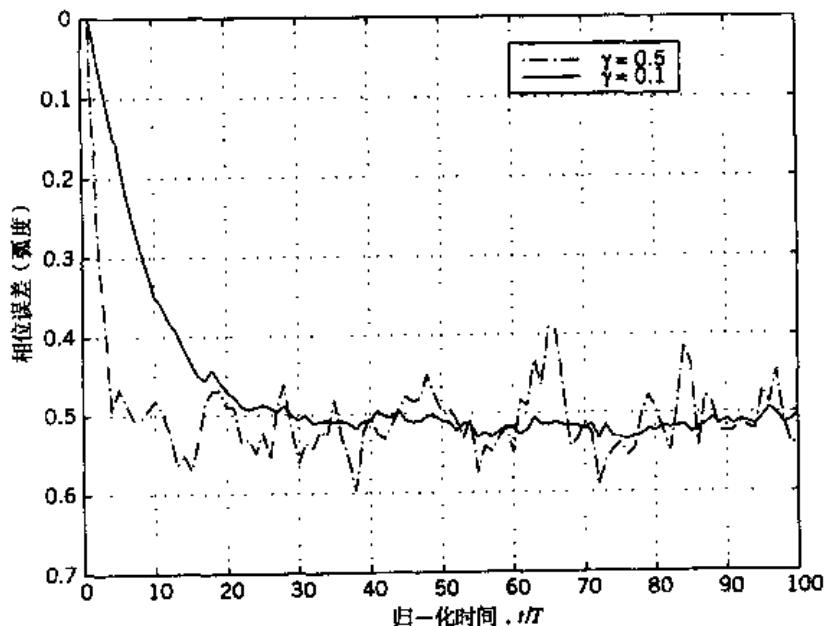


图 6.68 步长参数变化对递归科斯塔环收敛性能的影响

- (ii) 图 6.69 所示为 E/N_0 取三个不同值 5 dB, 10 dB 和 30 dB 时, 相位误差 φ 相对于归一化时间 t/T 的变化, 其中 $\gamma = 0.08$ 。由图可见, 信噪比越大, 相位误差 φ 的稳态波动就越小。更进一步, 算法的收敛速率也随着信噪比的增大而提高, 这是令人满意的。

图 6.70 为 $B_L T = 0.02$, $\gamma = 0.08$ 时相位误差的方差(对实验中的 100 个实验点取平均值而得)相对于 E/N_0 (以 dB 表示)的变化曲线。该图也显示了定义如下的修正的克拉默 - 拉奥边界(Cramér-Rao bound):^[24]

$$\text{MCRB}(\theta) = \frac{1}{2L_0(E/N_0)} \quad (6.293)$$

这是对普通克拉默 - 拉奥边界的修正。克拉默 - 拉奥边界是任意无偏估计值方差的下界。对克拉默 - 拉奥边界的修正是为了克服在实际同步问题中遇到的计算困难。图 6.70 中显示的实验结果与 $(E/N_0) \geq 10 \text{ dB}$ 时的理论结果是非常接近的。

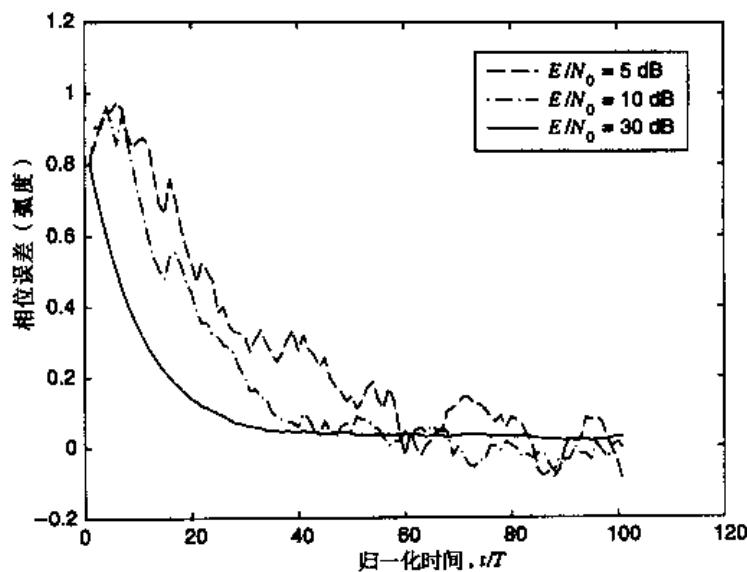


图 6.69 不同 E/N_0 下递归科斯塔环的收敛性能

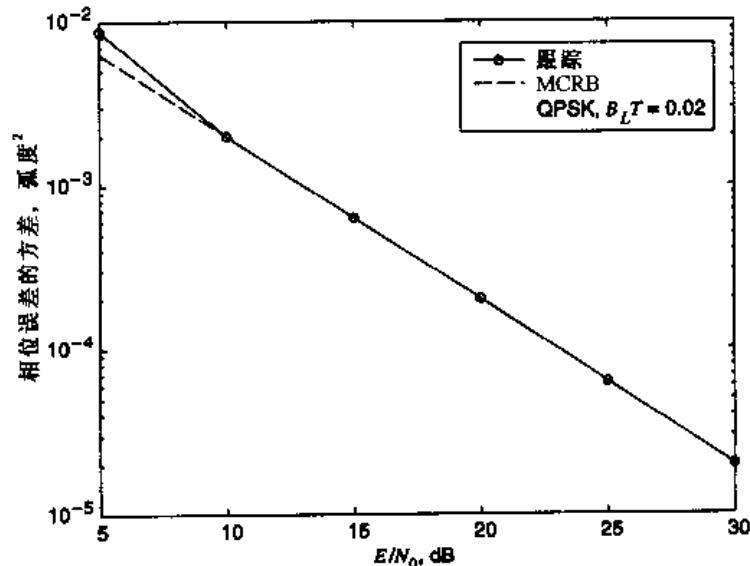


图 6.70 不同 E/N_0 下测得的递归科斯塔环跟踪误差的方差与理论值的比较

6.15.2 实验 2: 符号定时恢复

为测量无数据辅助早 - 晚延迟同步器进行定时恢复时的 S 曲线, 可采用图 6.71 所示的实验装置, 其中 $S(\delta)$ 中的 δ 代表定时偏移。图 6.72 中为在 $E/N_0 = 10 \text{ dB}$ 和 $E/N_0 = \infty$ 时用这种方法测得的 S 曲线。

图 6.73 显示了步长参数 γ 取两个不同值 $T/20$ 和 $T/200$ 时, 实验测得的符号定时误差归一化值相对于 E/N_0 的变化。该图也描绘了式(6.293)对应的修正克拉默 - 拉奥边界的理论曲线。

由图可见,随着步长参数 γ 的减小,归一化定时误差也相应减小,而由修正克拉默 - 拉奥边界决定的 E/N_0 范围(虽然是近似的)则增大。

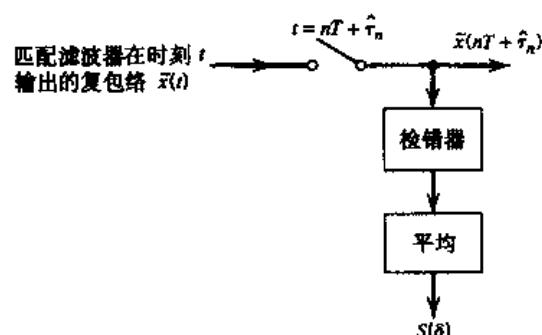


图 6.71 测量递归早 - 晚延迟同步器 S 曲线的方法

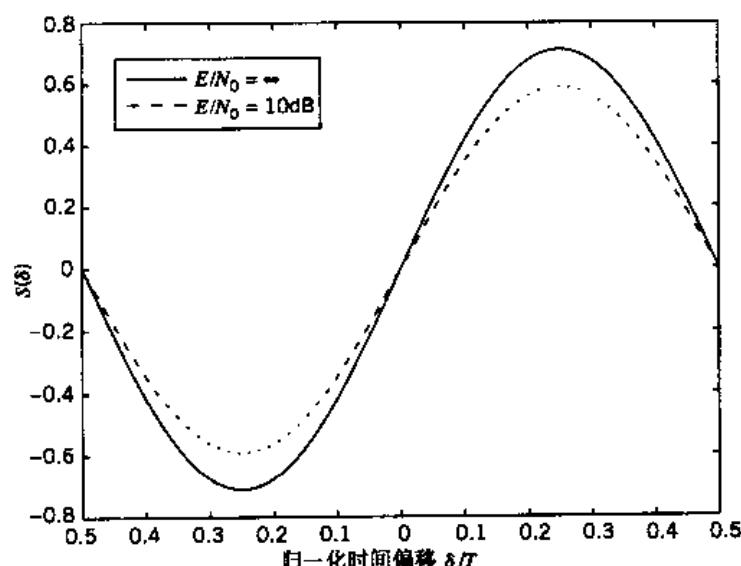


图 6.72 无噪声及噪声环境下测得的 NDA-ELD 同步器的 S 曲线

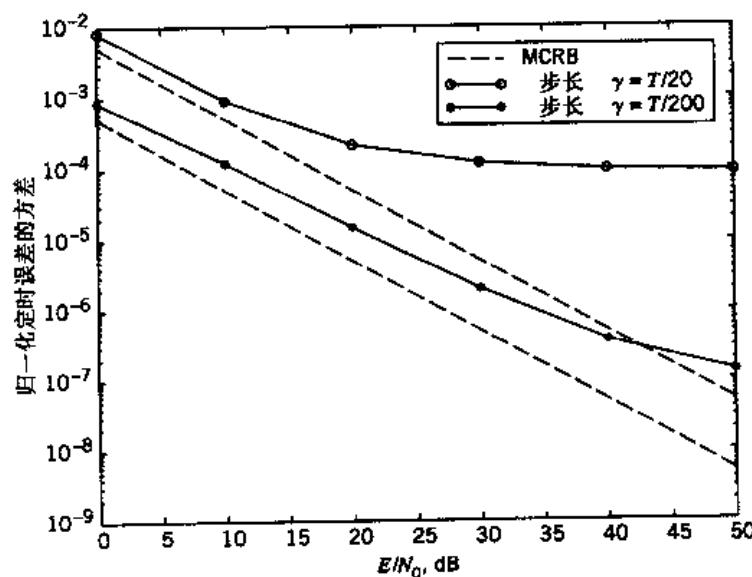


图 6.73 步长参数取两种值且 E/N_0 不同时,NDA-ELD 同步器跟踪误差的方差与理论值的比较

6.16 总结与讨论

本章利用第5章介绍的最佳接收机的基本背景理论,推导出了一些重要的数字调制技术在AWGN信道中的误比特率公式(或范围)。

1. 在PSK中,有:

- 相干BPSK。
- 相干QPSK及其变形,即偏移QPSK和 $\pi/4$ 移位QPSK。
- 相干 M 进制PSK,包括BPSK和QPSK这两个 $M=2$ 和 $M=4$ 的特例。相干 M 进制PSK用于数字卫星通信中。
- DPSK,可看成是PSK的伪非相干形式。

2. 相干 M 进制QAM,是结合了AM和PSK的混合调制方式。QPSK是 M 进制QAM在 $M=4$ 时的一个特例。 M 进制QAM是构造高速音频调制解调器的基础。

3. 在FSK中,有:

- 相十二进制FSK。
- 相干形式的MSK和高斯最小频移键控(GMSK)。其中GMSK是构造GSM无线通信的基础。
- 相干 M 进制FSK。
- 非相干二进制FSK。

本章还研究了两种可替换的通带数据传输技术:CAP和DMT。在AWGN信道中,CAP和DMT的性能是相同的,因为DMT可以看成是单载波调制信号的一个线性可逆变换。但在实际环境中,二者则体现出完全不同的性能。^[25] DMT在使用双绞线的ADSL中的应用已经被标准化。采用了判决反馈均衡的CAP则提供了解决ADSL问题的另一种方法。CAP也用于预分配系统的局域网数据传输。

DMT和OFDM都属于多信道调制。DMT和OFDM的基本差别是,DMT允许采用加载来优化信息容量,而OFDM则不允许。这一差别的产生是由于二者不同的应用领域。一般来讲,DMT用于ADSL之类的双向信道,而OFDM则用于广播和无线信道。

不考虑具体的数字调制系统,将接收机按发射机进行同步是系统工作的关键。无论接收机是相干的还是非相干的,符号定时恢复都是必须的。如果接收机是相干的,还需要提供载波恢复。本章后半部分讨论了用来满足这两种需要的无数据辅助同步器,并重点讨论了抑制载波的 M 进制PSK信号。介绍集中于适合数字信号处理的迭代同步技术。

注释与参考

- [1] 对于早期以几何学的观点为基础的各种不同的数字调制技术(ASK, FSK 和 PSK),可参阅 Arthurs 和 Dym(1962)。也可参阅下列书籍:
- Anderson(1998, Chapter 3)

- Benedetto 和 Biglieri(1999, Chapter 4, Chapter 5)
 - Lee 和 Messerschmitt(1994, 第 II 部分)
 - Proakis(1995, Chapter 5)
 - Sklar(1988, Chapter 3)
 - Viterbi 和 Omura(1979, pp.47 – 127)
- [2] 关于偏移 QPSK 的早期论文, 可参阅 Gitlin 和 Ho(1975)。
- [3] 在 Akaiwa 和 Negata 的公开演讲(1987)中, 第一次对 $\pi/4$ 移相 QPSK 进行了介绍。
- [4] Chennakeshu 和 Sauliner(1993)在一个数字无线通信环境中, 采用计算机仿真研究了 $\pi/4$ 移相 QPSK 的性能。产生 $\pi/4$ 移相 QPSK 信号时采用的脉冲波形信号是基于平方根升余弦谱的(见习题 4.38)。以后的论文证明了 $\pi/4$ 移相 QPSK 的性能在这种环境中将迅速衰减。图 6.13 的微分检测器源自 Chennakeshu 和 Sauliner(1993)。
- [5] 对式(6.65)的推导, 见 Cioffi(1998)。
- [6] 1975 年, Werner 在贝尔实验室的内部记录中首次对式(6.76)的推导进行了公布。稍后, Falconer(1975)发布了另一份贝尔实验室的记录, 其中指出: 如果不要求和已有的 QAM 或者其他带通信号匹配的话, 实际上并不需要进行符号旋转, 从而简化了 CAP 信号的数学表示(因此也简化了其实现), 如式(6.77)所示。然而, 直到 1987 年, 贝尔实验室的 Garry Smith 在权威论文中用 CAP 代替了无载波幅度/相位调制, 术语“CAP”才真正产生。在 Werner 的报告中(1992, 1993)第一次对 DSL 中的 CAP 进行了详细介绍。在 Chen, Im 和 Werner 各自的论文中(1992), 分别对 DSL 上 CAP 的可行性进行了研究。对此也可参见 Chen(1998, pp.461 – 473)。在 Im 和 Werner 的论文中(1995), 对 CAP 在局域网中的应用(包括长度小于 100 m 的情况下双绞线的应用)进行了讨论。100 m 的最大长度是按金属线标准规定的。
基带均衡器的数字使用与图 6.24 所示的在 Mueller 和 Werner(1982)中讨论的 CAP 接收机相似。
- [7] 在 Doelz 和 Heald 的论文中(1961), 第一次对 MSK 信号进行了讨论。在 Pasupathy(1979)中, 对 MSK 进行了研究并将其与 QPSK 进行了比较。由于 MSK 的频率间隔仅为二进制 FSK 信号相干检测的传统间隔 $1/T_b$ 的一半, 这种发信方式被称为快速 FSK, 见 deBuda(1972)。
- [8] 对高斯 MSK 的早期讨论, 见 Murota 和 Hirade(1981)和 Ishizuka 和 Hirade(1980)。
- [9] 除矩形调制脉冲的情况外, 数字 FM 的功率谱密度的规范都很难确定。Garrison(1975)的论文提出了一个基于为调制脉冲选择合适的有限持续时间/电平量化近似值的过程。其中提出的公式特别适用于数字 FM 信号功率谱的机器计算。见 Stüber(1996)。
- [10] Anderson 和 Salz 的论文中(1965)给出了任一频率偏移值下, 对 M 进制 FSK 频谱的详细分析。图 6.36 所示的结果为频率偏移 $k = 0.5$ 时, 该论文推导出的公式的特例。
- [11] 在 McDonough 和 Whalen 中(1995), 提出了推导非相干二进制 FSK 的误比特率的标准方法。在 Arthurs 和 Dym 中(1962), 提出了推导 DPSK 的误比特率的标准公式, 其中包括了对莱斯分布的应用。如果我们关注的是附加了高斯噪声的正弦信号形脉冲

的包络,就需要用到这种分布。6.6 节的推导避免了标准方法中的复杂问题。

- [12] 在 Simon 和 Divsalar(1992)中,对 DPSK 的最佳接收机进行了介绍。
- [13] 在 Lewart(1988)和 Hold(1997)中,对各种类型的调制解调器进行了技术讨论,并强调了它们的特性。
- [14] Wei(1984)将差分编码应用于卷积信道编码。其中,介绍了一些八状态卷积编码器,其编码输出与信号元素旋转无关。特别地,在论文的第 II 部分,Wei 介绍了应用于网格编码的 V.32 调制解调器,以及具有 90° 旋转不变性的卷积编码器的设计准则和过程。
- [15] Yen(1956),对带限信号的非均匀抽样进行了介绍。论文推导出的主要结果可以归纳为四个概括性的理论。式(6.188)就是基于 Yen 的论文的理论 III。
在 Kalet, Mazo 和 Saltzberg(1993)中,Yen 的这一理论被用于阐明设计双向数字调制解调器的基本原理。见 Ayanoglu 等人(1995)的专著。
- [16] 对数字调制解调器的第二种实现的介绍,参见 Humbert 和 Troulis(1996)。
- [17] 在 Forney 等人(1996)的著作中,对 V.34 高速调制解调器标准进行了详细描述。V.34 调制解调器中使用的网格图是基于 Wei(1984, 1987)的。
- [18] 多信道调制的思想可以追溯到 Chang(1966), Saltzberg(1967), 以及 Weinstein 和 Ebert(1971)的早期著作。在 Kalet(1989)中,对具有严重 ISI 的线性信道中的多音频调制的最优性进行了讨论。Cioffi 及其同事的工作,使 ADSL 的 DMT 有了统一的标准。细节请见 Ruiz 等人(1992)的著作, Chow 和 Cioffi(1995), Starr 等人(1999)著作的 7.2 节,以及 Cioffi(1998)的第 11 章。习题 6.44 改编自 Cioffi(1998)。
- [19] 在约束条件

$$y = f(x)$$

下,确定函数

$$\varphi(x) = 0$$

极值的拉格朗日因子方法源于下述理论:连续可微函数 $f(x)$ 有极值的充分必要条件是,它对 x 的微分在函数的临界(如最大和最小)点为零。因此,在 $f(x)$ 的临界点,有

$$\frac{\partial f}{\partial x} dx = 0 \quad (1)$$

更进一步,由于 $\varphi(x) = 0$,因此其微分同样为零,即

$$\frac{\partial \varphi}{\partial x} dx = 0 \quad (2)$$

将式(2)乘以参数 λ ,并将结果和(1)相加,可得

$$\left(\frac{\partial f}{\partial x} + \lambda \frac{\partial \varphi}{\partial x} \right) dx = 0$$

由于 dx 是一个独立增量,因此立刻可以推导出

$$\frac{\partial}{\partial x} (f(x) + \lambda \varphi(x)) = 0$$

这个公式是拉格朗日因子法的数学表述。参数 λ 被称为拉格朗日因子。这个注释里的内容源自 Sokolnikoff 和 Redheffer(1966, pp. 341 – 344)。

- [20] 在 DFT 中, 输入和输出由分别定义在时域和频域的非均匀分布的点上的数据序列组成。这一特性使 DFT 非常适合用 FFT 算法进行数字计算。FFT 算法与 DFT 的普通算法相比减少了很多算子, 因而效率很高。从本质上说, FFT 算法通过遵循“分解和解决”的策略, 也就是将原来的 DFT 计算过程连续分解为若干较小的 DFT 计算过程, 从而实现了计算的高效性。对于一个 N 点 DFT, $N = 2^L$ 的情况, FFT 算法需要 $L = \log_2 N$ 步计算, 其中, 每步计算都包括了复杂的 N 次乘法和加法。在 Oppenheim 和 Schafer(1989, Chapter 9) 中, 对 FFT 算法进行了详细介绍。
- [21] 对 VDSL 的概述见 Cioffi 等人(1999)的论文。该论文也对 VDSL 和音频调制解调器进行了比较性介绍。
- [22] 对 OFDM 及其应用的介绍, 参见 Casas 和 Leung(1991), LeFloch 等人(1989)的著作, 以及 Zou 和 Wu(1995)。对 OFDM 的研究注释以及参考列表, 参见 Cimini 和 Li(1999)。
- [23] 对传统同步系统的相位恢复和符号定时恢复的详细介绍, 参见 Stiffler(1971), 以及 Lindsey(1972) 和 Simon(1973, Chapter 2 和 Chapter 9)。对于强调离散时间信号处理算法的现代处理方法, 见 Mengali 和 D'Andrea(1997), Meyr, Moeneclaey 和 Fechtel(1998)。
- [24] 相位恢复时的修正克拉默 – 拉奥边界的式(6.293)的推导见 Mengali 和 D'Andrea(1997)。
- [25] Saltzberg(1998)讨论了 ADSL 应用中 CAP 和 DMT 的性能如何受到信道衰减和系统不完善的影响。其中考虑的衰减(不完善)包括脉冲噪声, 窄带干扰(例如空中 AM 无线电传输时射频的引入), 由于不完善的同步造成的定时抖动, 以及系统的非线性。

习题

ASK

- 6.1 在一个通断键控的 ASK 系统中, 发送幅度为 $\sqrt{2E_b/T_b}$ 的正弦载波对应于符号 1, 其中, E_b 为每比特的信号能量, T_b 为比特持续时间。不发送任何载波对应于符号 0。假设符号 1 和 0 等概地出现。

对于一个 AWGN 信道, 在下列假定下, 求该 ASK 系统的平均差错概率。

(a) 相干接收。

(b) 非相干接收, 具有大的比特信噪比 E_b/N_0 。

注意: 当 x 很大时, 修正零阶第一类贝塞尔函数可以近似地写为(见附录 3):

$$I_0(x) \approx \frac{\exp(x)}{\sqrt{2\pi x}}$$

PSK

- 6.2 PSK 信号和相位参考信号同时被送到相关器, 该相位参考信号和正确的载波相位有

不超过 φ 弧度的偏差。求相位误差 φ 对系统平均差错概率的影响。

- 6.3 考虑一个由乘法器、环路滤波器(VCO)和压控振荡器(VCO)组成的锁相环。输入乘法器的是定义如下的 PSK 信号：

$$s(t) = A_c \cos[2\pi f_c t + k_p m(t)]$$

其中, k_p 为相位灵敏度, 数据信号 $m(t)$ 对于二进制符号 1 和 0 分别取值 +1 和 -1。VCO 的输出为

$$r(t) = A_c \sin[2\pi f_c t + \theta(t)]$$

- (a) 假设该滤波器只去除载波频率为 $2f_c$ 的调制分量, 计算环路滤波器的输出。
 (b) 当相位已锁定即 $\theta(t) = 0$ 时, 证明环路滤波器的输出和数据信号 $m(t)$ 成正比例关系。

- 6.4 相干 PSK 系统的一个信号分量定义如下：

$$s(t) = A_c k \sin(2\pi f_c t) \pm A_c \sqrt{1 - k^2} \cos(2\pi f_c t)$$

其中 $0 \leq t \leq T_b$, “+”对应于符号 1 而“-”对应于符号 0。第一项代表为对接收机和发射机进行同步而包括的载波分量。

- (a) 绘制该方式的信号空间图。从图中能得到什么结论?
 (b) 当存在具有零均值且功率谱密度为 $N_0/2$ 的加性高斯白噪声时, 证明平均差错概率为

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}(1-k^2)}\right)$$

其中

$$E_b = \frac{1}{2} A_c^2 T_b$$

- (c) 假设发射信号能量的 10% 分配给载波分量。为了实现 10^{-4} 的差错概率, 确定所需要的 E_b/N_0 。
 (d) 比较这里的 E_b/N_0 和传统 PSK 系统为达到相同差错概率所需的 E_b/N_0 。
 6.5 (a) 假设输入二进制序列是 1100100010, 给出使用基于图 6.6 的信号组的 QPSK 得到的调制波形的同相分量和正交分量的波形。
 (b) 给出(a)中指定的二进制序列的 QPSK 波形。

- 6.6 用 P_{sI} 和 P_{sQ} 代表窄带数字通信系统同相和正交信道的符号差错概率。证明整个系统的平均符号差错概率为

$$P_e = P_{sI} + P_{sQ} - P_{sI}P_{sQ}$$

- 6.7 式(6.47)为相干 M 进制 PSK 平均符号差错概率的近似公式。这个公式是通过使用图 6.15(b)的信号空间图中的联合边界推导出的。假设发送信息点 m_1 , 证明可以直接由图 6.15(b)推导出近似式(6.47)。
 6.8 确定一个随机二进制序列产生的偏移 QPSK 信号的功率谱密度, 序列中符号 1 和 0 (分别用 ± 1 代表)等概地出现, 不同时间块内的符号统计独立且均匀分布。

6.9 第2章讨论的VSB提供了通带数据传输的另一种调制方法。

(a) 特别地,数字VSB传输系统可看成是以每秒 $2/T$ 维的速率工作的时变一维系统,其中 T 为符号持续时间。证明该结论的正确性。

(b) 证明数字VSB实际上和偏移QPSK具有相等的性能。

6.10 二进制数据流01101000被送入初始状态为($\phi_1 = \sqrt{E}$, $\phi_2 = 0$)的一个 $\pi/4$ 移位DQPSK调制器,如图6.11(a)所示。利用表6.2中总结的输入二位组和载波相位变化之间的关系,确定该数据流对应的调制器相位状态。

6.11 和普通QPSK调制器类似, $\pi/4$ 移位DQPSK调制器的输出可表示为如下同相分量和正交分量的形式:

$$s(t) = s_I(t) \cos(2\pi f_c t) - s_Q(t) \sin(2\pi f_c t)$$

用公式表示 $\pi/4$ 移位DQPSK信号的同相分量 $s_I(t)$ 和正交分量 $s_Q(t)$,从而归纳出产生 $\pi/4$ 移位DQPSK信号的方法。

6.12 $\pi/4$ 移位DQPSK信号的一个特性是可用FM鉴频器来解调。证明该特性。FM鉴频器在第2章中进行过介绍。

6.13 用 $\Delta\theta_k$ 表示 $\pi/4$ 移位DQPSK的差分编码相位。由此产生的符号对(I, Q)可以定义为

$$\begin{aligned} I_k &= I_{k-1} \cos(\Delta\theta_k) - Q_{k-1} \sin(\Delta\theta_k) \\ Q_k &= I_{k-1} \sin(\Delta\theta_k) + Q_{k-1} \cos(\Delta\theta_k) \end{aligned}$$

其中 I_k 和 Q_k 是对应于第 k 个符号的同相和正交分量。证明该关系对可以简化为

$$\begin{aligned} I_k &= \cos \theta_k \\ Q_k &= \sin \theta_k \end{aligned}$$

其中, θ_k 是第 k 个符号的绝对相位。

QAM

6.14 图6.53为一个240-QAM信号星座图。该图可看成是QAM十字形星座图的扩展形式。

(a) 指出图6.53中属于QAM正方形星座图的部分。

(b) 在(a)的基础上指出图6.53中属于QAM十字形星座图的部分。

(c) 指出图6.53中属于QAM十字形星座图扩展的部分。

6.15 证明256-QAM比64-QAM节省传输带宽和平均信号能量。

6.16 比较两个通带数据传输系统。一个系统使用16-PSK,另一个使用16-QAM。两个系统都要求得到 10^{-3} 的平均符号差错概率。比较这两个系统的信噪比要求。

CAP

6.17 二维CAP和M进制QAM方式很接近。

(a) 给定一个有指定幅度级数的QAM系统,推导出等价的CAP系统。

(b) 完成和(a)相反的操作。

6.18 证明共有 L 级幅度的一个 CAP 信号的功率谱定义如下：

$$S(f) = \frac{\sigma_A^2}{T} |P(f)|^2$$

其中, $|P(f)|$ 为通带同相脉冲 $p(t)$ 的幅度谱, σ_A^2 为复数符号 $A_i = a_i + jb_i$ 的方差, 其定义如下:

$$\sigma_A^2 = \frac{1}{L} \sum_{i=1}^L (a_i^2 + b_i^2)$$

6.19 给定一个具有滚降因子 α 的基带升余弦谱 $G(f)$, 叙述计算相应 CAP 信号特性的通带同相脉冲 $p(t)$ 和正交脉冲 $\hat{p}(t)$ 的频域过程。

FSK

6.20 在相干二进制 FSK 系统中, 分别用信号矢量 s_1 和 s_2 代表二进制符号 1 和 0。接收机在

$$\mathbf{x}^T s_1 > \mathbf{x}^T s_2$$

时, 判决为符号 1, 其中, $\mathbf{x}^T s_i$ 为观察矢量 \mathbf{x} 和信号矢量 s_i 的内积, $i = 1, 2$ 。证明判决准则等价于条件 $x_1 > x_2$, 其中, x_1 和 x_2 为观察矢量 \mathbf{x} 的分量。假设信号矢量 s_1 和 s_2 具有相等能量。

6.21 一个 FSK 系统以 2.5×10^6 bps 的速率传输二进制数据。在传输期间, 信号附加有均值为零, 功率谱密度为 10^{-20} W/Hz 的高斯白噪声。当不存在噪声时, 数字 1 或 0 对应的接收正弦信号的幅度为 1 mV。确定采用下列系统结构时的平均符号差错概率:

- (a) 相干二进制 FSK
- (b) 相干 MSK
- (c) 非相干二进制 FSK

6.22 (a) 在一个相干 FSK 系统中, 代表符号 1 和 0 的信号 $s_1(t)$ 和 $s_2(t)$ 的定义如下:

$$s_1(t), s_2(t) = A_c \cos \left[2\pi \left(f_c \pm \frac{\Delta f}{2} \right) t \right], \quad 0 \leq t \leq T_b$$

假设 $f_c > \Delta f$, 证明信号 $s_1(t)$ 和 $s_2(t)$ 的相关系数近似为:

$$\rho = \frac{\int_0^{T_b} s_1(t)s_2(t) dt}{\sqrt{\int_0^{T_b} s_1^2(t) dt}} \approx \text{sinc}(2\Delta f T_b)$$

- (b) 信号 $s_1(t)$ 和 $s_2(t)$ 正交时, 频偏 Δf 的最小值是多少?
- (c) 当 Δf 取何值时, 具有最小的平均符号差错概率?
- (d) 对于(c)中所得的 Δf 值, 确定使该相干 FSK 系统具有和相干二进制 PSK 系统相同的噪声性能所需要增加的 E_b/N_0 。

6.23 具有离散相位的二进制 FSK 信号定义如下:

$$s(t) = \begin{cases} \sqrt{\frac{2E_b}{T_b}} \cos\left[2\pi\left(f_c + \frac{\Delta f}{2}\right)t + \theta_1\right] & \text{发送符号1时} \\ \sqrt{\frac{2E_b}{T_b}} \cos\left[2\pi\left(f_c - \frac{\Delta f}{2}\right)t + \theta_2\right] & \text{发送符号0时} \end{cases}$$

其中, E_b 为每比特的信号能量, T_b 为比特持续时间, θ_1 和 θ_2 为均匀分布的随机变量在时间间隔 0 到 2π 内的抽样值。两个分别提供 $f_c \pm \Delta f/2$ 的发射频率的振荡器彼此独立工作。假设 $f_c \gg \Delta f$ 。

- (a) 计算 FSK 信号的功率谱密度。
 - (b) 证明当频率远离载波频率 f_c 时, 功率谱密度随着频率平方的倒数衰落。
- 6.24 利用式(6.104), 画出产生具有连续相位的桑德 FSK 信号 $s(t)$ 的框图, 式(6.104)如下:

$$s(t) = \sqrt{\frac{2E_b}{T_b}} \cos\left(\frac{\pi t}{T_b}\right) \cos(2\pi f_c t) \mp \sqrt{\frac{2E_b}{T_b}} \sin\left(\frac{\pi t}{T_b}\right) \sin(2\pi f_c t)$$

6.25 讨论 MSK 和偏移 QPSK 的相似性以及区别。

6.26 有两种检测 MSK 信号的方法。一种方法是使用相干接收机来充分利用 MSK 信号的载波相位信息。另一种方法是使用非相干接收机并忽略相位信息。第二种方法可以简化实现, 但它的噪声性能较第一种有所下降。为了实现 10^{-5} 的平均符号差错概率, 第二种方法需要比第一种方法增加多少 dB 的比特信噪比 E_b/N_0 ?

- 6.27 (a) 画出相应于输入二进制序列 1100100010 的 MSK 信号的同相和正交分量的波形。
 (b) 画出(a)中指定的二进制序列对应的 MSK 信号波形。

6.28 一个不归零数据流(幅度电平为 ± 1)通过其冲激响应由下面高斯函数定义的低通滤波器:

$$h(t) = \frac{\sqrt{\pi}}{\alpha} \exp\left(-\frac{\pi^2 t^2}{\alpha^2}\right)$$

其中, α 为根据滤波器的 3 dB 带宽设计参数:

$$\alpha = \sqrt{\frac{\log 2}{2}} \frac{1}{W}$$

(a) 证明滤波器的传输函数为

$$H(f) = \exp(-\alpha^2 f^2)$$

从而, 证明滤波器的 3 dB 带宽实际上等于 W 。可以利用表 A6.3 中的傅里叶变换对。

- (b) 证明滤波器对中心在原点的单位幅度的、持续时间为 T 的方波的响应为式(6.135)。
- 6.29 画出对应于二进制序列 1101000 的 GMSK 调制器产生的波形, 假设使用的增益带宽积 $WT_b = 0.3$ 。将结果和例 6.5 进行比较。
- 6.30 总结标准 MSK 和高斯滤波 MSK 信号之间的相似点和差别。

非相干接收机

6.31 在 6.8 节, 我们推导了非相干二进制 FSK 的误比特率公式, 现在回顾这个问题。假设用发射信号 $s_1(t)$ 代表二进制符号 1。根据 6.8 节, 有:

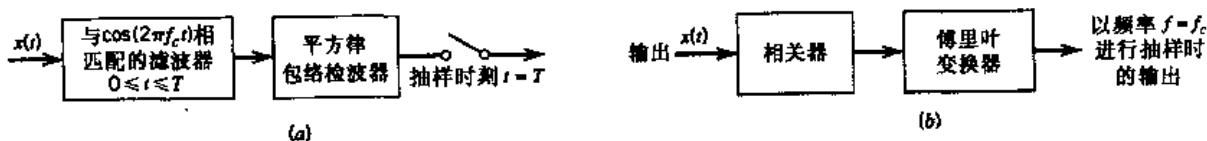
- 式(6.164)中, 由抽样值 l_2 代表的随机变量 L_2 是瑞利分布的。
- 式(6.170)中, 由抽样值 l_1 代表的随机变量 L_1 是莱斯分布的。

瑞利分布和莱斯分布在第 1 章中进行了讨论。利用第 1 章定义的概率分布, 推导出非相干二进制 FSK 的 BER 公式, 即式(6.181)。

6.32 图 P6.32(a) 为采用匹配滤波器来检测附加了高斯白噪声的、具有确定频率和随机相位的正弦信号的非相干接收机。该接收机的另一种实现是如图 P6.32(b) 所示的工作在频域的频谱分析接收机, 其中相关器具有有限时间自相关函数 $R_x(\tau)$:

$$R_x(\tau) = \int_0^{T-\tau} x(t)x(t+\tau) dt, \quad 0 \leq \tau \leq T$$

证明图 P6.32(a) 中平方律包络检波器在 $t = T$ 时刻的输出抽样值是图 P6.32(b) 中傅里叶变换器以 $f = f_c$ 进行频率抽样时的输出值的两倍。



P6.32

6.33 二进制序列 1100100010 被送入图 6.43(a) 的 DPSK 发射机。

- (a) 画出发射机输出端的波形。
 - (b) 将该波形送入图 6.43(b) 所示的 DPSK 接收机中, 证明在不存在噪声的情况下, 在接收机输出端重新得到的是原始的二进制序列。
- 6.34 差分 M 进制 PSK 是二进制 DPSK 的 M 进制扩展。符号时间 n 对应的调制器相角 θ_n 由关系式递归定义:

$$\theta_n = \theta_{n-1} + \left(\frac{2\pi}{M} \right) m_n, \quad \text{mod } 2\pi$$

其中, θ_{n-1} 为先前的相角, $m_n \in \{0, 1, \dots, M-1\}$ 为调制器当前的输入。该 M 进制调制方式的符号差错概率大约是

$$P_e \approx \operatorname{erfc} \left(\sqrt{\frac{2E}{N_0}} \sin \left(\frac{\pi}{2M} \right) \right), \quad M \geq 4$$

其中假设 E/N_0 足够大。

- (a) 在 $M \geq 4$ 时, 求为使差分 M 进制 PSK 获得和相干 M 进制 PSK 相同的符号差错概率, 每个符号的发送能量需要增大的倍数。
- (b) 当 $M = 4$ 时, 差分 QPSK 的性能比相干 QPSK 差多少 dB?

使用单个载波的数字调制方式的比较

6.35 二进制数据以 10^6 bps 的速率通过微波链路传输, 接收机输入端噪声功率谱密度是 10^{-10} W/Hz。确定采用

- (a) 相干二进制 PSK
- (b) DPSK

为维持平均差错概率 $P_e \leq 10^{-4}$, 所需要的平均载波功率。

6.36 为实现 $P_e = 10^{-4}$ 的平均符号差错概率, 相干二进制 PSK 和相干 FSK(传统的)系统需要的 E_b/N_0 值分别为 7.2 和 13.5。利用近似式

$$\text{erfc}(u) \approx \frac{1}{\sqrt{\pi u}} \exp(-u^2)$$

确定采用下列系统时, 为使 $P_e = 10^{-4}$, 所需的 E_b/N_0 的值:

- (a) 相干二进制 PSK 和 DPSK。
- (b) 相干二进制 PSK 和 QPSK。
- (c) 相干二进制 FSK(传统的)和非相干二进制 FSK。
- (d) 相干二进制 FSK(传统的)和相干 MSK。

6.37 6.10 节中以误比特率为基准, 比较了相干二进制 PSK, 相干二进制 FSK, QPSK, MSK, DPSK 和非相干 FSK 的噪声性能。本题采用平均符号差错概率 P_e 来进行比较。试述各调制方式下 P_e 相对于 E_b/N_0 的变化情况, 并评论你的结果。

6.38 带通信号的噪声等价带宽就是满足下面关系式的带宽值:

$$2BS(f_c) = P/2$$

其中, $2B$ 为在中心在中频 f_c 的噪声等价带宽, $S(f_c)$ 为信号在 $f = f_c$ 时功率谱密度的最大值, P 为信号的平均功率。证明二进制 PSK, QPSK 和 MSK 的噪声等价带宽分别如下:

调制类型	噪声带宽/比特率
二进制 PSK	1.0
QPSK	0.5
MSK	0.62

注意: 可以使用表 A6.10 中定义的积分。附录 2 中有对噪声等价带宽的讨论。

音频调制解调器

6.39 (a) 参考图 6.48(a)所示的差分编码器, 表 6.10 定义了不同输入二位组带来的 V.32 调制解调器的相位变化。通过引入差分编码器相应输出的先前值和当前值, 可扩展该表。对于每个输入二位组 $Q_{1,n}Q_{2,n}$, 差分编码二位组 $I_{1,n}I_{2,n}$ 及其先前值 $I_{1,n-1}I_{2,n-1}$ 都有四种可能的取值。

(b) 送入 V.32 调制解调器的非冗余编码的当前四位组为 0001。调制解调器先前的输出值为 01, 确定该调制解调器的输出码字及其坐标。

6.40 非冗余编码的 V.32 调制解调器标准采用正方形的 16-QAM 星座图。该模型的具体

参数如下：

载波频率 = 1800 Hz

符号速率 = 2400 波特

数据速率 = 9600 bps

试计算：

(a) 平均信噪比：

(b) 此调制解调器的平均符号差错概率。

其中假设 $E_s/N_0 = 20$ dB。

多信道线路码

6.41 考虑式(6.196)定义的通带基本函数，其中 $\phi(t)$ 如式(6.197)所定义。证明通带基本函数的性质 1, 2 和 3 的有效性。

6.42 加载问题的注水方案由满足式(6.210)的约束条件的式(6.213)定义。利用这对关系式，构造出在 N 个子信道间分配发射功率 P 的递归算法。分别以下列方式作为算法的开始：

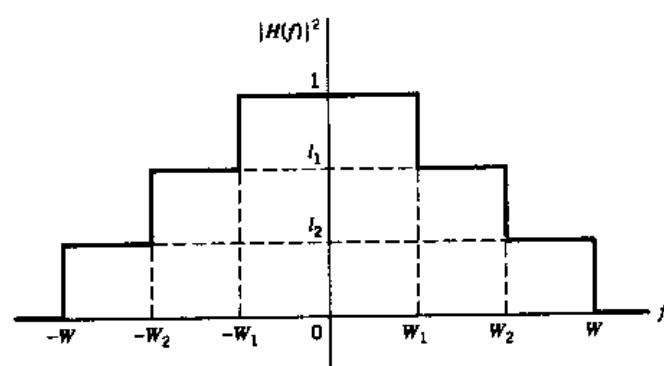
(a) 迭代系数 $i = 0$ 时的初始整体噪信比 $NSR(i) = 0$ ；

(b) 子信道按照分配的功率由最小到最大。

6.43 图 P6.43 为一个线性信道的平方幅度响应 $|H(f)|^2$ 的曲线。假设对所有子信道都有间隙 $\Gamma = 1$ 且噪声方差 $\sigma_n^2 = 1$ 。

(a) 推导在三个通带分别为 $(0, W_1)$, (W_1, W_2) 和 (W_2, W) 的子信道间最优的功率分配 P_1 , P_2 和 P_3 。

(b) 假设整体发射功率 $P = 10$, $l_1 = 2/3$, $l_2 = 1/3$, 计算 P_1 , P_2 和 P_3 。



P6.43

6.44 本题研究采用特征值分解(SVD)代替离散傅里叶变换(DFT)进行矢量编码的方法。该方法不需要对循环前缀，其中信道矩阵为

$$\mathbf{H} = \begin{bmatrix} h_0 & h_1 & h_2 & \cdots & h_v & 0 & \cdots & 0 \\ 0 & h_0 & h_1 & \cdots & h_{v-1} & h_v & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & h_0 & h_1 & \cdots & h_v \end{bmatrix}$$

其中,序列 b_0, b_1, \dots, b_v 代表信道冲激响应的抽样值。矩阵 \mathbf{H} 的特征值分解定义为

$$\mathbf{H} = \mathbf{U}[\Lambda : \mathbf{O}_{N,v}] \mathbf{V}^\dagger$$

其中 \mathbf{U} 为 $N \times N$ 维酉矩阵, \mathbf{V} 为 $(N+v) \times (N+v)$ 维酉矩阵;即

$$\mathbf{U}\mathbf{U}^\dagger = \mathbf{I}$$

$$\mathbf{V}\mathbf{V}^\dagger = \mathbf{I}$$

其中, \mathbf{I} 为单位矩阵,上标 \dagger 代表埃尔米特共轭变换。 Λ 是特征值为 $\lambda_n, n = 1, 2, \dots, N$ 的 $N \times N$ 维对角矩阵。 $\mathbf{O}_{N,v}$ 为 N 维零矩阵。

(a) 使用特征值分解,证明使用矢量编码的 N 个子信道的输出为

$$\mathbf{X}_n = \lambda_n \mathbf{A}_n + \mathbf{W}_n$$

其中, X_n 为矩阵积 $\mathbf{U}^\dagger \mathbf{x}$ 的一个分量, \mathbf{x} 为接收信号(信道输出)矢量。 A_n 为第 n 个符号 $a_n + jb_n$, \mathbf{W}_n 为由于信道噪声引入的随机变量。

(b) 证明矢量编码的信噪比为

$$(\text{SNR})_{\text{vector coding}} = \Gamma \left(\prod_{n=1}^{N^*} \left(1 + \frac{(\text{SNR})_n}{\Gamma} \right) \right)^{1/(N+v)} - \Gamma$$

其中 N^* 为分配的发射功率为非负值的信道数, $(\text{SNR})_n$ 为子信道 n 的信噪比, Γ 为指定的间隙。

(c) 随着块长 N 接近于无穷,特征值接近信道傅里叶变换的幅值。利用这一结论,总结矢量编码和DMT的关系。

6.45 在下列信道干扰下比较DMT和CAP的性能:

(a) 突发噪声

(b) 窄带干扰

其中假设:

(1) DMT有大量的子信道数。

(2) CAP系统是未编码的,其接收机使用一对自适应滤波器来实现。

6.46 OFDM技术可看成是 M 进制FSK的概括形式。证明上述结论的合理性。

同步

6.47 图P6.47为 M 进制PSK接收机中用于相位恢复的连续时间 M 次环框图。

(a) 证明该 M 次方律设备的输出包括频率为 Mf_c 的分量,其中, f_c 为初始载频。

(b) 锁相环中的振荡器的频率为 Mf_c ,说明这样设置的合理性。

(c) M 次环由于在间隔 $[0, 2\pi]$ 内有 M 进制的相位模糊,因而存在相位模糊的问题。

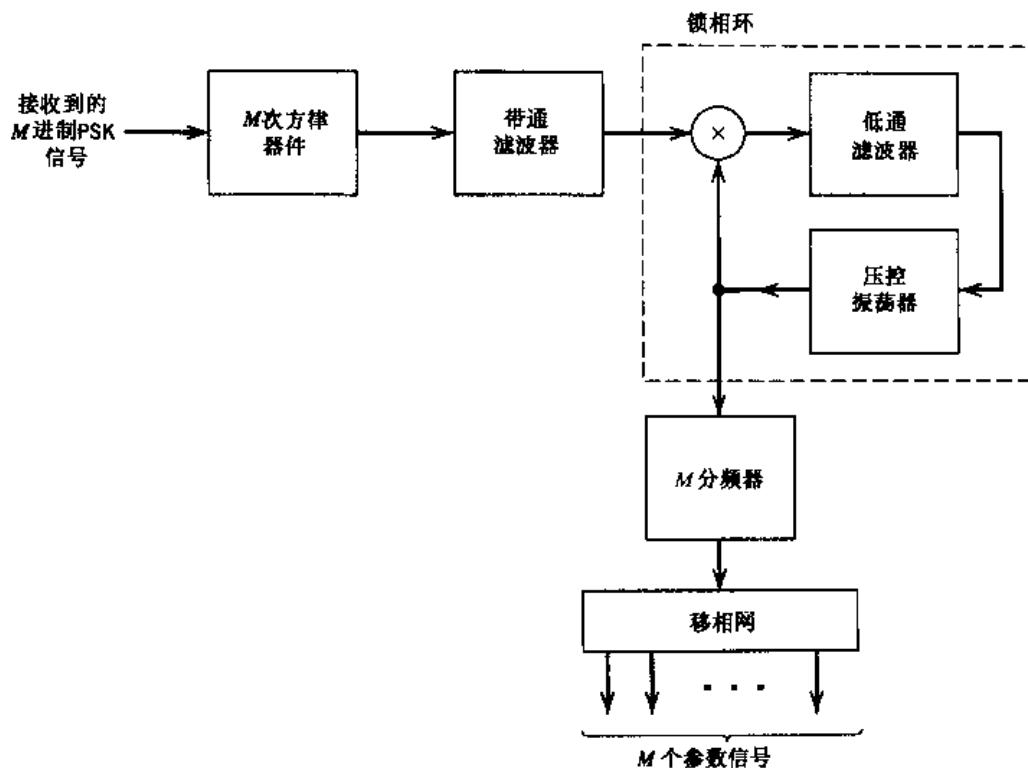
解释 M 次环的这个问题是如何产生的。如何克服该问题?

6.48 (a) 在式(6.272)所示的递归算法中,载波相位 θ 的旧估计值 $\hat{\theta}[n]$ 和当前估计值 $\hat{\theta}[n+1]$ 都用弧度表示。给出误差信号 $e[n]$ 和步长参数 γ 的单位。

(b) 在式(6.286)所示的递归算法中,控制信号 $c[n]$ 和 $c[n+1]$ 都是无量纲的。给出误差信号 $e[n]$ 和步长参数 γ 的单位。

6.49 利用式(6.264)和式(6.265)对 \tilde{x}_k 和 a_k 的定义,证明似然函数 $L(a_k, \theta, \tau)$ 的指数可

以表示为式(6.273)。



P6.47

6.50 6.14节研究了用于载波相位恢复的，基于式(6.260)的对数似然函数的无数据辅助方法。本题研究该公式用于数据辅助载波相位恢复的情况。

- (a) 考虑一个为线性调制系统设计的接收机。假设接收机已知长度 L_0 的预定值，证明最大似然估计的载波相位定义如下：

$$\hat{\theta} = \arg \left\{ \sum_{k=0}^{L_0-1} a_k^* \bar{x}(k) \right\}$$

其中 $\{a_k\}_{k=0}^{L_0-1}$ 为已知的复数符号序列， $\{\bar{x}(k)\}_{k=0}^{L_0-1}$ 为相应接收信号的复包络。

- (b) 利用(a)中得到的结果，给出最大似然相位估计器的框图。

计算机实验

6.51 本实验的目的是比较弥散信道对下列通带调制技术产生的波形的影响：

- (a) BPSK
- (b) QPSK
- (c) MSK
- (d) 时间-带宽积 $WT_b = 0.3$ 的高斯 MSK

信道中包括一个阶数 $2N = 10$, 3 dB 带宽 $2B$ 的中心频率在 f_c 的带通巴特沃兹滤波器。该信道的低通等价形式具有如下平方幅度响应：

$$|H(f)|^2 = \frac{1}{1 + (f/B)^{2N}}$$

信道带宽是可变的。试证明它对滤波后的调制波形的影响。

假设使用相干接收机,在(a),(b),(c)和(d)条件下,分别画出具有下列信道带宽的调制信号的波形:

- (i) $2B = 12 \text{ kHz}$
- (ii) $2B = 16 \text{ kHz}$
- (iii) $2B = 20 \text{ kHz}$
- (iv) $2B = 24 \text{ kHz}$
- (v) $2B = 30 \text{ kHz}$

并对结果做出评述。

提示:计算中可利用附录2中介绍的带通到低通的变换,将计算转移至基带进行。

第7章 扩频调制

本章介绍一种称为扩频调制的调制技术。扩频调制与前面各章介绍的调制技术有根本的差别。在扩频调制中,为了通信的安全牺牲了信道带宽和发射功率。

本章将涉及下列内容:

- 以伪噪声序列形式出现的扩频序列,其特性及产生方法。
- 扩频调制的基本概念。
- 两种普遍使用的扩频调制方法:直接序列和跳频。

本章内容是第8章介绍码分多址(CDMA)无线通信的基础。

7.1 引言

正如第4,5,6章所述,研究数字通信所关注的主要问题是如何有效地利用带宽和功率。尽管这两种主要的通信资源都十分重要,但在某些情况下有必要牺牲其效率以迎合其他的设计目标。例如,要求系统在敌对环境中提供一种安全的通信,从而使发射信号不容易被非法窃听者检测到或识别出。通过使用总称为扩频调制的一类调制技术就可以满足这一要求。

扩频通信系统的主要优点是,无论是在另一个用户试图同时通过同一信道发信引起的无意干扰环境下,还是在敌对发射者试图堵塞传输而造成的故意干扰环境下,它都能够抵抗干扰。

扩频调制^[1]的定义可以表述为两部分:

1. 扩频是一种使数据序列所占带宽超过发送它所需最小带宽的传输方法。
2. 扩频是在传输之前通过使用独立于数据序列的编码来完成的。同样的码用于接收机(已经与发射机进行了同步)以对接收信号进行解扩,从而恢复出原始数据序列。

尽管FM和PCM之类的标准调制技术都可以满足上述定义的第一部分,但它们都不能满足定义的第二部分,因而不属于扩频技术。

扩频调制最初是为军用目的而开发出来的,主要关注的是抗干扰的问题。后来,一些民用领域也从扩频调制的独特性质中受益。例如,扩频调制可以在基于地面的移动无线环境中防止多径现象。扩频调制还可用在多址通信系统中,这类系统有大量需要共享信道的独立用户,且系统不提供额外的同步机制,例如包括必须与中心站进行通信的移动设备的地面无线环境。第8章将具体介绍此类应用。

本章将讨论扩频调制的原理,重点是直接序列和跳频这两种技术。在直接序列扩频技术中,需进行两步调制。首先,输入数据序列被用来调制宽带码,这种码将窄带数据序列变换为类似噪声的宽带信号。得到的宽带信号用PSK技术进行第二步调制。在跳频技术中,通过伪随机地改变载波频率来拓宽数据调制载波的频谱。在应用时,这两种技术都依赖于称为伪随机序列或者伪噪声序列的如同噪声一样的扩展码的采用。由于伪噪声序列是扩频调制的基础,在研究开始前,我们先介绍伪噪声序列的产生和性质。

7.2 伪噪声序列

伪噪声(PN)序列是一种具有类似噪声波形的周期二进制序列,通常由反馈移位寄存器产生,图7.1为该寄存器的通用框图。反馈移位寄存器是由一个普通移位寄存器和一个逻辑电路连接而成的多回路反馈电路,其中,普通移位寄存器由 m 个受同一时钟控制的触发器(二状态记忆设备)构成。在每个时钟脉冲,每个触发器的状态被移入下一级的触发器中。在每个时钟脉冲,逻辑电路都计算出触发器状态的布尔函数,计算结果将被反馈到第一级触发器的输入端,从而防止移位寄存器变空。如此产生的PN序列由移位寄存器的长度 m 、其初始状态和反馈逻辑共同决定。

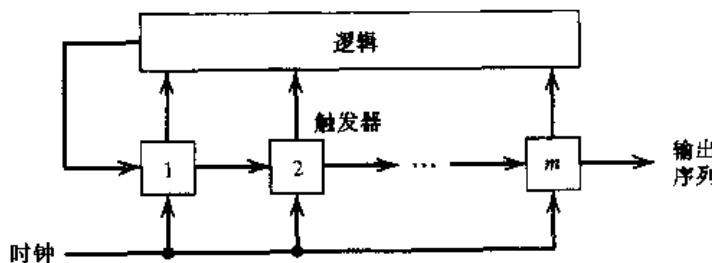


图 7.1 反馈移位寄存器

用 $s_j(k)$ 代表第 j 个触发器在第 k 个时钟脉冲之后的状态。这一状态可以用符号0或1代表。移位寄存器在第 k 个时钟脉冲后的状态则由 $|s_1(k), s_2(k), \dots, s_m(k)|$ 确定,其中 $k \geq 0$ 。对于初始状态, $k=0$ 。由对移位寄存器的定义可得

$$s_j(k+1) = s_{j-1}(k), \quad \begin{cases} k \geq 0 \\ 1 \leq j \leq m \end{cases} \quad (7.1)$$

其中 $s_0(k)$ 为第 k 个时钟脉冲后第一级触发器的输入。根据图7.1所示的配置, $s_0(k)$ 为单个状态 $s_1(k), s_2(k), \dots, s_m(k)$ 的布尔函数。对于指定的长度 m ,该布尔函数惟一确定后继的状态序列,因此,也就确定了在移位寄存器最后一级触发器输出端产生的PN序列。共有 m 个触发器的移位寄存器可能的状态数至多为 2^m ,因此,反馈移位寄存器产生的PN序列必须是周期性的,且周期不超过 2^m 。

当反馈逻辑完全由模2加法器组成时,该反馈移位寄存器就称为是线性的。此时,零状态(例如所有触发器都处于状态0)是不允许的。因为当处于零状态时,反馈逻辑产生的输入 $s_0(k)$ 将是0,移位寄存器将继续保持零状态,因此,输出将全是0。所以,具有 m 个触发器的线性反馈移位寄存器产生的PN序列的周期不可能超过 $2^m - 1$ 。当周期恰为 $2^m - 1$ 时,该PN序列被称为最大长度序列或者简称为 m 序列。

例 7.1

考虑图7.2中包括三个触发器的线性反馈移位寄存器。被送入第一级触发器的输入 s_0 等于 s_1 和 s_3 的模2和。假设移位寄存器的初始状态是100(三个触发器的内容从左到右排列),则后继状态将为

100,110,111,011,101,010,001,100,...

因此,输出序列(移位寄存器每个状态的最后一位)为

00111010...

输出以周期 $2^3 - 1 = 7$ 重复。

注意,选择 100 作为初始状态是随意的,另外 6 个允许状态都可以被选为初始状态,相应产生的输出序列只会有简单的循环移位。

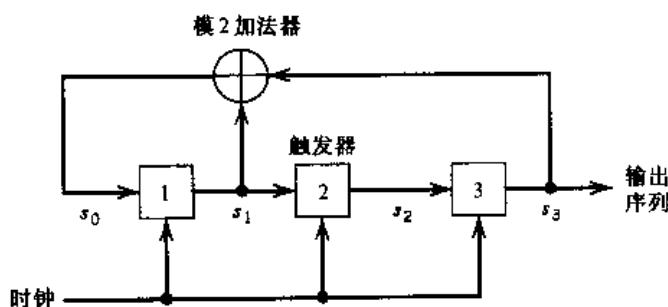


图 7.2 $m=3$ 的 m 序列发生器

7.2.1 m 序列^[2]的性质

m 序列具有随机二进制序列的许多性质。一个随机二进制序列是二进制符号 1 和 0 等概出现的序列。 m 序列的部分性质为:

1. 在 m 序列的每个周期里,1 的数目总是比 0 的数目多 1 个。这一性质被称为平衡性。
2. 在 m 序列的每个周期的连 1 游程和连 0 游程中,长度为 1 的占游程总数的一半,长度为 2 的占游程总数的四分之一,长度为 3 的占游程总数的八分之一,以此类推,直至游程总数取的分数不再代表有意义的数。这一性质称为游程特性。所谓“游程”是指在一个序列周期中,同样符号(1 或 0)组成的子序列。子序列的长度也就是游程的长度。长度为 m 的线性反馈移位寄存器产生的 m 序列,共有 $(N+1)/2$ 个游程,这里, $N=2^m-1$ 。
3. m 序列的自相关函数是周期性的二值函数。该性质称为相关性。

m 序列的周期定义如下:

$$N = 2^m - 1 \quad (7.2)$$

其中, m 为移位寄存器的长度。用电平 -1 和 +1 分别代表序列中的二进制符号 0 和 1,用 $c(t)$ 代表产生的 m 序列的波形,图 7.3(a)为 $N=7$ 时的波形示例。波形 $c(t)$ 的周期为

$$T_b = NT_c \quad (7.3)$$

其中, T_c 为 m 序列中分配给符号 1 或 0 的持续时间。根据定义,周期为 T_b 的周期信号 $c(t)$ 的自相关函数为

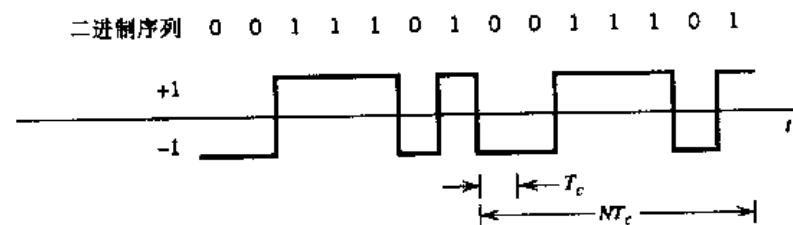
$$R_c(\tau) = \frac{1}{T_b} \int_{-T_b/2}^{T_b/2} c(t)c(t-\tau) dt \quad (7.4)$$

其中时延 τ 在时间间隔 $(-T_b/2, T_b/2)$ 内。式(7.4)为式(1.26)的特例。将式(7.4)用于 $c(t)$ 代

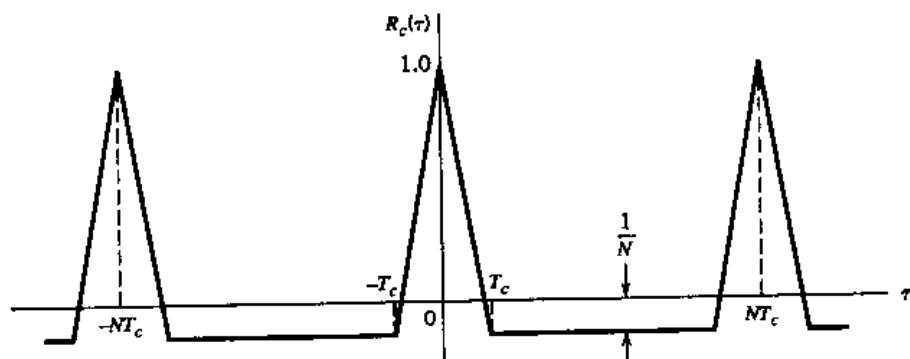
表的 m 序列, 可得

$$R_c(\tau) = \begin{cases} 1 - \frac{N+1}{NT_c} |\tau|, & |\tau| \leq T_c \\ -\frac{1}{N}, & \text{其他} \end{cases} \quad (7.5)$$

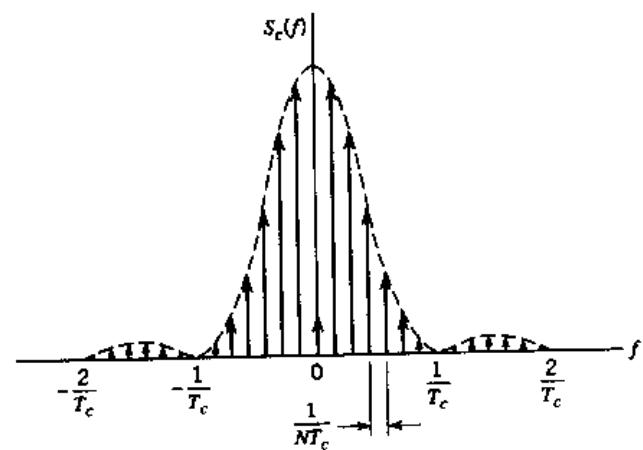
在图 7.3(b) 为 $m=3$ 或者说 $N=7$ 时的结果。



(a) $m=3$ 或周期 $N=7$ 的 m 序列波形



(b) 自相关函数



(c) 功率谱密度

图 7.3 图 7.2 所示反馈移位寄存器的输出

由傅里叶变换理论可知,时域的周期性对应于频域抽样点的均匀性。时域和频域的关系可通过 m 序列波形 $c(t)$ 的功率谱密度得到证实。特别地,利用式(7.5)的傅里叶变换,得到抽样点频谱:

$$S_c(f) = \frac{1}{N^2} \delta(f) + \frac{1+N}{N^2} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \operatorname{sinc}^2\left(\frac{n}{N}\right) \delta\left(f - \frac{n}{NT_c}\right) \quad (7.6)$$

图 7.3(c)为 $m=3$ 或者说 $N=7$ 时的抽样点频谱。

比较图 7.3 所示 m 序列的结果和图 1.11 所示随机二进制序列的相应结果,可得如下结论:

- 对于 m 序列的一个周期,自相关函数 $R_c(\tau)$ 和随机二进制波形的自相关函数是相似的。
- 两种序列波形的功率谱密度具有相同的包络 $\operatorname{sinc}^2(fT_c)$ 。它们之间的根本差别是,随机二进制序列具有连续谱密度,而 m 序列的谱密度则由相距 $1/NT_c$ Hz 的三角形函数组成。

随着移位寄存器长度 m ,或者说 m 序列的周期 N 的增大, m 序列和随机二进制序列越来越接近。实际上,在 N 取无穷大的极限情况下,这两种序列是完全一样的。然而,增大 N 的代价是存储量的增加,这使得实际可用的最大 N 值受到了限制。

7.2.2 选择 m 序列

现在,我们理解了 m 序列的性质,并知道它可由线性反馈移位寄存器来产生。下面要解决的关键问题是:如何寻找指定周期 N 的反馈逻辑?我们可以从第 10 章介绍的差错控制编码理论中找到解决方案。利用已经编辑好的、不同长度的移位寄存器需要的反馈连接扩展表,可以很容易找到需要的反馈逻辑。表 7.1 列举了对应移位寄存器长度 $m=2,3,\dots,8$ 的最大(反馈)抽头。^[3]需注意的是,随着 m 的增大,可用码的数量也在增大。此外,表中的各组反馈连接按时间反转得到的“镜像”组,也产生相同的 PN 码。

表 7.1 中加 * 的组,对应于 Mersenne 素数长度序列,该序列的周期 N 为素数。

表 7.1 移位寄存器长度为 2~8 的 m 序列

移位寄存器长度 m	反馈连接
2*	[2,1]
3*	[3,1]
4*	[4,1]
5*	[5,2],[5,4,3,2],[5,4,2,1]
6	[6,1],[6,5,2,1],[6,5,3,2]
7*	[7,1],[7,3],[7,3,2,1],[7,4,3,2],[7,6,4,2],[7,6,3,1],[7,6,5,2],[7,6,5,4,2,1],[7,5,4,3,2,1]
8	[8,4,3,2],[8,6,5,3],[8,6,5,2],[8,5,3,1],[8,6,5,1],[8,7,6,1],[8,7,6,5,2,1],[8,6,4,3,2,1]

例 7.2

考虑用长度 $m=5$ 的线性反馈移位寄存器产生的一个 m 序列。从表 7.1 中选择组 [5,2] 的反馈抽头, 相应的码生成器配置如图 7.4(a)所示。假设初始状态为 10000, 该方法产生的 m 序列的一个周期变化情况如表 7.2(a)所示, 由图可知; 经过 31 个循环后, 生成器回到了初始状态 10000。也就是说, 周期为 31, 这与式(7.2)的计算结果是一致的。

下面假设从表 7.1 中选择另一组反馈抽头, 即 [5,4,2,1], 相应的码生成器如图 7.4(b)所示。当初始状态为 10000 时, 可见, m 序列的变换如表 7.2(b)所示。生成器仍是在 31 个循环后回到了初始状态 10000, 也理应如此。但是产生的 m 序列则与表 7.2(a)中的不同。

显然, 图 7.4(a)所示的码生成器优于图 7.4(b)所示的码生成器, 因为前者只需要较少的反馈连接。

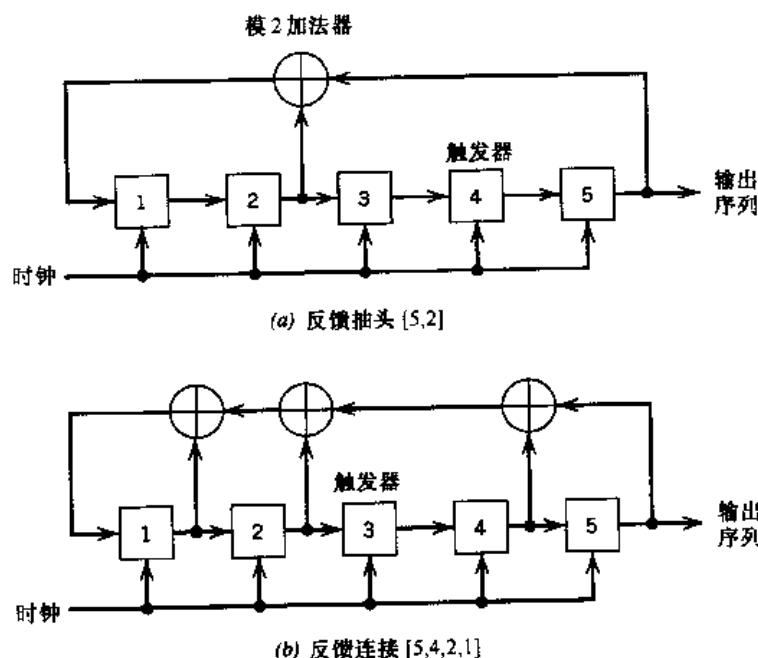


图 7.4 长度 $m=5$ 的反馈移位寄存器的两种不同配置

表 7.2(a) 图 7.4(a)所示反馈移位寄存器生成 m 序列的过程演示

反馈符号	移位寄存器状态					输出符号
	1	0	0	0	0	
0	0	1	0	0	0	0
1	1	0	1	0	0	0
0	0	1	0	1	0	0
1	1	0	1	0	1	0
1	1	1	0	1	0	1
1	1	1	1	0	1	0
0	0	1	1	1	0	1

(续表)

反馈符号	移位寄存器状态					输出符号
	1	0	0	0	0	
1	1	0	1	1	1	0
1	1	1	0	1	1	1
0	0	1	1	0	1	1
0	0	0	1	1	0	1
0	0	0	0	1	1	0
1	1	0	0	0	1	1
1	1	1	0	0	0	1
1	1	1	1	0	0	0
1	1	1	1	1	0	0
1	1	1	1	1	1	0
0	0	1	1	1	1	1
0	0	0	1	1	1	1
1	1	0	0	1	1	1
1	1	1	0	0	1	1
0	0	1	1	0	0	1
1	1	0	1	1	0	0
0	0	1	0	1	1	0
0	0	0	1	0	1	1
0	0	0	0	1	0	0
0	0	0	0	0	1	0
1	1	0	0	0	0	1

码:0000101011101100011111001101001

表 7.2(b) 图 7.4(b)所示反馈移位寄存器生成 m 序列的过程演示

反馈符号	移位寄存器状态					输出符号
	1	0	0	0	0	
1	1	1	0	0	0	0
0	0	1	1	0	0	0
1	1	0	1	1	0	0
0	0	1	0	1	1	0
1	1	0	1	0	1	1
0	0	1	0	1	0	1
0	0	0	1	0	1	0
0	0	0	0	1	0	0
0	0	0	0	0	1	0

(续表)

反馈符号	移位寄存器状态					输出符号
	1	0	0	0	0	
1	1	0	0	1	0	1
0	0	1	0	0	1	0
0	0	0	1	0	0	1
0	0	0	0	1	0	0
1	1	0	0	0	1	0
0	0	1	0	0	0	1
1	1	0	1	0	0	0
1	1	1	0	1	0	0
1	1	1	1	0	1	0
1	1	1	1	1	0	1
0	0	1	1	1	1	1
1	1	0	1	1	1	1
1	1	1	0	1	1	1
0	0	1	1	0	1	1
0	0	0	1	1	0	1
1	1	0	0	1	1	0
1	1	1	0	0	1	1
1	1	1	1	0	0	0
0	0	1	1	1	0	0
0	0	0	1	1	1	0
0	0	0	0	1	1	1
1	1	0	0	0	0	1

码:000011010100100010111101100111

7.3 扩频的概念

扩频调制的一个重要特点就在于它能够以有限的功率来抵抗外来干扰信号。干扰信号可能是为了破坏通信而送入接收机的,由具有相当功率的宽带噪声和多音波形组成。通过故意使信息承载信号占据的带宽远大于传输所需的最小带宽,就可以有效避免干扰波形的干扰。这使得发射信号具有类似噪声的特点,可以混入背景,因此,发射信号在信道中传播时可以不被任何监听者发现。因此,可将扩频看成是一种“伪装”信息承载信号的方法。

一种增大信息承载(数据)序列带宽的方法涉及到调制。用 $\{b_k\}$ 代表二进制数据序列, $\{c_k\}$ 代表PN序列。由于它们可取的两极电平大小一样,极性相反,即为 ± 1 ,故用波形 $b(t)$ 和 $c(t)$ 分别代表它们的极性不归零形式。用 $b(t)$ 代表信息承载(数据)信号,用 $c(t)$ 代表PN信号。通过将数据信号 $b(t)$ 和PN信号 $c(t)$ 送入乘积调制器或乘法器就可以进行所需的调制,如图7.5(a)所示。由傅里叶变换理论可知,两个信号相乘所得信号的频谱等于这两个信号频谱的卷积。因此,如果信息信号 $b(t)$ 是窄带的,而PN信号 $c(t)$ 是宽带的,乘积(调制)信号 $m(t)$ 将具有和宽带PN信号相似的频谱。换言之,在当前的应用中,PN序列担当了扩频码的角色。

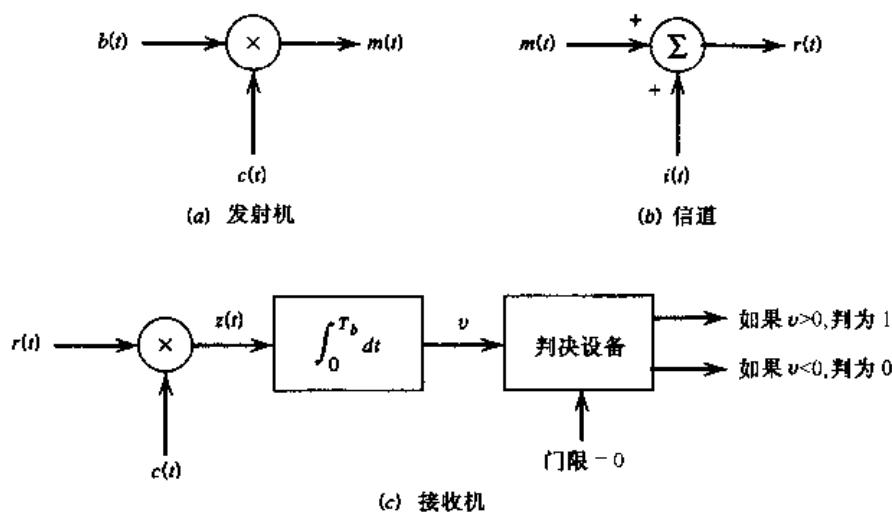


图 7.5 基带扩频系统的理想模型

由图 7.6 所示的波形可见,通过将信息承载信号 $b(t)$ 和 PN 信号 $c(t)$ 相乘,每个信息位都被分解为一些小的时间增量。这些小的时间增量通常称为码片。

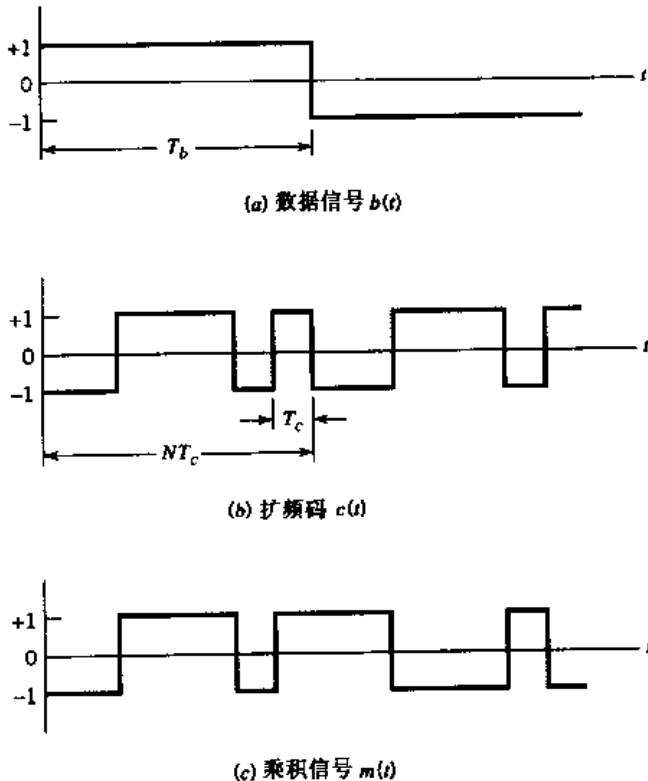


图 7.6 图 7.5(a)所示发射机的波形

对于基带传输,乘积信号 $m(t)$ 代表了发射信号。因此,可将发射信号表示为

$$m(t) = c(t)b(t) \quad (7.7)$$

接收信号 $r(t)$ 由发射信号 $m(t)$ 和用 $i(t)$ 代表的附加干扰组成,如图 7.5(b) 中的信道模型所示。因此

$$\begin{aligned} r(t) &= m(t) + i(t) \\ &= c(t)b(t) + i(t) \end{aligned} \quad (7.8)$$

为了恢复初始的信息信号 $b(t)$, 接收信号 $r(t)$ 被送入图 7.5(c) 所示的一个由乘法器、积分器和判决设备级联而成的解调器中。本地产生的与发射机中 PN 序列完全一致的 PN 序列也被送入乘法器。更进一步, 假设接收机和发射机的工作完全同步, 这就意味着接收机中的 PN 序列和发射机中的 PN 序列恰好对齐。因而, 接收机中乘法器的输出为

$$\begin{aligned} z(t) &= c(t)r(t) \\ &= c^2(t)b(t) + c(t)i(t) \end{aligned} \quad (7.9)$$

式(7.9)表明: 数据信号 $b(t)$ 两次和 PN 信号 $c(t)$ 相乘, 而非期望的信号 $i(t)$ 则只和 $c(t)$ 相乘一次。PN 信号 $c(t)$ 在 -1 和 +1 两者之间变化, 但二者的平方相等。因此

$$c^2(t) = 1 \quad \text{对所有 } t \quad (7.10)$$

可将式(7.9)简化为

$$z(t) = b(t) + c(t)i(t) \quad (7.11)$$

由式(7.11)可见, $b(t)$ 在接收机中乘法器输出端产生, 只不过还伴随有附加项 $c(t)i(t)$ 的干扰。把干扰 $i(t)$ 和本地产生的 PN 信号 $c(t)$ 相乘, 意味着扩频码对于干扰产生了与在发射机中对初始信号相似的影响。数据分量 $b(t)$ 是窄带的, 而分量 $c(t)i(t)$ 则是宽带的。因此, 通过将乘法器的输出送入一个带宽刚好能够适应数据信号 $b(t)$ 恢复需要的基带(低通)滤波器, 分量 $c(t)i(t)$ 的绝大部分功率将被滤除。因而, 干扰 $i(t)$ 对接收机输出端的影响也被大大降低。

在图 7.5(c) 所示的接收机中, 由积分器计算出乘法器输出信号下方的面积, 从而完成低通滤波操作。积分在比特持续时间 $0 \leq t \leq T_b$ 内进行, 产生抽样值 v 。最后, 接收机进行判决: 如果 v 大于门限值 0, 则接收机判定在时间间隔 $0 \leq t \leq T_b$ 内发送的是初始数据序列中的二进制符号 1; 如果 v 小于 0, 则接收机判决为发送了符号 0; 如果 v 恰等于 0, 则接收机随机判决为 1 或 0。

总而言之, 在发射机中使用扩频码(具有伪随机性质)产生的宽带发射信号, 对不知道扩频码的接收者来说类似于噪声。由 7.2 节的讨论可知, 对于给定的数据速率, 扩频码的周期越长, 发射信号也就越接近真正的随机二进制波形, 因而也就越难被检测到。自然, 为了提高抗干扰性能, 就不得不增大传输带宽、系统复杂度和处理时延。然而, 当主要目标是传输的安全性时, 这些代价都是值得的。

7.4 相干二进制 PSK 的直接序列扩频

前面各节介绍的扩频技术称为直接序列扩频, 其中关注的是基带传输。为了将这种技术用于卫星信道之类的通带传输, 可在发射机和接收机中结合相干二进制 PSK, 如图 7.7 所示。图 7.7(a) 所示的发射机首先把输入的二进制数据序列 $\{b_k\}$ 转换为极性 NRZ 波形 $b(t)$, 随后需进行两步调制。第一步调制是将数据信号 $b(t)$ (代表数据序列) 和 PN 信号 $c(t)$ (代表 PN 序列) 输入一个乘积调制器或者乘法器。第二步调制是采用一个二进制 PSK 调制器。因此,

发射信号 $x(t)$ 就是直接序列扩频二进制相移键控(DS/BPSK)信号。 $x(t)$ 的相位调制 $\theta(t)$ 根据 t 时刻信息信号 $b(t)$ 和 PN 信号 $c(t)$ 的极性, 按照表 7.3 的真值表取两个可能值 0 和 π 之一。

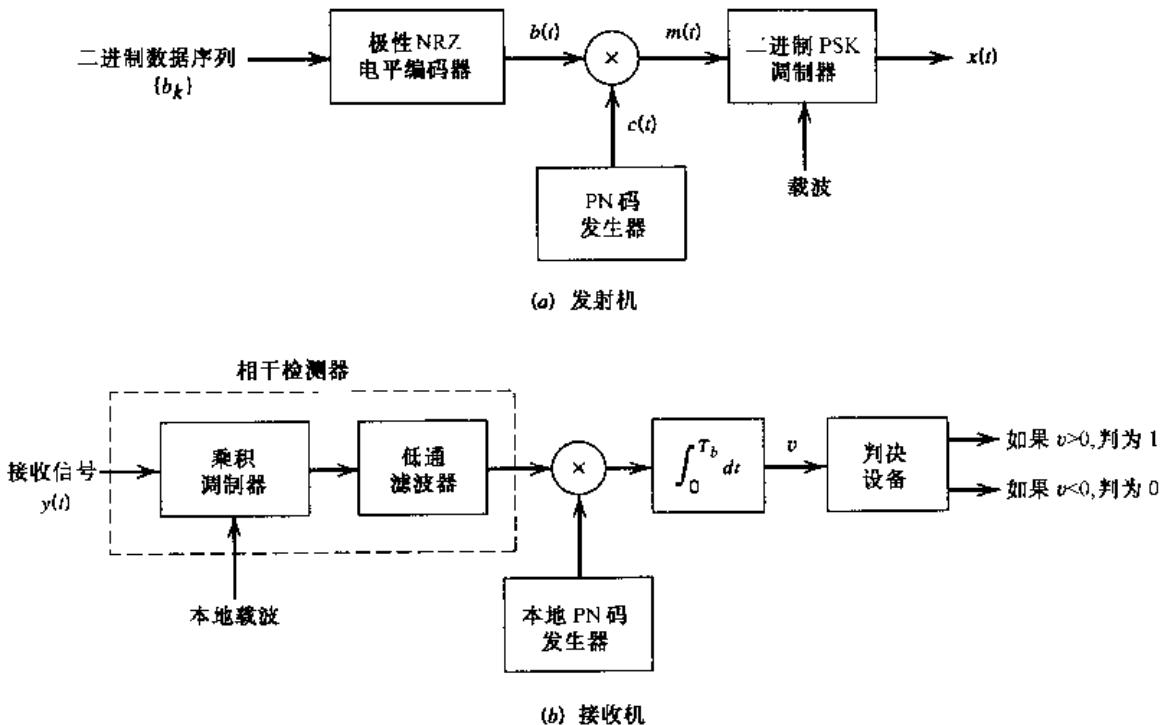


图 7.7 直接序列扩频相干 PSK

图 7.8 为第二步调制的波形, 其中包括图 7.6(c)中的部分调制波形。这里的波形对应于 PN 序列的一个周期。图 7.8(b)为正弦载波的波形, 而图 7.8(c)则为第二步调制产生的 DS/BPSK 波形。

图 7.7(b)所示的接收机中, 包括两步解调。第一步, 接收信号 $y(t)$ 和本地产生的载波被送入一个乘积调制器, 调制器后跟随一个带宽等于初始信息信号 $m(t)$ 谱宽的低通滤波器。这一步解调是对发射信号的 PSK 的逆过程。第二步解调是将低通滤波器的输出与本地产生的 PN 信号 $c(t)$ 的副本相乘来进行频谱解扩, 然后在比特持续时间 $0 \leq t \leq T_b$ 内进行积分, 最后, 按 7.3 节中介绍的方法进行判决。

表 7.3 PM 真值表

		数据序列 $b(t)$ 在时刻 t 的极性	
		+	-
PN 序列 $c(t)$ 在时刻 t 的极性	+	0	π
	-	π	0

7.4.1 分析模型

在图 7.7(a)所示的普通发射机中, 扩频在相位调制之前进行。但是我们发现, 如果将这两步操作交换顺序, 将更便于分析, 正如图 7.9 中的模型所示。可以这样做的原因是, 扩频和二进制 PSK 都是线性操作, 相位解调和频谱解扩也是一样。但为了使操作顺序的交换实际可

行,有必要对输入数据序列和 PN 序列进行同步。图 7.9 所示的模型中还包括了信道和接收机。在这个模型中,假设干扰 $j(t)$ 限制了系统性能,因此,信道噪声的影响可以忽略。于是,信道输出为

$$\begin{aligned} y(t) &= x(t) + j(t) \\ &= c(t)s(t) + j(t) \end{aligned} \quad (7.12)$$

其中, $s(t)$ 为二进制 PSK 信号, $c(t)$ 为 PN 信号。在图 7.9 所示的信道模型中,干扰信号用 $j(t)$ 表示。选择这一符号的目的是为了与图 7.5(b) 中的干扰表示形式相区分。图 7.9 中的信道模型是频域带通的,而在图 7.5(b) 中则是低通的。

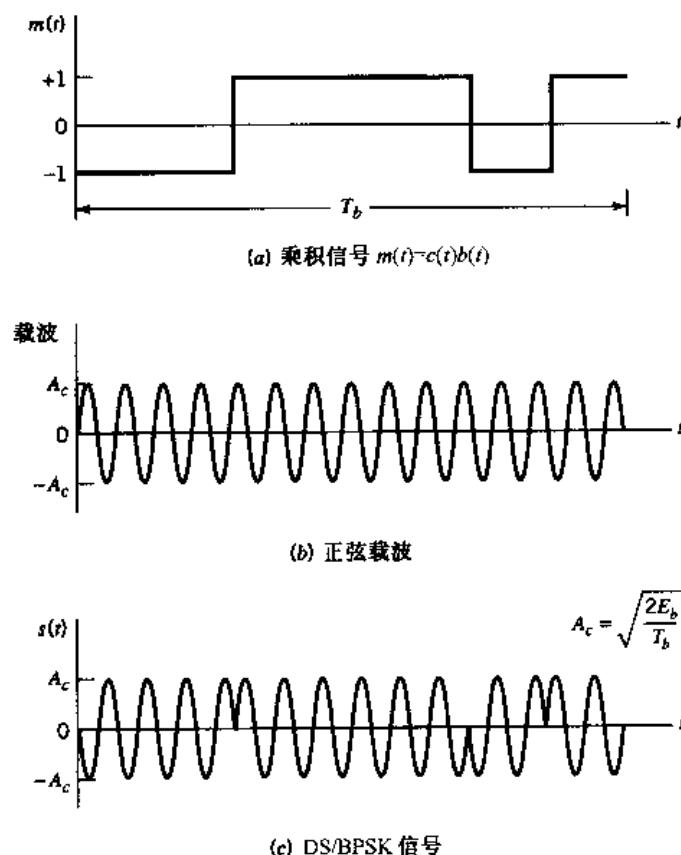


图 7.8 直接序列扩频第二步调制的波形

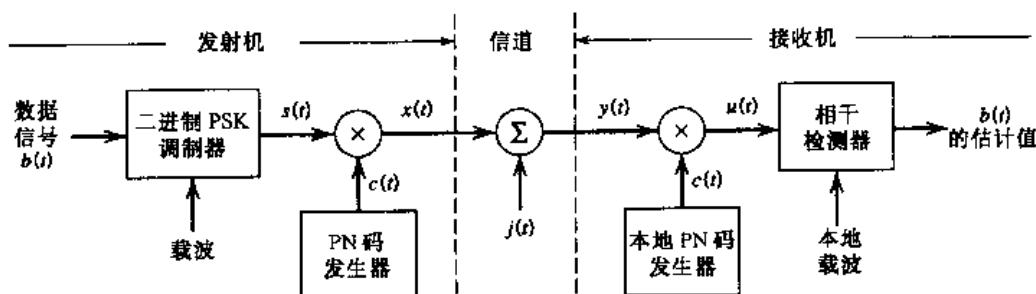


图 7.9 DS/BPSK 系统的模型

在接收机中,接收信号 $y(t)$ 首先与 PN 信号 $c(t)$ 相乘得到作为相干检测器输入信号的输

出 $u(t)$, 即

$$\begin{aligned} u(t) &= c(t)y(t) \\ &= c^2(t)s(t) + c(t)j(t) \\ &= s(t) + c(t)j(t) \end{aligned} \quad (7.13)$$

在式(7.13)的最后一行中, 不难注意到, PN 信号 $c(t)$ 满足式(7.10)中介绍的性质, 即

$$c^2(t) = 1 \quad \text{对所有 } t$$

式(7.13)表明: 相干检测器的输入 $u(t)$ 由一个包含 $c(t)j(t)$ 代表的加性码调制干扰信号的二进制 PSK 信号 $s(t)$ 组成。后面一项的调制使得干扰信号频谱扩展, 因此, 在接收机输出端对信息比特检测的可靠性得到了提高。

7.4.2 同步

为了正确操作, 扩频通信系统要求接收机中(用来解扩接收信号本地产生)的 PN 序列与发射机中(用来扩展发射信号)的 PN 序列之间必须进行同步。^[4] 该同步问题的解决方案包括两个部分: 捕获和跟踪。在捕获(也就是粗同步化)阶段, 两个 PN 码在尽可能短的时间内被对齐到码片的一帧内。一旦捕获了输入的 PN 码, 跟踪也就是精确同步即开始。PN 捕获过程通常包括两步。首先, 接收信号与本地产生的 PN 码相乘, 得到该 PN 码与发射机中使用的 PN 码之间的相关量。接着, 用合适的判决准则和搜索策略处理该相关量, 从而确定这两个码是否同步, 并确定不同步时应有的操作。跟踪是通过采用与产生本地相干载波参考时类似的相位锁定技术实现的, 二者之间的主要差别就在于实现鉴相的方法不同。

7.5 信号空间的维度和处理增益

在理解了扩频调制的概念及其实现方法后, 就可以对该技术进行更加细致的数学分析了。这里考虑采用的方法是基于第 5 章的信号空间理论的。特别地, 我们对发射信号和干扰信号(干扰者)都给出信号空间表示。

这里, 考虑正交基本函数集

$$\phi_k(t) = \begin{cases} \sqrt{\frac{2}{T_c}} \cos(2\pi f_c t), & kT_c \leq t \leq (k+1)T_c \\ 0, & \text{其他} \end{cases} \quad (7.14)$$

$$\tilde{\phi}_k(t) = \begin{cases} \sqrt{\frac{2}{T_c}} \sin(2\pi f_c t), & kT_c \leq t \leq (k+1)T_c \\ 0, & \text{其他} \end{cases} \quad (7.15)$$

$$k = 0, 1, \dots, N - 1$$

其中, T_c 为码片持续时间, N 为每比特的码片数。因此, 可将一个信息比特持续时间内的发射信号 $x(t)$ 表示为如下形式:

$$\begin{aligned}
 x(t) &= c(t)s(t) \\
 &= \pm \sqrt{\frac{2E_b}{T_b}} c(t) \cos(2\pi f_c t) \\
 &= \pm \sqrt{\frac{E_b}{N}} \sum_{k=0}^{N-1} c_k \phi_k(t), \quad 0 \leq t \leq T_b
 \end{aligned} \tag{7.16}$$

其中, E_b 为每比特的信号能量, “+”对应于信息位 1, 而“-”对应于信息位 0。码序列 $|c_0, c_1, \dots, c_{N-1}|$ 代表 PN 序列, 其中, $c_k = \pm 1$ 。发射信号 $x(t)$ 至少需要用 N 个正交函数表示, 因此是 N 维的。

下面考虑干扰信号(干扰者) $j(t)$ 的表示形式。理想情况下, 干扰信号的所有能量将恰好分布于发射信号 $x(t)$ 的 N 维信号空间里。否则将浪费部分能量。然而, 干扰者至多只能知道发射信号带宽, 但无法知道信号相位。因此, 可用通式来表示干扰信号:

$$j(t) = \sum_{k=0}^{N-1} j_k \phi_k(t) + \sum_{k=0}^{N-1} \tilde{j}_k \tilde{\phi}_k(t), \quad 0 \leq t \leq T_b \tag{7.17}$$

其中

$$j_k = \int_0^{T_b} j(t) \phi_k(t) dt, \quad k = 0, 1, \dots, N - 1 \tag{7.18}$$

且

$$\tilde{j}_k = \int_0^{T_b} j(t) \tilde{\phi}_k(t) dt, \quad k = 0, 1, \dots, N - 1 \tag{7.19}$$

因此, 干扰 $j(t)$ 是 $2N$ 维的, 即其维数为表示发送的 DS/BPSK 信号 $x(t)$ 所需维数的两倍。根据式(7.17), 可将干扰 $j(t)$ 的平均功率表示为

$$\begin{aligned}
 J &= \frac{1}{T_b} \int_0^{T_b} j^2(t) dt \\
 &= \frac{1}{T_b} \sum_{k=0}^{N-1} j_k^2 + \frac{1}{T_b} \sum_{k=0}^{N-1} \tilde{j}_k^2
 \end{aligned} \tag{7.20}$$

更进一步, 由于缺乏信号相位信息, 干扰者所能采用的最佳策略就是将相等的能量分配给式(7.18)和式(7.19)定义的余弦和正弦坐标。因此, 完全可以假设

$$\sum_{k=0}^{N-1} j_k^2 = \sum_{k=0}^{N-1} \tilde{j}_k^2 \tag{7.21}$$

于是, 可将式(7.20)简化为

$$J = \frac{2}{T_b} \sum_{k=0}^{N-1} j_k^2 \tag{7.22}$$

我们的目标是通过寻找图 7.9 所示的 DS/BPSK 接收机输入、输出端的信噪比, 并将上述结果结合在一起。为此, 利用式(7.13)将相干检测器的输出表示为

$$\begin{aligned}
 v &= \sqrt{\frac{2}{T_b}} \int_0^{T_b} u(t) \cos(2\pi f_c t) dt \\
 &= v_s + v_{ci}
 \end{aligned} \tag{7.23}$$

其中,分量 v_s 和 v_{ci} 分别对应于解扩的二进制 PSK 信号 $s(t)$ 和扩频干扰 $c(t)j(t)$ 。这两个分量分别定义为

$$v_s = \sqrt{\frac{2}{T_b}} \int_0^{T_b} s(t) \cos(2\pi f_c t) dt \quad (7.24)$$

和

$$v_{ci} = \sqrt{\frac{2}{T_b}} \int_0^{T_b} c(t)j(t) \cos(2\pi f_c t) dt \quad (7.25)$$

首先,考虑对应于信号的分量 v_s 。解扩的二进制 PSK 信号 $s(t)$ 等于

$$s(t) = \pm \sqrt{\frac{2E_b}{T_b}} \cos(2\pi f_c t), \quad 0 \leq t \leq T_b \quad (7.26)$$

其中,“+”对应于信息位 1,而“-”对应于信息位 0。因此,假设载频 f_c 是 $1/T_b$ 的整数倍,有

$$v_s = \pm \sqrt{E_b} \quad (7.27)$$

接下来,考虑对应于干扰的分量 v_{ci} 。将 PN 信号 $c(t)$ 直接表示为 $[c_0, c_1, \dots, c_{N-1}]$, 可相应地将式(7.25)写为如下的形式:

$$v_{ci} = \sqrt{\frac{2}{T_b}} \sum_{k=0}^{N-1} c_k \int_{kT_c}^{(k+1)T_c} j(t) \cos(2\pi f_c t) dt \quad (7.28)$$

利用式(7.14)求出 $\phi_k(t)$,再利用式(7.18)求出系数 j_k ,就可以将 v_{ci} 重新定义为

$$\begin{aligned} v_{ci} &= \sqrt{\frac{T_c}{T_b}} \sum_{k=0}^{N-1} c_k \int_0^{T_b} j(t) \phi_k(t) dt \\ &= \sqrt{\frac{T_c}{T_b}} \sum_{k=0}^{N-1} c_k j_k \end{aligned} \quad (7.29)$$

下面将 PN 序列近似为独立同分布的(i.i.d.)二进制序列。将式(7.29)重写为如下形式:

$$V_{ci} = \sqrt{\frac{T_c}{T_b}} \sum_{k=0}^{N-1} C_k j_k \quad (7.30)$$

其中, V_{ci} 和 C_k 分别为对应于抽样值 v_{ci} 和 c_k 的随机变量。在式(7.30)中,假设干扰是固定的。将 C_k 看成 i.i.d. 随机变量,得到的事件 $C_k = \pm 1$ 的概率为

$$P(C_k = 1) = P(C_k = -1) = \frac{1}{2} \quad (7.31)$$

于是,随机变量 V_{ci} 的均值为零,因为对于固定的 k ,有

$$\begin{aligned} E[C_k j_k | j_k] &= j_k P(C_k = 1) - j_k P(C_k = -1) \\ &= \frac{1}{2} j_k - \frac{1}{2} j_k \\ &= 0 \end{aligned} \quad (7.32)$$

对于代表了系数组 j_0, j_1, \dots, j_{N-1} 的固定的矢量 j , V_{ci} 的方差为

$$\text{var}[V_{ej}|j] = \frac{1}{N} \sum_{k=0}^{N-1} j_k^2 \quad (7.33)$$

由扩展因子 $N = T_b/T_c$, 可利用式(7.22)将上述方差表示为平均干扰功率 J 的函数, 即

$$\text{var}[V_{ej}|j] = \frac{JT_c}{2} \quad (7.34)$$

因此, 随机变量 V_{ej} 具有零均值, 且方差为 $JT_c/2$ 。

由式(7.27)不难看出, 相干检测器输出端(在每个比特持续时间内)的信号分量等于 $\pm\sqrt{E_b}$, 其中, E_b 为每比特的信号能量。因此, 信号分量的最大瞬时功率为 E_b 。从而, 可将输出信噪比定义为最大瞬时功率 E_b 与式(7.34)中平均噪声分量方差的比值, 因此有

$$(\text{SNR})_O = \frac{2E_b}{JT_c} \quad (7.35)$$

接收机输入端的平均信号功率等于 E_b/T_b , 因此, 可将输入信噪比定义为

$$(\text{SNR})_I = \frac{E_b/T_b}{J} \quad (7.36)$$

因此, 将式(7.35)和式(7.36)联立, 消去 E_b/J , 可将输出信噪比表示为输入信噪比的函数, 即

$$(\text{SNR})_O = \frac{2T_b}{T_c} (\text{SNR})_I \quad (7.37)$$

习惯上将信噪比表示为 dB 的形式。为此, 引入处理增益(PG), 处理增益就是采用扩频后在信噪比上获得的增益。特别地, 有

$$\text{PG} = \frac{T_b}{T_c} \quad (7.38)$$

这代表了扩频信号相对于非扩频信号的处理增益。因此, 可将式(7.37)写为如下等价形式:

$$10 \log_{10}(\text{SNR})_O = 10 \log_{10}(\text{SNR})_I + 3 + 10 \log_{10}(\text{PG}) \text{ dB} \quad (7.39)$$

式(7.39)右端的 3 dB 项说明了使用相干检测(假定接收机已知信号相位)在信噪比上获得的增益。信噪比的这一增益与是否采用扩频无关。实际上, 最后一项即 $10 \log_{10}(\text{PG})$ 就是处理增益。注意到处理增益 PG 和扩展因子 N (PN 序列长度)都等于 T_b/T_c 。也就是说, PN 序列越长(或者说相应 T_c 的码片时间越短), 处理增益也就越大。

7.6 差错概率

用图 7.9 所示的 DS/BPSK 系统的相干检测器的输出 v 代表随机变量 V 的抽样值, 并用外部干扰引起的等价噪声分量 v_{ej} 代表随机变量 V_{ej} 的抽样值。则由式(7.23)和式(7.27)可得

$$V = \pm\sqrt{E_b} + V_{ej} \quad (7.40)$$

其中, E_b 为每比特的发射信号能量。“+”对应于发送符号(信息比特)1, 而“-”对应于发送符号0。图 7.9 中相干检测器使用的判决准则是, 当检测器输出大于门限值 0 时, 判决时间间隔

$(0, T_b)$ 内接收的是 1; 如果检测器输出小于门限值, 则判决为 0; 如果检测器输出恰好为 0, 则接收机随机判决为 1 或 0。当两种信息比特等概出现时, 由于问题的对称性, 平均差错概率 P_e 等于当发送符号 0 时, 接收机判决为符号 1 的条件差错概率。也就是说

$$\begin{aligned} P_e &= P(V > 0 \mid \text{发送符号 } 0) \\ &= P(V_{ci} > \sqrt{E_b}) \end{aligned} \quad (7.41)$$

当然, 差错概率 P_e 依赖于式(7.30)中定义的随机变量 V_g 。根据定义, V_g 为 N 个均匀分布的随机变量之和。因此, 由中心极限理论可知, 当 N 很大时, 随机变量 V_g 将为高斯分布。实际上, 实用中直接序列扩频系统的扩展因子或者说 PN 序列的长度 N 都是很大的, 因此都可以应用中心极限理论。

前面计算了 V_g 的均值和方差, 参见式(7.32)和式(7.34)。因此, 可以说相干检测器输出中包含的等价噪声分量 V_g 可以近似为具有零均值且方差为 $JT_c/2$ 的高斯随机变量, 其中, J 为平均干扰功率, T_c 为码片持续时间。根据这一近似, 可继续计算事件 $V_g > \sqrt{E_b}$ 的概率, 并根据式(7.41)将平均差错概率表示成

$$P_e \approx \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{JT_c}}\right) \quad (7.42)$$

这个引用了高斯假设的简单公式对于具有很大扩展因子 N 的 DS/BPSK 二进制系统来说是很适用的。

7.6.1 抗干扰特性

比较式(7.42)和相干二进制 PSK 系统的平均差错概率公式, 即式 6.20:

$$P_e = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_b}{N_0}}\right) \quad (7.43)$$

由此可见, 在计算 DS/BPSK 系统的误比特率时, 干扰可被看成是功率谱密度为 $N_0/2$ 的宽带噪声, 其定义如下:

$$\frac{N_0}{2} = \frac{JT_c}{2} \quad (7.44)$$

这个关系式只是式(7.34)中结果的简单重复。

由于每比特的信号能量 $E_b = PT_b$, 其中, P 为平均信号功率, T_b 为比特持续时间, 可将单位比特信噪比表示为

$$\frac{E_b}{N_0} = \left(\frac{T_b}{T_c}\right)\left(\frac{P}{J}\right) \quad (7.45)$$

利用式(7.38)中对处理增益 PG 的定义, 可将结果写为

$$\frac{J}{P} = \frac{PG}{E_b/N_0} \quad (7.46)$$

比例式 J/P 称为干扰容限。因此, 用 dB 表示的干扰容限和处理增益之间的关系是

$$(\text{Jamming 余量})_{\text{dB}} = (\text{处理增益})_{\text{dB}} - 10 \log_{10} \left(\frac{E_b}{N_0} \right)_{\text{min}} \quad (7.47)$$

其中 $(E_b/N_0)_{\text{min}}$ 为提供指定的平均差错概率所需要的最小值。

例 7.3

扩频通信系统具有如下参数：

信息比特持续时间 $T_b = 4.095 \text{ ms}$

PN 码片持续时间 $T_c = 1 \mu\text{s}$

因此,由式(7.38)可得处理增益为

$$\text{PG} = 4095$$

因此,需要的 PN 序列的周期为 $N = 4095$, 移位寄存器的长度为 $m = 12$ 。

为了进行令人满意的接收,假设平均差错概率不超过 10^{-5} 。从相干二进制 PSK 接收机的公式可知, $E_b/N_0 = 10$ 对应的平均差错概率为 0.387×10^{-5} 。因此,利用这个 E_b/N_0 值,以及计算出的处理增益值,由式(7.47)可得干扰容限为

$$\begin{aligned} (\text{干扰容限})_{\text{dB}} &= 10 \log_{10} 4095 - 10 \log_{10}(10) \\ &= 36.1 - 10 \\ &= 26.1 \text{ dB} \end{aligned}$$

也就是说,即使接收机输入端的噪声或干扰高达接收信号功率的 409.5 倍,在接收机输出端,信息比特仍可以被可靠地检测出来。显然,这说明扩频通信具有极强的抗干扰性能。

7.7 跳频扩频

在 7.4 节介绍的扩频系统中,采用 PN 序列调制 PSK 信号,从而使发射带宽瞬时扩展,这可由系统的处理增益确定。处理增益是 PN 序列周期的函数。通过使用具有很窄的码片持续时间的 PN 序列,可以获得很大的处理增益,但这同时也伴随着更大的传输带宽和每比特更多的码片数。然而,用来产生 PN 扩频信号的物理设备性能限制了实际可获得的处理增益。实际上,这样得到的处理增益可能仍不足以克服我们关注的一些干扰的影响,因此不得不寻求其他方法。一种替代方法就是使数据调制的载波随机跳频,从而使干扰信号覆盖更大的频率范围。此时,发射信号的频谱扩展是有序的,而不是瞬时的。术语“有序的”是指跳频点的 PN 序列。

这类载波随机跳频的扩频方式被称为跳频(FH)扩频。跳频系统的一种常见调制方式就是 M 进制频移键控(MFSK)。这两种技术的结合简称为 FH/MFSK。(对 M 进制 FSK 的介绍见第 6 章。)

由于跳频并不是瞬时覆盖整个频谱范围的,因此需要研究跳频发生的速率。为此,定义两种基本的(技术独立的)跳频:

- 慢跳频,此时 MFSK 信号的符号速率 R_s 为跳频速率 R_h 的整数倍。也就是说,在每个跳频点上都发送多个符号。

- 快跳频,此时跳频速率 R_h 为 MFSK 符号速率 R_s 的整数倍。也就是说,在发送一个符号

期间，载波频率将改变或跳转多次。

显然，慢跳频和快跳频是相反的，下面将对两者分别进行讨论。

7.7.1 慢跳频

图 7.10(a)为 FH/MFSK 发射机的框图，该发射机由频率调制器和混频器连接而成。首先，输入二进制序列被送入一个 M 进制 FSK 调制器，得到的调制波形和一个数字频率合成器的输出被送入乘法器和带通滤波器级联而成的混频器。滤波器选择通过乘法器得到的和频率分量作为发射信号。特别地，PN 序列的连续 k 比特驱动了频率合成器，从而使载波频率可以在 2^k 个离散值上跳变。在一个跳频点上，发射信号的带宽与采用 $M = 2^k$ 个正交信号的传统 MFSK 相同。然而，对于 2^k 个跳频点来说，发送的 FH/MFSK 信号就占据了较宽的带宽。实际上，采用该技术，可以实现几千兆赫的跳频带宽，这比直接序列扩频可以实现的带宽大了几个数量级。这一较大的跳频带宽还意味着，相干检测只能用于每个跳频点，因为频率合成器无法维持连续跳频点之间的相位连续性。因此，绝大多数的跳频扩频通信系统使用非相干 M 进制调制方式。

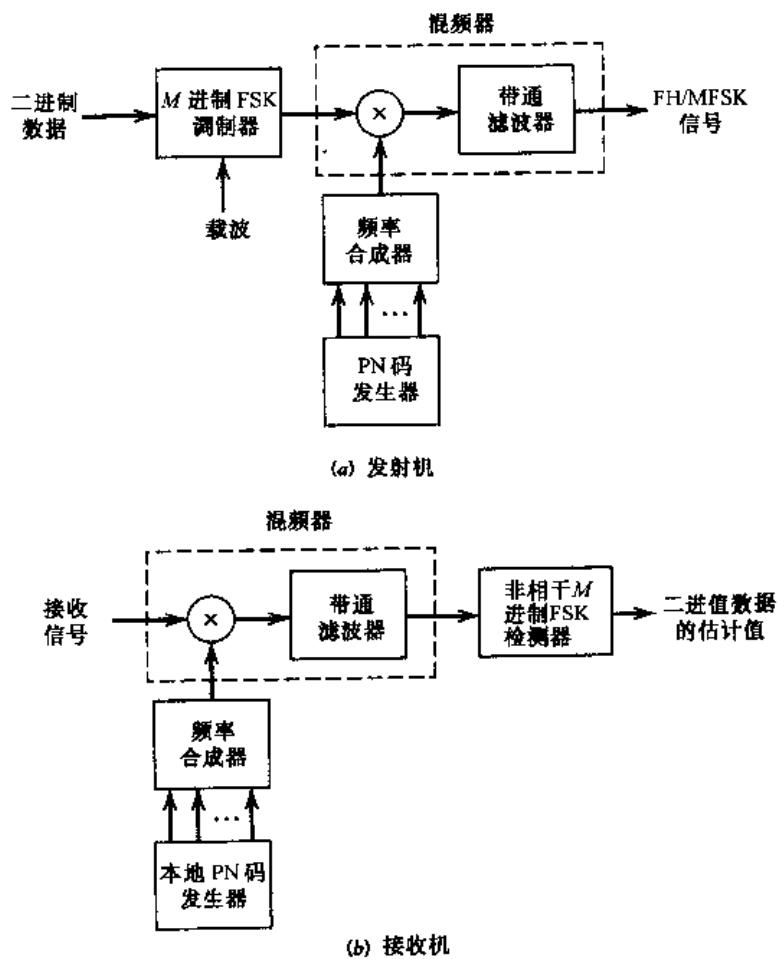


图 7.10 跳频 M 进制 FSK

在图 7.10(b)所示的接收机中，首先将接收信号和本地频率合成器产生的、与发射机中的、

跳变频率同步变化的频率进行混频(下行转换)。将输出进行带通滤波,然后由一个非相干 M 进制 FSK 检测器进行处理。为了实现这一 M 进制检测器,可采用一组 M 个非相干匹配滤波器,其中,每一个都和 MFSK 频率之一相匹配,非相干匹配滤波器在第 6 章进行了介绍。通过选择滤波器输出的最大值得到初始发送符号的估计值。

具有最短的持续时间的一个 FH/MFSK 频率称为码片。注意:不要将这一术语和 7.4 节中介绍 DS/BPSK 的术语相混淆。FH/MFSK 系统的码片速率 R_c 定义为

$$R_c = \max(R_b, R_s) \quad (7.48)$$

其中, R_b 为跳频速率, R_s 为符号速率。

慢 FH/MFSK 的特点是在每个跳频点发送多个符号。因此,慢 FH/MFSK 信号的每个符号都是一个码片。因此,在慢 FH/MFSK 系统中,输入二进制数据的比特速率 R_b , MFSK 信号的符号速率 R_s , 码片速率 R_c 和跳频速率 R_b 之间的关系为

$$R_c = R_s = \frac{R_b}{K} \geq R_b \quad (7.49)$$

其中, $K = \log_2 M$ 。

在每个跳频点,MFSK 的频率按照码片速率 $R_c = R_s$ 的整数倍进行划分,并确保其正交性。这一条件意味着任意发送符号都不会对构成图 7.10(b)所示接收机中 MFSK 检测器的另外 $M - 1$ 个非相干匹配滤波器造成串扰。这里的“串扰”指一个滤波器的输出溢出到邻近的滤波器。得到的慢 FH/MFSK 系统的性能和传统的(非跳频)MFSK 信号在加性高斯白噪声的情况下进行非相干检测的性能是一样的。因此,干扰信号对 FH/MFSK 接收机是有影响的,就平均符号差错概率来看,这个影响就相当于无干扰时加性高斯白噪声对传统的非相干 M 进制 FSK 接收机的影响。根据这一等价性,可将式(6.140)作为 FH/MFSK 系统符号差错概率的近似公式。

假设干扰者决定将平均功率 J 分散到整个跳频范围内,则干扰者的影响就相当于一个功率谱密度为 $N_0/2$ 的加性高斯白噪声,其中, $N_0 = J/W_c$, W_c 为跳频带宽。因此,扩频系统可用符号能量与噪声功率谱密度的比值来描述:

$$\frac{E}{N_0} = \frac{P/J}{W_c/R_s} \quad (7.50)$$

其中,比例式 P/J 是干扰容限的倒数。式(7.50)的分母即为慢 FH/MFSK 系统的处理增益,其定义为

$$\begin{aligned} PG &= \frac{W_c}{R_s} \\ &= 2^k \end{aligned} \quad (7.51)$$

也就是说,处理增益(用 dB 表示)为 $10 \log_{10} 2^k \approx 3k$, 其中, k 为用来选择跳频点的 PN 段的长度。

这个结果的前提是干扰者在整个跳频频谱范围内分配自己的功率。但是,如果干扰者决定将功率集中在几个跳频点上,接收机可实现的处理增益将小于 $3k$ dB。

例 7.4

图 7.11(a)显示了一个完整的 PN 序列周期里慢 FH/MFSK 信号频率的变化情况。PN 序列的周期为 $2^4 - 1 = 15$ 。FH/MFSK 信号具有如下参数:

每个 MFSK 符号的比特数	$K = 2$
MFSK 频率数	$M = 2^K = 4$
每个跳频点 PN 段的长度	$k = 3$
全部跳频点数	$2^k = 8$

本例中，在发送了 2 个符号，或者说 4 个信息比特后，载波频率就跳变到新的跳频点。图 7.11(a)还显示了输入二进制数据，以及用来控制 FH 载波频率选择的 PN 序列。值得注意的是，尽管有 8 个离散的可用跳频点，却只有 3 个被 PN 序列选用。

图 7.11(b)显示了随时间变化的解跳频率的变化情况。可见，这个变化与该输入数据得到的传统 MFSK 信号的变化是一样的。

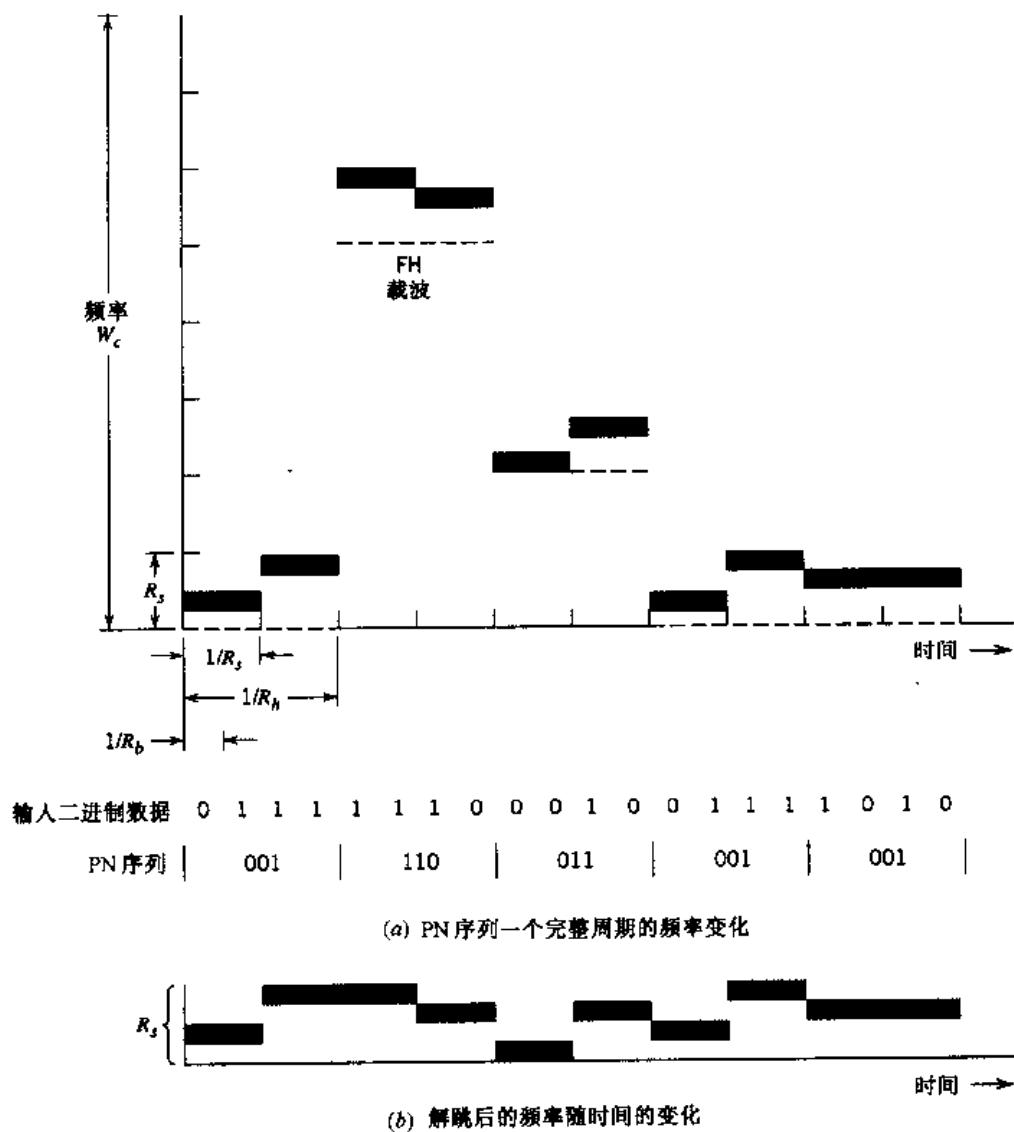


图 7.11 慢跳频示意图

7.7.2 快跳频

快 FH/MFSK 系统和慢 FH/MFSK 系统的差别在于,前者的每个 M 进制符号都对应着多个跳频点。因此,在快 FH/MFSK 系统中,每个跳频点都是一个码片。一般来说,快跳频用来对付具有以下功能的智能干扰策略:测量发射信号的频谱区域,然后在该通带送入干扰信号。显然,为了克服干扰,发射信号必须在干扰信号完成上述两步处理之前跳频到新的载波频率。

接收机里的数据恢复中采用的是非相干检测。但是,检测过程和慢 FH/MFSK 接收机中用到的检测过程是截然不同的。特别地,需要考虑两个步骤:

1. 对于每个 FH/MFSK 符号,分别对接收到的 K 个跳频码片进行判决,并使用简单的择多判决来对解跳 MFSK 信号做出估计。
2. 对于每个 FH/MFSK 符号,计算出接收到的 K 个码片的全部信号的似然函数,并选择其中的最大值。

基于第二个步骤的接收机是最优的,因为在给定的 E_b/N_0 下,对应于平均符号差错概率最小。

例 7.5

图 7.12(a)显示了快 FH/MFSK 信号的发射频率随时间的变化情况。信号具有下列参数:

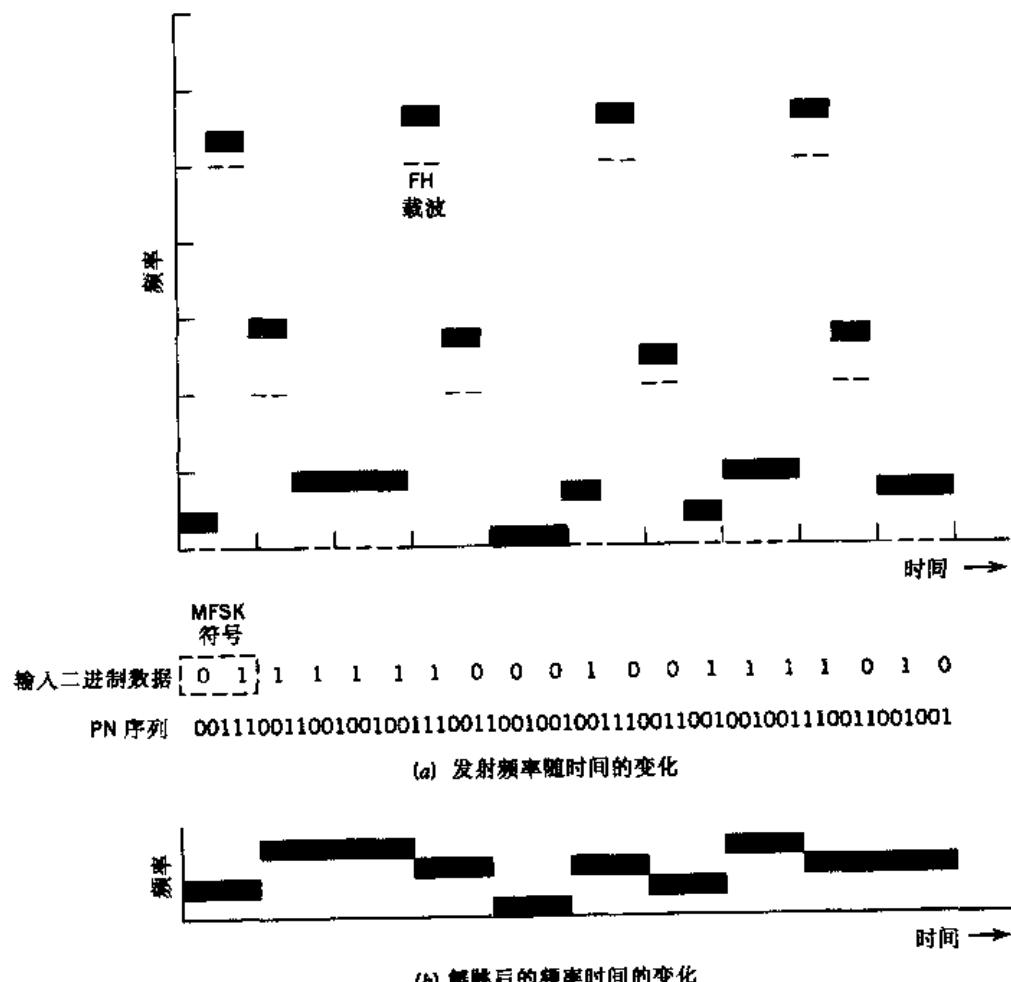


图 7.12 快跳频示意图

每个 MFSK 符号的比特数	$K = 2$
MFSK 频率数	$M = 2K = 4$
每个跳频点 PN 序列段的长度	$k = 3$
全部跳频点数	$2^k = 8$

在这个例子里, 每个 MFSK 信号具有相同的比特数和码片数。也就是说, 码片速率 R_c 和比特速率 R_b 是相等的。在每个码片之后, 发送的 MFSK 信号的载波频率就跳变到不同的值, 除了 PN 序列的 k 码片段重复的少数情况以外。

图 7.12(b) 显示了解跳的 MFSK 信号的频率随时间的变化情况, 这与例 7.4 的结果是一致的。

7.8 计算机实验: 最大长度码和 Gold 码

码分复用(CDM)是传统的频分复用(FDM)和时分复用(TDM)的一个替代。码分复用不需要频分复用那么大的带宽(第 2 章进行了介绍), 也不需要时分复用所需的同步(第 3 章进行了介绍)。实际上, 在扩频调制的基础上, 通过给每个独立用户分配“扩展码”, 可使得公共信道的用户可以访问信道。本实验的目的就是研究能提供令人满意性能的 CDM 系统的一类扩展码。

在理想 CDM 系统中, 系统的任意两个用户之间的互相关函数是 0。为了实现这一理想条件, 要求系统任意两个用户的扩展码之间的互相关函数(在所有循环移位时)都等于 0。遗憾的是, 普通的 PN 序列由于相对较差的互相关性而不能满足这一要求。

为了弥补普通 PN 序列的上述缺陷, 可采用称为 Gold 序列(码)^[5]的一类特殊 PN 序列, 其实现基于下述理论:

用 $g_1(X)$ 和 $g_2(X)$ 表示期望的一对本原 n 次多项式, 与之对应的移位寄存器产生的 m 序列周期为 $2^n - 1$, 与之对应的互相关函数的幅度小于或等于

$$2^{(n+1)/2} + 1 \quad n \text{ 为奇数} \quad (7.52)$$

或

$$2^{(n+2)/2} + 1 \quad n \text{ 为偶数且 } n \neq 0 \pmod{4} \quad (7.53)$$

则对应于乘积多项式 $g_1(X)g_2(X)$ 的移位寄存器将产生 $2^n + 1$ 个不同的序列, 每个序列的周期都是 $2^n - 1$, 而任意两个序列之间的互相关函数都满足上述条件。

本书将该理论称为 Gold 理论。

为了理解 Gold 理论, 需要先定义本原多项式。考虑一个定义在二进制区域(例如遵循二进制算法的一个有限二元组 0 和 1)上的多项式 $g(X)$ 。如果多项式 $g(X)$ 不能被该二进制区域的任何多项式分解因式, 则该多项式就是本原多项式。如果一个 m 次的既约多项式可整除 $X^n + 1$ 且 m 是满足 $n = 2^m - 1$ 的最小整数, 则该多项式就称为本原多项式。更进一步的讨论见第 8 章, 特别是例 8.3。

7.8.1 实验 1: PN 序列的相关性

考虑一对能产生周期为 $2^7 - 1 = 127$ 的 PN 序列的移位寄存器。其中一个反馈移位寄存器

的反馈抽头为[7,1],而另一个的反馈抽头为[7,6,5,4]。两个序列都具有图7.13(a)所示的自相关函数,这可由式(7.5)得到。

但是,PN序列之间互相关函数的计算是一个更为复杂的问题,特别是在 n 很大时。为了进行计算,对时间间隔 $0 < \tau \leq 2^n - 1$ 内的不同循环移位 τ 进行计算机仿真。计算结果如图7.13(b)所示。此图说明:与各自的自相关函数相比,PN序列之间的互相关函数具有较差的性能。互相关函数的幅度超过了40。

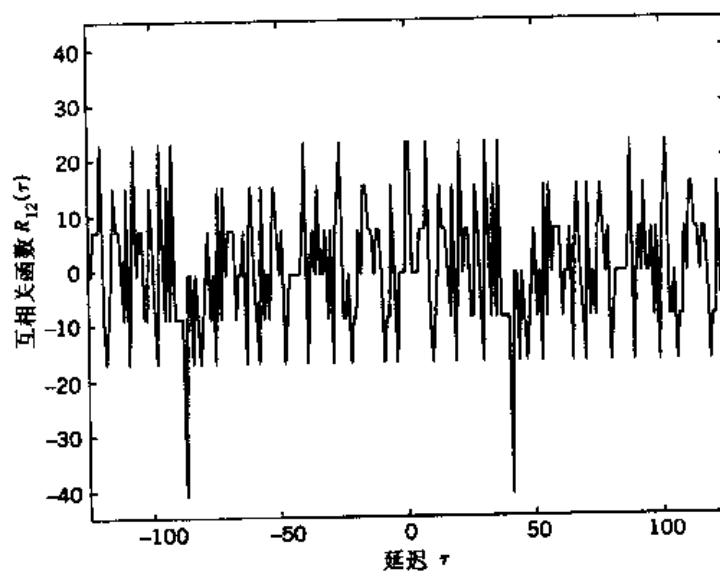
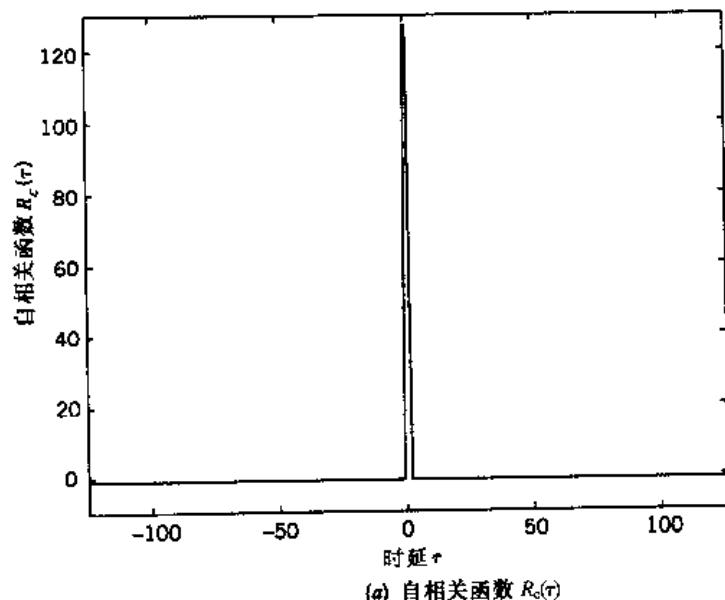


图7.13 PN序列的自相关函数和互相关函数

7.8.2 实验2:Gold序列的相关特性

在接下来的实验里,考虑周期为 $2^7 - 1 = 127$ 的Gold序列。为了在 $n = 7$ 时产生这样的序

列,需要用到满足式(7.52)(n 为奇数)的一对PN序列,即

$$2^{(n+1)/2} + 1 = 2^4 + 1 = 17$$

反馈抽头为[7,4]和[7,6,5,4]的PN序列都可以满足该要求。Gold序列生成器如图7.14所示,图中还包括对这两个序列相加的模2加法器。根据Gold理论,共有

$$2^n + 1 = 2^7 + 1 = 129$$

个序列可以满足式(7.52)。任意两个这样的序列之间的互相关函数如图7.15所示,它和Gold理论是完全一致的。特别地,相关函数的幅度小于或等于17。

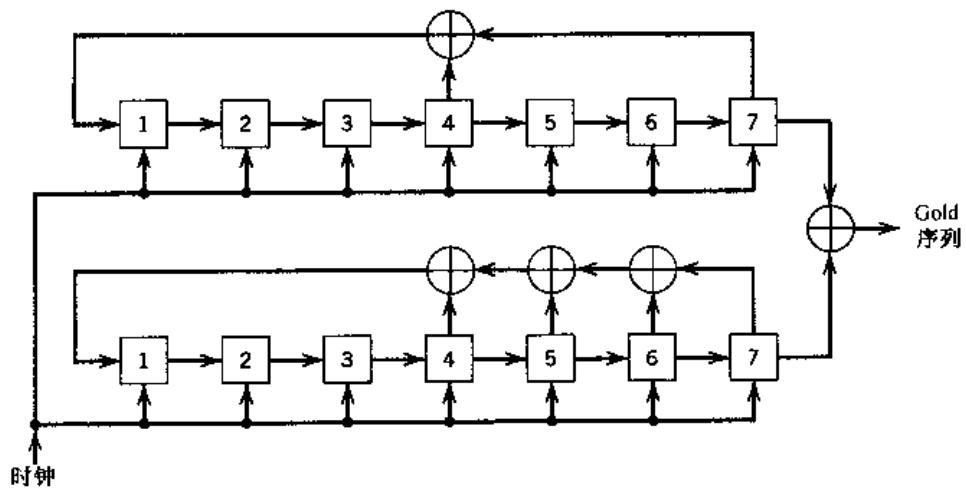


图 7.14 周期为 $2^7 - 1 = 127$ 的 Gold 序列发生器

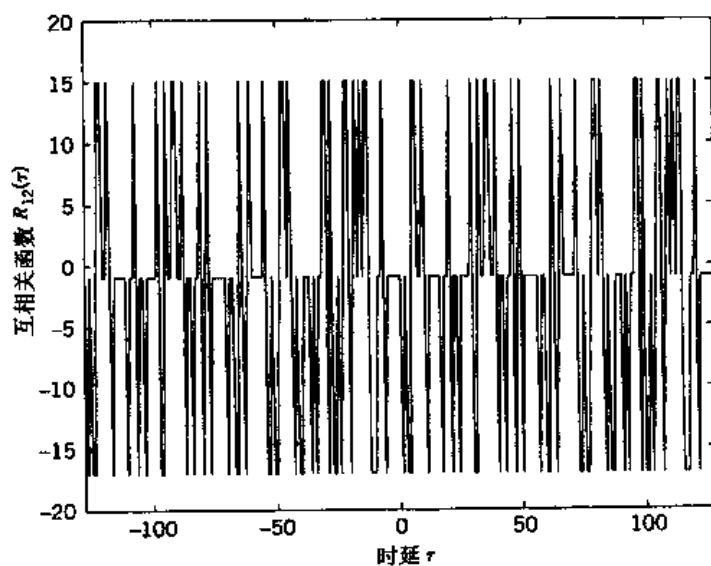


图 7.15 分别基于 PN 序列 [7,4] 与 [7,6,5,4] 的一对 Gold 序列的互相关函数 $R_{12}(\tau)$

7.9 总结与讨论

DS/MPSK 和 FH/MFSK 代表了扩频通信的两个基本类别。它们都要求采用 PN 序列,但是,该序列在两种扩频调制中的应用方式是不同的。

在 DS/MPSK 系统中,PN 序列将发射信号的频谱瞬时扩展到较宽的通带上,从而使发射信号呈现出类似噪声的特性。可采用单载波的二进制 PSK($M = 2$)。一个可替代的方案是采用 QPSK($M = 4$),其中用一对正交载波来发送数据,PSK 和 QPSK 都在 6.3 节进行了介绍。使用 QPSK 的目的通常是提供较高的带宽效率。但在扩频系统中,带宽效率不是关注的重点。此时,采用 QPSK 的目的则是由于它对一些干扰比较不敏感。

在 FH/MFSK 系统中,PN 序列使载频以伪随机的方式在一组频率点上跳变,从而使发射信号的频谱顺序地扩展。

当然,直接序列和跳频扩频技术也可用于同一个系统,这样的系统称为混合 DS/FH 扩频系统。采用混合方法的好处在于:在同一个系统中,可同时实现直接序列和跳频扩频技术的优点。

如果不考虑干扰波形,那么,对扩频通信的讨论就是不完整的。实际应用中遇到的干扰包括下列类型:

1. **阻塞噪声干扰**,由具有高平均功率的带限高斯白噪声组成。阻塞噪声干扰是一种强有力干扰,且除了频谱带宽它不需要知道抗干扰通信系统的任何信息。
2. **局部噪声干扰**,当全部功率平均分配在整个扩展频谱时,由某些通带中的噪声组成。由于通常的带宽较小,局部噪声干扰比阻塞噪声干扰更容易产生。
3. **突发噪声干扰**,其中包括具有如下部分发射功率:

$$J_{\text{peak}} = \frac{J}{p}$$

的宽带噪声,式中 p 表示该干扰存在时间占总时间的比例。因此,在占总时间比例为 $1 - p$ 的那部分时间里,该干扰的发射功率为 0。突发噪声干扰的平均噪声功率等于 J 。

4. **单频干扰**,由其频率位于扩展带宽内的正弦信号组成,是最容易产生的干扰信号。
5. **多频干扰**,是局部噪声干扰的频率等价形式。

除了这 5 种干扰以外,实际的通信环境中还存在其他多种干扰波形。但总的来说,没有一种干扰波形能够对所有的扩频系统都造成最严重的干扰,也没有一种扩频系统对所有可能的干扰波形都是最佳的。

注释与参考

- [1] 引言中关于扩频调制的定义选自 Pickholtz, Schilling 以及 Milstein(1982)。该论文对扩频通信理论进行了回顾性的研究。
- 介绍该主题的论文有 Viterbi(1979), Cook 和 Marsh(1983)。关于该主题的著作,有 Dixon(1984), Holmes(1982) Ziemer 和 Peterson(1985, pp. 327 – 649), Cooper 和 McGillem (1986, pp. 269 – 411), 以及 Simon, Omura, Scholtz 和 Levitt(1985, Volumes I, II, III)。

Simon 等人所著的共三卷的著作是目前研究扩频通信的最详尽的公开资料。扩频通信的发展可以回顾到 20 世纪 50 年代中期。对该技术发展史的介绍, 参见 Scholtz (1982)。该论文追溯了 20 世纪 20 年代扩频通信的起源。该论文中关于发展史的大部分内容见 Simon 等人的著作卷 I 第 2 章。

Tantaratana 和 Ahmed 的书中(1998)包括了扩频通信无线应用的介绍性论文和高级论文。这些论文被分为以下类别: 扩频技术、蜂窝移动系统、卫星通信、无线局域网和全球定位系统(GPS)。

- [2] 对 m 序列更详细的介绍, 参见 Golomb(1964, pp. 1 – 32), Simon, Omura, Scholtz 和 Levitt (1985, pp. 283 – 295), 以及 Peterson 和 Weldon(1972)。最新的参考资料如产生 m 序列的多项式的详细列表, 可参见 Dixon(1984)。对 PN 序列的研究论文, 参见 Sarwate 和 Pursley(1980)。
- [3] 表 7.1 摘自 Dixon 的著作(1984, pp. 81 – 83)。该著作列出了移位寄存器长度 $m \leq 89$ 的所有 m 序列的反馈抽头。
- [4] 对扩频通信中同步问题的详细讨论参见 Ziemer 和 Peterson(1985, Chapter 9-10)以及 Simon 等人的著作(1985, Volume III)。
- [5] 关于 Gold 序列的最初的论文是 Gold(1967, 1968)。而 Holmes(1982)对 Gold 序列进行了详细介绍。

习题

PN 序列

- 7.1 一个 PN 序列由长度 $m = 4$ 的反馈移位寄存器产生。码片速率为 10^7 码片/秒。试确定以下参数:
 - (a) PN 序列长度
 - (b) PN 序列的码片持续时间
 - (c) PN 序列的周期
- 7.2 图 P7.2 为一个四级反馈移位寄存器。寄存器的初始状态为 1000。试确定移位寄存器的输出序列。

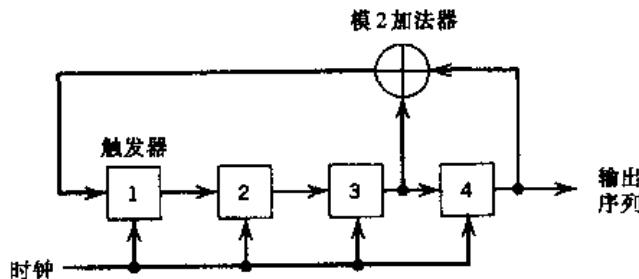


图 P7.2

- 7.3 对于习题 7.2 所示的反馈移位寄存器, 证明 PN 序列的平衡性和游程性质。同时, 计算并画出该移位寄存器产生的 PN 序列的自相关函数。

- 7.4 参考表 7.1,推导出周期 $N = 63$,反馈抽头分别为 $[6,1]$, $[6,5,2,1]$ 和 $[6,5,3,2]$ 的三级反馈移位寄存器的 m 序列。
- 7.5 图 P7.5 为图 7.4(b)所示的线性反馈移位寄存器的多抽头形式。证明该方式产生的 PN 序列与表 7.2(b)中的结果完全一致。

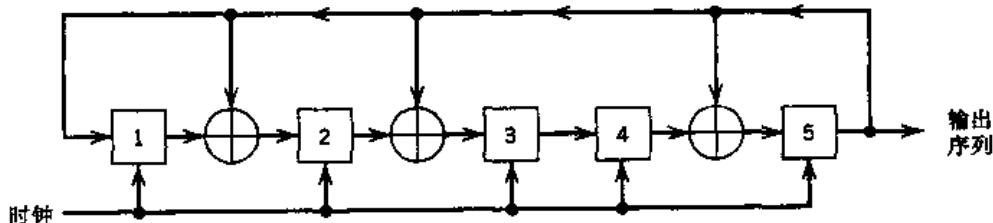


图 P7.5

DS/PSK 系统

- 7.6 证明表 7.3 中的真值表可通过如下两个步骤得到:
- 信息信号 $b(t)$ 与 PN 信号 $c(t)$ 进行模 2 加。
 - 模 2 加法器输出端的符号 0 和 1 分别移相 0° 和 180° 表示。
- 7.7 将一个单频干扰信号
- $$j(t) = \sqrt{2}J \cos(2\pi f_c t + \theta)$$
- 送入 DS/BPSK 系统。 N 维发射信号 $x(t)$ 如式(7.16)所示。确定干扰信号 $j(t)$ 的 $2N$ 个坐标。
- 7.8 扩频系统的处理增益可表示为发射信号的扩展带宽和接收信号的解扩带宽的比值。对于 DS/BPSK 系统,证明上述结论。
- 7.9 一个 DS/BPSK 系统采用长度为 19 的反馈移位寄存器来产生 PN 序列。计算该系统的处理增益。
- 7.10 在 DS/BPSK 系统中,用来产生 PN 序列的反馈移位寄存器的长度 $m = 19$ 。要求由于外部产生的干扰信号引起的系统平均符号差错概率不超过 10^{-5} 。计算下列系统参数并用 dB 表示:
- 处理增益
 - 抗干扰容限
- 7.11 7.5 节中对采用二进制 PSK 的一个直接序列扩频系统的信号空间维度和处理增益进行了分析。对采用 QPSK 的系统重复上述分析。

跳频扩频

- 7.12 慢 FH/MFSK 系统具有如下参数:
- 每个 MFSK 符号的比特数 = 4
- 每个跳频点的 MFSK 符号数 = 5
- 计算该系统的处理增益。
- 7.13 快跳频 FH/MFSK 系统具有如下参数:
- 每个 MFSK 符号的比特数 = 4

每个 MFSK 符号的跳频点数 = 4

计算该系统的处理增益。

计算机实验

7.14 考虑周期 $N = 63$ 的两个 PN 序列。一个序列的反馈抽头为 $[6, 1]$ ，另一个序列的反馈抽头为 $[6, 5, 2, 1]$ （选自表 7.1）。

(a) 计算这两个序列的自相关函数和互相关函数。

(b) 比较(a)中计算所得的互相关函数和序列 $[6, 5, 2, 1]$ 与镜像序列 $[6, 5, 4, 1]$ 之间的互相关函数。

对结果做出总结。

7.15 (a) 计算反馈抽头为 $[5, 2]$ 的 PN 序列与由反馈抽头为 $[5, 3]$ 定义的镜像序列之间的局部互相关函数。

(b) 对反馈抽头为 $[5, 2]$ 的 PN 序列和反馈抽头为 $[5, 4, 2, 1]$ 的 PN 序列重复上述计算。

(c) 对反馈抽头为 $[5, 4, 3, 2]$ 的 PN 序列和反馈抽头为 $[5, 4, 2, 1]$ 的 PN 序列重复上述计算。

反馈抽头 $[5, 2]$, $[5, 4, 3, 2]$ 和 $[5, 4, 2, 1]$ 是根据表 7.1 得到的周期为 31 的 m 序列的可能抽头值。

第8章 多用户无线通信

顾名思义,所谓多用户通信就是一个通信信道被多个用户同时使用。本章所讨论的通信系统是通过无线电传输来连接接收机和发射机的。

本章将集中讨论以下问题:

- 多址技术,这是多用户通信系统的基础。
- 卫星通信,提供全球覆盖。
- 无线链路分析,强调了发射和接收天线的任务以及自由空间传播等问题。
- 无线通信,侧重于移动性和多径现象的研究。
- 无线通信的语音编码。
- 无线通信的自适应天线。

8.1 引言

前面章节中所介绍的通信理论,大多数内容都是基于对通信信道特定的理想化模型来进行的,即带宽受限和受加性高斯白噪声干扰的信道模型。如此发展起来的经典通信理论在数学方面相当成熟,对通信系统不断扩展的领域提供了坚实的基础。卫星通信信道就是由这种模型描述的一种物理信道,所以,本章首先讨论的多用户通信类型即为卫星通信。

在对地同步轨道中的卫星通信系统中,依靠可视无线传播来处理从一个地面终端至转发器的上行通信,以及从转发器至另一个地面终端的下行通信。因此,对卫星通信的讨论将引出对自由空间中无线电传播的分析,因为后者连接了接收天线和发射天线。

卫星通信的应用提供了全球范围的覆盖。本章研究的另一个多用户通信系统,即无线通信系统,则提供了移动性。无线通信系统与现有的电话网络和卫星通信系统一起,使得一个移动单元能够与世界上任何地点的任何人进行通信。无线通信系统的另一个特性就是不受限(即允许完全的自由定位)。因此可考虑将其应用于局域网(即数据网被局限在几公里内的大楼中)。这样做的原因在于,与传统的电缆相比,无线通信系统具有明显的优势:不需要布线或重复布线,可灵活地创建新的通信服务器以及用户的可移动性。

由于多径的存在,无线通信的无线电传播信道已不再是理想的 AWGN 信道。多径是信号相关现象的一种非高斯形式,其出现原因是由于传输信号在固定或移动物体上的反射。多径的存在带来了无线电传播信道实际应用的困难和数学分析的复杂性。简单地说,多径是无线通信的户内和户外工作所固有的物理现象。

在继续讨论卫星通信和无线通信的特性之前,我们先介绍多址技术。多址技术使不同的用户同时(或几乎同时)地接入到一个共同的信道。

8.2 多址技术

凭借多址技术,多个用户或本地站可以同时或几乎同时地共享通信信道,尽管各自的传输信号来源于不同的位置。换言之,多址技术使得信道的通信资源被寻求互相通信的大量用户所共享。

多址技术和复用技术有几点细微差别需要注意:

- 多址技术是指通信信道(例如卫星或无线电信道)被分布在不同位置上的用户远程共享,而复用技术是指信道(例如电话信道)被局部范围内的用户所共享。
- 在复用系统中,用户的要求通常是固定的。相反地,在多址系统中,用户的要求会随时间而动态变化。在这种情况下,就必须提供动态的信道分配。

正因如此,多址系统中信道资源的共享应当在不造成系统用户之间严重相互干扰的情况下进行。多址可分为四种基本类型:

1. 频分多址(FDMA)

在 FDMA 中,不相交的子频带在连续时间上分配给不同的用户。为了减小处于相邻信道频带的用户之间的干扰,采用保护频带作为缓冲区间,如图 8.1(a)所示。这些保护频带是必须的,因为不可能实现理想滤波以分离不同的用户。

2. 时分多址(TDMA)

在 TDMA 中,每个用户都占用信道的整个频带,但只占据给被称为时隙的很短的持续时间。如图 8.1(b)所示,采用保护时间作为缓冲区间插入到所分配的时隙之间。这是为了减小由于系统非理想而产生的时间不确定性所造成的用户间的干扰,尤其是在同步机制下。

3. 码分多址(CDMA)

在 FDMA 中,信道资源的共享是通过将其沿频率坐标分割为不同频带而实现的,如图 8.1(a)所示。而在 TDMA 中,信道资源的共享则是通过将其沿时间坐标分割为不同时隙来实现,如图 8.1(b)所示。图 8.1(c)介绍了另一种共享信道资源的技术,其中结合了 FDMA 和 TDMA,这种技术称为 CDMA。如果在 CDMA 中采用跳频机制,就能够确保在每个连续的时隙内,分配给用户的频带以一种本原随机的方式重新排序。具体表述为,在时隙 1,用户 1 占用频带 1,用户 2 占用频带 2,用户 3 占用频带 3,以此类推。在时隙 2,用户 1 跳至频带 3,用户 2 跳至频带 1,用户 3 跳至频带 2,以此类推。这种机制类似于玩音乐座椅的游戏。CDMA 与 FDMA 和 TDMA 相比,有一个重要优势,就是能够提供安全通信。在如图 8.1(c)所示的 CDMA 类型中,跳频机制可通过采用 PN 序列来实现。

4. 空分多址(SDMA)

在 SDMA 中,资源分配是通过各用户的空间分离实现的。特别地,将多波束天线指向不同的方向来分离无线电信号。所以,不同用户能够在相同的频率或相同时隙下同时访问信道。

这些多址技术具有一个共同的特点:采用时间、频率或空间的分离技术(或一般概念上的

正交性)来分配信道的通信资源。

有了这些背景知识,接下来就可以讨论一些重要的多用户通信系统了。

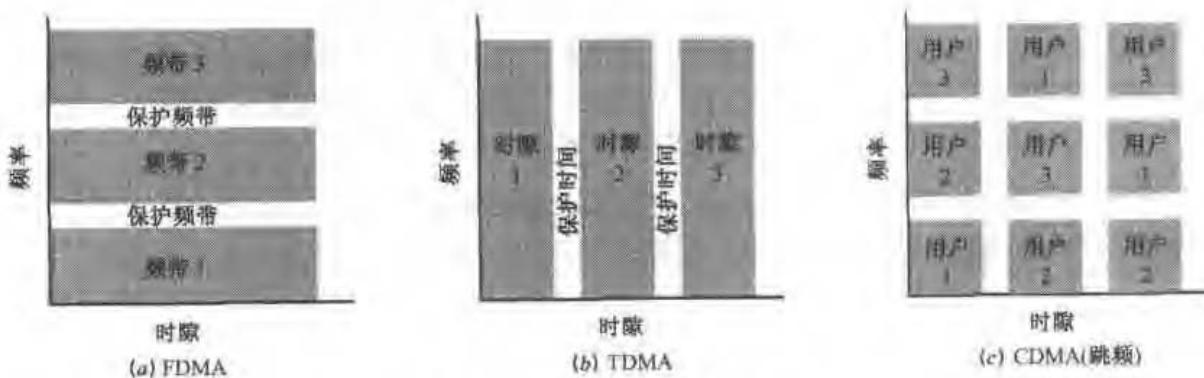


图 8.1 多址技术的基本思想

8.3 卫星通信

在静地卫星通信系统中,^[1]消息信号通过上行链路从地面站传送至卫星,经卫星上的转发器(电子电路)进行放大,然后通过下行链路从卫星传送至另一个地面站,如图 8.2 所示。卫星通信的最普遍的频带是上行链路为 6 GHz(C 频带),下行链路为 4 GHz。这种频带分配方式具有如下优势:

- 相对便宜的微波设备。
- 降水引起的衰落很小。降水是信号衰落的一个主要大气因素。
- 可忽略的天空背景噪声。天空背景噪声(由从银河系、太阳和陆地发出的随机噪声产生)在 1~10 GHz 之间时达到最低水平。

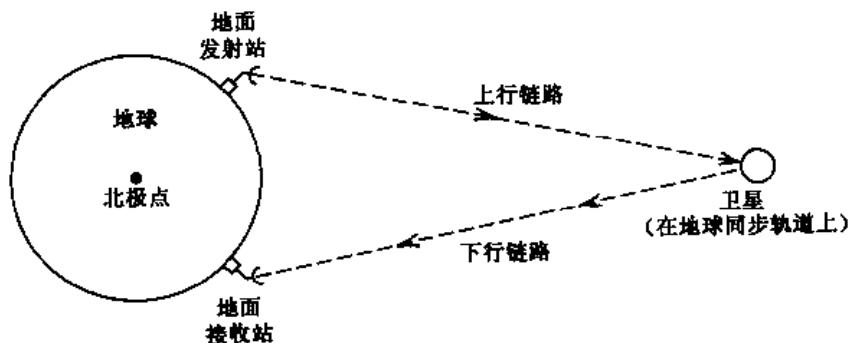


图 8.2 卫星通信系统

但是,无线电干扰限制了通信卫星在 6/4 GHz 的频带中的应用,因为这个频带的传输频率与陆地微波系统是相同的。在工作于 14/12 GHz 频带(即 Ku 频带)的更强大的“第二代”通信卫星中,这个问题得到了解决。同时更高频率的使用使得构造更小(从而更便宜)的天线成为可能。

图 8.3 所示的框图为一个典型通信卫星的单个转发器信道的基本组成部分。特别地,上行链路的接收天线输出被送至下列部分的级联中:

- 带通滤波器,用于从不同无线电信道中分离出接收信号。
- 低噪声放大器。
- 下变频器,将接收到的射频信号转换为所需的下行链路频率。
- 行波管(TWT)放大器,提供宽频带的高增益。在行波管中,电磁信号在螺旋(即弹簧形状的电线卷)中传播,而高压电波束中的电子以接近信号波的速度通过螺旋。结果是功率由电子传递到波,信号波在螺旋中的传输过程中迅速增强。

图 8.3 所示的信道采用了单频率传输。不采用单频率传输的信道则通过两个步骤来进行上行链路到下行链路的频率转换:下变频为中频,接着进行放大,然后上变频为所需的传输频率。

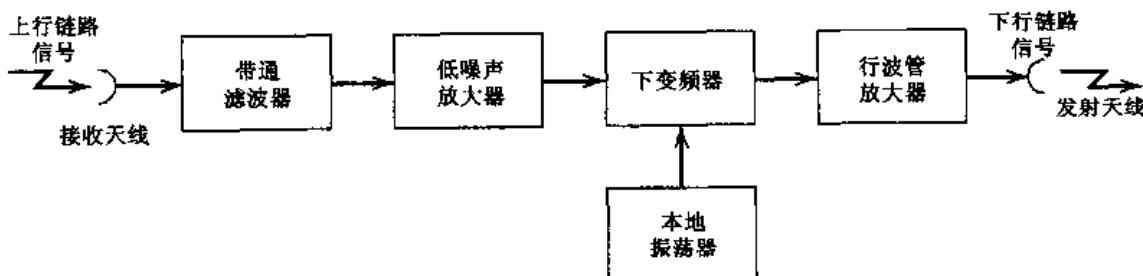


图 8.3 转发器框图

由于长距离的原因,传输时延在卫星信道中变得非常明显。特别地,由卫星传输的语音信号会存在大约 270 ms 的时延。所以,对于语音信号,卫星链路接收端的任何阻抗不匹配都会造成讲话者声音的回音,这个回音在大约 540 ms 的巡回时延后将在发射终端听到。可用回波抵消器来解决这个问题,将回声的估计从返回路径中减掉。回音的消除由特殊的滤波器来完成,该滤波器能够根据信道特性的变化自适应地变化。

卫星信道可由加性高斯白噪声(AWGN)模型来近似表示,该模型可应用于卫星通信系统的上行链路和下行链路部分。因此,第 6 章所介绍的关于通常数据传输系统的大部分内容,特别是 PSK 和 FSK 技术,都可直接应用于数字卫星通信。

卫星转发器与传统微波视距中继器的区别在于,许多地面工作站可以同时或几乎同时地从不同位置访问卫星。这可通过 8.2 节中所讨论的任何一种多址技术来实现。关于这一点,可得到如下的观察结果:

- 在卫星信道中,转发器的非线性是用户间相互干扰的主要原因。为控制这个严重的问题,转发器中的行波管放大器工作在其限度以下。因此,在 FDMA 系统中,系统效率将减小,这是由于在行波管放大器中存在不可避免的功率衰落。
- 在 TDMA 系统中,用户逐个访问卫星转发器。因此,只要允许行波管放大器运行在饱和状态,卫星转发器就能够以接近全功率的方式工作。这意味着 TDMA 比 FDMA 更有效地利用了转发器,因此 TDMA 在数字卫星通信系统得到广泛的应用。
- SDMA 利用地面站的空间位置进行工作,这是通过机载交换实现的。特别地,转发器装配有多个天线,应选择适当的天线束以便需要使用转发器的地面站能够进行无线电传输。

除了多址,卫星信道的另一个功能是强调大覆盖范围的广播。这里提到的广播卫星,其特性就在于能够向廉价的接收机进行高功率传输。此特性在为实现大范围电视服务的家庭接收

而设计的直接广播卫星(DBS)中得到了利用。与卫星通信的大型地面站相比,DBS 的地面站结构非常简单,因此也相当廉价。

8.4 无线链路分析

卫星通信系统设计的一个重要问题在于其链路预算分析。^[2]顾名思义,链路预算(更准确地说是“链路功率预算”)就是通信链路运行中的所有增益和损耗的总和。特别地,构造链路预算的平衡表必须提供以下三个广义项目的详细说明:

1. 可用资源在发射机和接收机之间的分配。
2. 造成信号功率损耗的原因。
3. 噪声来源。

将所有的这些项目都包括在链路预算中,然后就可以用估算程序来计算无线电链路的性能了。这里所说的链路可能是一个卫星通信系统的上行或下行链路。显然,此节所提出的通信链路分析要素同样适用于其他基于视距工作的无线电链路。正因如此,此节对无线电链路分析的处理方法具有普遍性。下面以一个数字卫星通信系统的下行链路预算分析的例子来结束本节。

由第6章所给出的资料,可得到一个存在加性高斯白噪声的数字通信系统的性能,这可由具有如图8.4所示的“瀑布”曲线的公式来定义。该图形显示了符号差错概率 P_e 对信噪比 E_b/N_0 的变化曲线。选择调制方法后,首先必须明确下列两个 E_b/N_0 的值:

1. 所需的 E_b/N_0

假设预先设定符号差错概率 $P_e = 10^{-3}$ 。采用图8.4所示的瀑布曲线所对应的调制方法,实现预设的 P_e 所需的 $(E_b/N_0)_{req}$ 是确定的。用 $(E_b/N_0)_{req}$ 来表示此次计算得到的 E_b/N_0 的值。预设的 P_e 和计算得到的 $(E_b/N_0)_{req}$ 定义了图8.4的瀑布曲线上的一点,假定为工作点1。

2. 接收到的 E_b/N_0

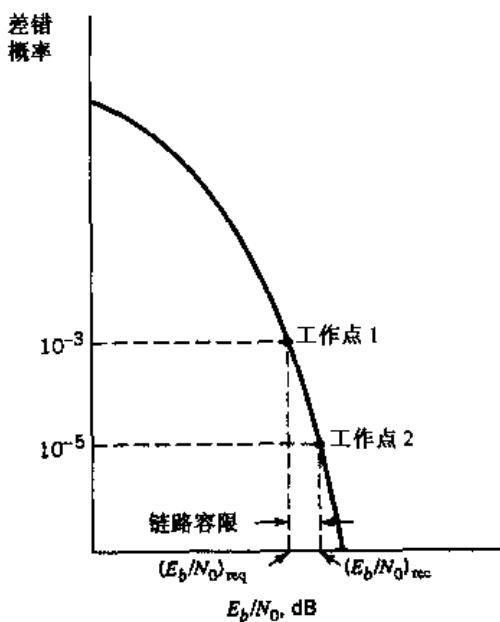
为确保通信链路的可靠工作,链路预算包括一个称为链路容限的安全度量。链路容限提供了对改变和突发状况的保护措施。因此,系统实际接收到的 (E_b/N_0) 要比 $(E_b/N_0)_{req}$ 大些。以 $(E_b/N_0)_{rec}$ 表示实际的或接收到的 E_b/N_0 。 $(E_b/N_0)_{rec}$ 定义了图8.4的瀑布曲线上第二个点,假定为工作点2。图8.4中对应于工作点2的 P_e 的值为 10^{-5} ,仅仅是为了说明的需要,不具有任何意义。将链路容限用 M 表示,则

$$\left(\frac{E_b}{N_0}\right)_{rec} = M \left(\frac{E_b}{N_0}\right)_{req} \quad (8.1)$$

同样,将这两个 E_b/N_0 表示为dB的形式,可将链路容限定义为

$$M(\text{dB}) = \left(\frac{E_b}{N_0}\right)_{rec}(\text{dB}) - \left(\frac{E_b}{N_0}\right)_{req}(\text{dB}) \quad (8.2)$$

显然,链路容限 M 越大,通信链路的安全性越高。但是,链路可靠性的增加需要很高的 E_b/N_0 。

图 8.4 差错概率关于 E_b/N_0 的“瀑布”曲线

8.4.1 自由空间传播模型

链路预算的下一个步骤是计算接收到的信号功率。计算中需考虑载波信号的发送和接收过程中所有的增益和损耗。

在无线通信系统中,调制信号的传播是通过发射天线实现的,发射天线的功能包括两部分:

- 将已调电信号转换到一个电磁场中。此时,发射天线扮演的是“阻抗变换器”的角色,从而使天线阻抗与自由空间阻抗相匹配。
- 在所需的方向上辐射电磁功率。

接收端用接收天线来实现与发射天线相反的功能:将电磁场转换为电信号,并从中提取调制信号。此外,可能还需要抑制某些方向上的无用辐射。

接收机通常位于发射天线的远端,在这种情况下,为方便起见,可认为发射天线是虚构的、无体积的发射机或是点源。将此远距离点源作为时间和空间的函数进行详尽描述需要电磁场方面的知识。但是,就链路计算而言,只需要了解天线功率密度的变化就足够了。

根据定义,坡印廷矢量(Poynting vector)或功率密度就是每单位面积上能量流的速率,其单位为 W/m^2 。将发射天线看成点源可大大简化问题,这是因为点源的功率密度只有一个径向分量,即辐射能量从点源沿着径向流出。

发射和接收天线的性能可与一个“参考”天线做比较。通常假设参考天线是一个各向同性源,也就是在所有方向均匀辐射(即完全无方向)的天线。各向同性源是假设的,实际上所有无线电天线都具有一些微小的方向性。尽管如此,各向同性源的概念仍然是非常有用的,尤其是在进行增益比较时。

假设各向同性源辐射的所有功率用 P_t 表示, 单位为 W(瓦)。辐射功率均匀地通过一个表面积为 $4\pi d^2$ 的球体, 其中, d 为到源点的距离, 单位为 m(米)。以 $\rho(d)$ 表示的球体表面任一点的功率密度如下:

$$\rho(d) = \frac{P_t}{4\pi d^2} \text{ W/m}^2 \quad (8.3)$$

式(8.3)说明功率密度与到源点距离的平方成反比例的关系。这个结论是我们所熟悉的反平方律, 它限制了电磁波在自由空间中的传播。

将功率 $\rho(d)$ 与距离 d 的平方相乘, 可得到辐射强度, 以 Φ 表示:

$$\Phi = d^2 \rho(d) \quad (8.4)$$

其中, $\rho(d)$ 的单位为 W/m^2 , 而辐射强度 Φ 的单位为 W/sr (瓦每立体弧度)。

在采用典型发射或接收无线电天线的情况下, 辐射强度由图 8.5 所示的球坐标 θ 和 ϕ 的函数表示。因此, 通常将辐射强度表示为 $\Phi(\theta, \phi)$, 即所谓的辐射强度图样。无限小立体角 $d\Omega$ 的辐射功率可由 $\Phi(\theta, \phi) d\Omega$ 给出, 此处有(参考图 8.5)

$$d\Omega = \sin \theta \, d\theta \, d\phi \text{ sr} \quad (8.5)$$

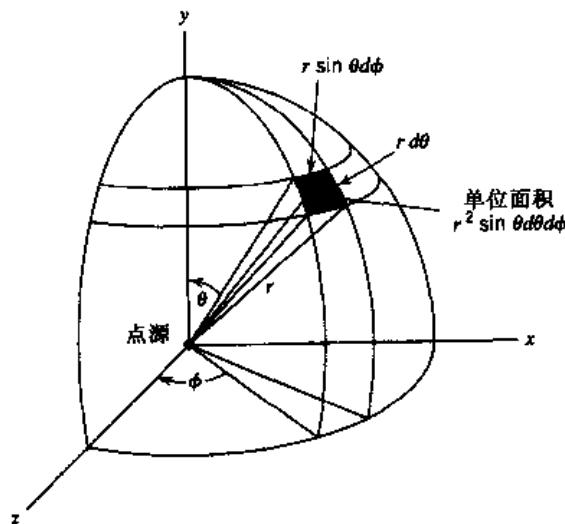


图 8.5 点源的球面坐标

因此, 总辐射功率为

$$P = \int \Phi(\theta, \phi) \, d\Omega \text{ W} \quad (8.6)$$

上式为功率定理理论的数学表达式。功率定理告诉我们: 给定 (θ, ϕ) 的值, 就可得到辐射强度图样 $\Phi(\theta, \phi)$, 总辐射功率可由 $\Phi(\theta, \phi)$ 在 4π 立体弧度的立体角上积分得到。每单位立体角上的平均辐射功率为

$$\begin{aligned} P_{av} &= \frac{1}{4\pi} \int \Phi(\theta, \phi) \, d\Omega \\ &= \frac{P}{4\pi} \text{ W/sr} \end{aligned} \quad (8.7)$$

上式表示由辐射与总功率 P 相同的各向同性源所提供的辐射强度。

8.4.2 方向增益、方向性和功率增益^[3]

发射天线将辐射功率集中在指定方向上的能力,或接收天线从给定方向上有效吸收入射功率的能力,用方向增益或方向性来表示。在由角度对(θ, ϕ)所确定的方向上,天线的方向增益由 $g(\theta, \phi)$ 表示,其定义为此方向上的辐射强度和平均辐射功率的比值,即

$$\begin{aligned} g(\theta, \phi) &= \frac{\Phi(\theta, \phi)}{P_{av}} \\ &= \frac{\Phi(\theta, \phi)}{P/4\pi} \end{aligned} \quad (8.8)$$

天线的方向性由 D 表示,其定义为天线上的最大辐射强度与各向同性源的辐射强度的比值。即方向性 D 是方向增益 $g(\theta, \phi)$ 的最大值。所以,尽管天线的方向增益是角度对(θ, ϕ)的函数,方向性 D 却是一个常量,即某个特定方向的最大值。

方向性的定义是基于辐射强度图样 $\Phi(\theta, \phi)$ 的,其中不包括由耗散损失和阻抗不匹配所造成的天线非理想性产生的影响。功率增益包括了天线的辐射效率。天线的功率增益(由 G 表示),可定义为在相同的输入功率下,天线的最大辐射强度和无损耗的各向同性源的辐射强度的比值。用 $\eta_{radiation}$ 表示天线的辐射效率因数,则功率增益 G 与方向性 D 的关系为

$$G = \eta_{radiation} D \quad (8.9)$$

因此,无损耗的各向同性源上的天线功率增益等于具有 100% 效率(即 $\eta_{radiation} = 1$)的天线的方向性,但如果天线有损耗(即 $\eta_{radiation} < 1$),天线功率增益将小于该方向性。以后的讨论中都假设天线具有 100% 的效率,因此只涉及到天线的功率增益。

根据互易原理,基于发射功率图样的功率增益概念可以扩展至接收天线的情况。如果天线的传输介质是线性、无源且各向同性的,该天线就可认为是互易的。对于一个给定的天线结构,发射和接收天线的功率增益是相等的。

天线的功率增益是将功率密度集中在一个小于 4π 立体弧度的区域中的结果,如图 8.6 所示。由此图可以引入如下的两个参量:

1. 各向同性源的有效辐射功率(EIRP)。EIRP 的定义为发射功率 P_t 与发射天线功率增益 G_t 的乘积,即

$$EIRP = P_t G_t \quad W \quad (8.10)$$

2. 天线波束宽度,这是天线立体角的“平面”度量,单位为度或弧度。波束宽度的定义为场功率图的主瓣上对应于最高场功率衰落 3 dB 的两点的夹角。天线的功率增益越高,此天线波束宽度就越窄。

从图 8.6 可以看出:另一个值得注意的问题就是场功率图的旁瓣。遗憾的是,每个物理天线都有旁瓣,旁瓣的作用是吸收无用的干扰辐射。

8.4.3 有效孔径

天线的有效孔径是接收天线的一个非常重要的概念,其定义为天线终端提供的功率与每

单位面积内经过适当极化的人射电磁波功率的比值。有效孔径 A 可由天线功率增益 G 表达为

$$A = \frac{\lambda^2}{4\pi} G \quad (8.11)$$

其中, λ 为载波波长。波长 λ 和频率 f 成倒数关系:

$$\lambda = \frac{c}{f} \quad (8.12)$$

其中, c 为光速(约为 3×10^8 m/s)。

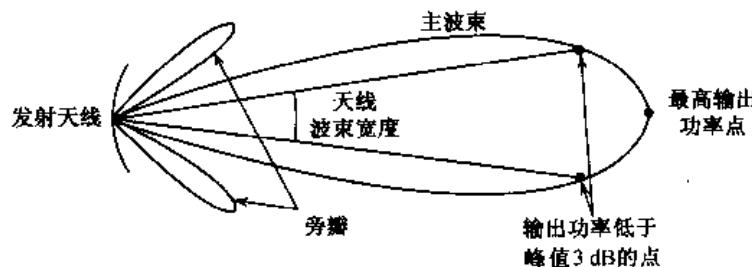


图 8.6 小于 4π 立体角的区域中发射天线的功率密度集中图示

有效孔径的概念对于反射天线和具有明确孔径的喇叭形天线都具有很重要的意义。对于这些天线, 天线的有效孔径与物理孔径的比值是在向目标方向辐射功率或从目标方向上吸收功率时天线孔径效率 η_{aperture} 的直接度量。反射天线效率 η_{aperture} 的理论值位于 45% ~ 75% 之间。

8.4.4 Friis 自由空间公式

有了这些关于天线的知识, 就能够总结出无线电通信链路的基本传播公式。假设发射天线具有式(8.10)所定义的 EIRP, 运用式(8.3)的反平方律, 可将发射天线的功率密度表示为 $\text{EIRP}/(4\pi d^2)$, 其中, d 为接收天线与发射天线之间的距离。接收天线所吸收的功率 P_r 为此功率密度与天线有效面积 A_r 的乘积, 即

$$\begin{aligned} P_r &= \left(\frac{\text{EIRP}}{4\pi d^2} \right) A_r \\ &= \frac{P_t G_t A_r}{4\pi d^2} \quad \text{W} \end{aligned} \quad (8.13)$$

根据互补原理, 可利用式(8.11)将接收天线的有效面积表达为

$$A_r = \frac{\lambda^2}{4\pi} G_r,$$

其中, G_r 为接收天线的功率增益, 将此 A_r 的表达式代入式(8.13), 可将接收信号功率表示为如下的等价形式:

$$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi d} \right)^2 \quad (8.14)$$

式(8.14)称为 Friis 自由空间公式。^[4]

路径损耗 PL 以 dB 的形式表示信号在整个通信链路上的“衰落”，其定义为发射信号功率 P_t 和接收信号功率 P_r 的差值(以 dB 表示)，即

$$\begin{aligned} PL &= 10 \log_{10} \left(\frac{P_t}{P_r} \right) \\ &= -10 \log_{10}(G_t G_r) + 10 \log_{10} \left(\frac{4\pi d}{\lambda} \right)^2 \end{aligned} \quad (8.15)$$

式(8.15)中第一项中的“-”表明此项代表“增益”。第二项由 $(4\pi d/\lambda)^2$ 项引起，称为自由空间损耗，表示为 L_{free_space} 。值得注意的是，增加接收天线和发射天线之间的距离 d 将造成自由空间损耗的增加，因此必须使无线电通信链路工作在较低的频率下，以维持路径损失在一个易处理的等级上。

由 Friis 自由空间公式，可以根据给定的功率增益 G_t 和 G_r 、载波波长 λ 和距离 d 计算出路径损失 PL。为了完成链路预算分析，还需计算接收信号的平均噪声功率，下面就对此进行讨论。

8.4.5 噪声系数

要对通信系统的接收端进行噪声分析，需要一个方便的工具来度量线性双端设备的噪声性能。这种度量可通过一个所谓的噪声系数来完成。假设一个线性双端设备与一个输入端内部阻抗为 $Z(f) = R(f) + jX(f)$ 的信源相连，如图 8.7 所示。噪声电压 $v(t)$ 表示与信源内部电阻 $R(f)$ 相关的热噪声。设备的输出噪声由两部分组成，一部分由信源产生，而另一部分产生于设备本身。将中心频率为 f 且带宽为 Δf 的输出噪声功率定义为设备输出端得到的频带中的最大平均噪声功率。当负载阻抗为设备输出阻抗的共轭复数，即电阻相匹配、电抗相抵消时，就可以获得由双端设备传送到外部负载的最大噪声功率。定义双端设备的噪声系数为每单位带宽的全部输出噪声功率(由设备和信源产生)与仅由信源产生的噪声功率的比值。

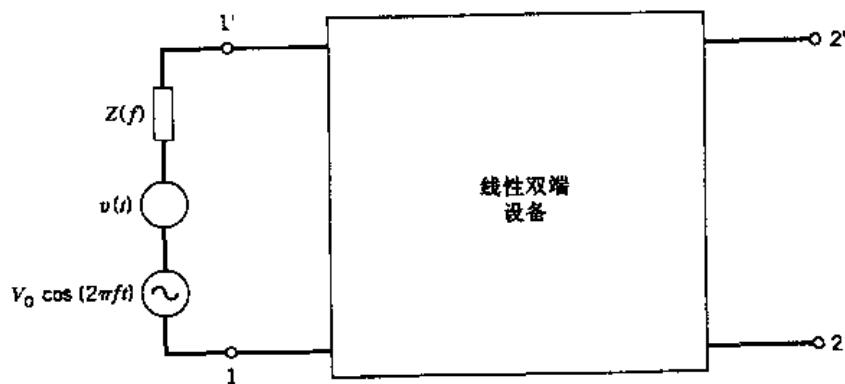


图 8.7 线性双端设备

将设备输出的全部有效噪声功率的频谱密度表示为 $S_{no}(f)$ ，由信源产生的设备输入的噪声有效功率的频谱密度表示为 $S_{ns}(f)$ 。用 $G(f)$ 表示双端设备的有效功率增益，其定义为设备输出的有效信号功率与频率为 f 的正弦信号对应的信源有效信号功率的比值。接下来，可将

设备的噪声系数 F 表示为

$$F = \frac{S_{NO}(f)}{G(f)S_{NS}(f)} \quad (8.16)$$

如果设备中不存在噪声,此时 $S_{NO}(f) = G(f)S_{NS}(f)$, 噪声系数等于 1。但在物理设备中, $S_{NO}(f)$ 大于 $G(f)S_{NS}(f)$, 因此, 噪声系数通常大于 1。噪声系数一般以 dB 表示, 即为 $10 \log_{10} F$ 。

噪声系数还可以表示为另一种形式。用 $P_s(f)$ 表示信源能提供的信号功率, 这是能够获得的最大平均功率。假设信源提供单一频率信号分量, 且开环电压为 $V_0 \cos(2\pi f t)$, 当信源连接如下负载时, 可得到有效信号功率:

$$Z^*(f) = R(f) - jX(f)$$

其中, * 表示共轭复数。在此条件下, 得到

$$\begin{aligned} P_s(f) &= \left[\frac{V_0}{2R(f)} \right]^2 R(f) \\ &= \frac{V_0^2}{4R(f)} \end{aligned} \quad (8.17)$$

因此, 设备输出端的有效信号功率为

$$P_O(f) = G(f)P_s(f) \quad (8.18)$$

将式(8.16)右端的分子和分母都乘以 $P_s(f)\Delta f$, 得

$$\begin{aligned} F &= \frac{P_s(f)S_{NO}(f) \Delta f}{G(f)P_s(f)S_{NS}(f) \Delta f} \\ &= \frac{P_s(f)S_{NO}(f) \Delta f}{P_O(f)S_{NS}(f) \Delta f} \\ &= \frac{\rho_s(f)}{\rho_O(f)} \end{aligned} \quad (8.19)$$

其中

$$\rho_s(f) = \frac{P_s(f)}{S_{NS}(f) \Delta f} \quad (8.20)$$

$$\rho_O(f) = \frac{P_O(f)}{S_{NO}(f) \Delta f} \quad (8.21)$$

$\rho_s(f)$ 为信源的有效信噪比, 而 $\rho_O(f)$ 为设备输出的有效信噪比, 两者都是在以 f 为中心、带宽为 Δf 的窄带上度量的。由于噪声系数总是大于 1, 由式(8.19)可得: 随着噪声系数的增大, 信噪比总是减小的, 这是一个很显然的结论。

噪声系数 F 是工作频率 f 的函数, 所以称为点噪声系数。相反地, 定义双端设备的平均噪声系数 F_0 为设备输出的全部噪声功率与仅由信源引起的噪声功率的比值, 即

$$F_0 = \frac{\int_{-\infty}^{\infty} S_{NO}(f) df}{\int_{-\infty}^{\infty} G(f)S_{NS}(f) df} \quad (8.22)$$

显然,当输入电路的 $R(f)$ 恒定,且在固定带宽上具有恒定增益,而在其他频率处增益为零时,热噪声的点噪声系数 F 与平均噪声系数 F_0 是一致的。

8.4.6 等价噪声温度

噪声系数 F 的缺点在于当用于具有较低噪声的设备时,所有的 F 值都接近于 1,因此很难进行比较。在这种情况下,等价噪声温度更为可取。考虑一个线性双端设备,其输出阻抗与图 8.8 所示信源的内部阻抗相匹配。图中还包括了噪声电压发生器和信源内部阻抗 R_s 。其中噪声电压的均方值为 $4kT_s\Delta f$, k 为巴特沃兹常量。因此,输入设备的噪声功率为

$$N_1 = kT \Delta f \quad (8.23)$$

用 N_d 表示双端设备的噪声功率在总有效输出噪声功率 N_2 中的对应项。即

$$N_d = GkT_e \Delta f \quad (8.24)$$

其中, G 为设备的功率增益, T_e 为等价噪声温度。从而,得到总输出噪声功率为

$$\begin{aligned} N_2 &= GN_1 + N_d \\ &= Gk(T + T_e) \Delta f \end{aligned} \quad (8.25)$$

因此,设备的噪声系数为(见图 8.8 中的输出端):

$$F = \frac{N_2}{N_2 - N_d} = \frac{T + T_e}{T} \quad (8.26)$$

由此解出等价噪声温度为

$$T_e = T(F - 1) \quad (8.27)$$

噪声系数 F 是在输入匹配的条件下度量的,噪声源的温度为 T 。 T 通常取为“室温”,即 290 K, K 代表开尔文温度。

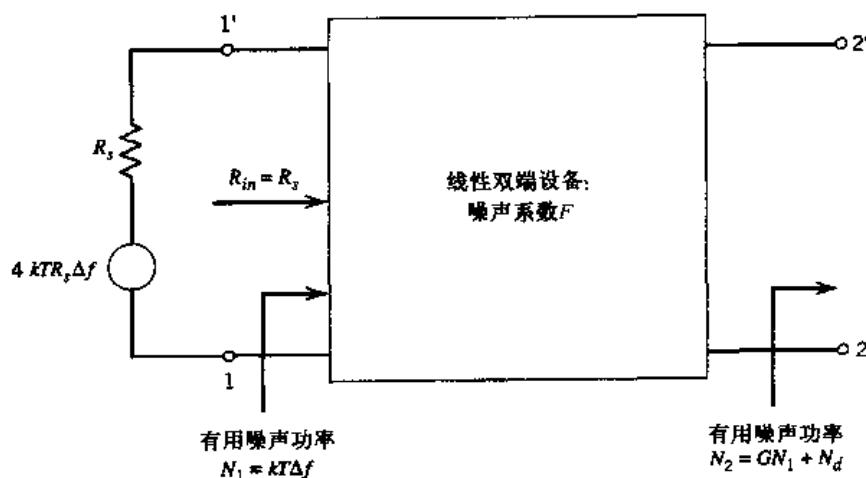


图 8.8 匹配于连接到输入的信源内部阻抗的线性双端设备

8.4.7 双端网络的级联

通常需要估计双端网络级联的噪声系数,但其单个噪声系数是已知的。在图 8.9 中,双端

网络具有一对噪声系数 F_1 和 F_2 以及功率增益 G_1 和 G_2 , 两者串行连接。假设设备是匹配的, 第二个网络的噪声系数 F_2 是在假定输入噪声功率为 N_1 的情况下定义的。



图 8.9 两个有噪声的双端网络的级联

在第一个网络的输入端, 信源产生的噪声功率 N_1 与由网络自身所产生的等价噪声功率 $(F_1 - 1)N_1$ 相加。第一个网络的输出噪声功率为 $F_1 N_1 G_1$ 。将此噪声功率加在第二个网络的输入端, 已知由第二个网络自身产生的等价附加功率为 $(F_2 - 1)N_1$ 。所以, 第二个网络的输出噪声功率为 $F_1 G_1 N_1 G_2 + (F_2 - 1)N_1 G_2$ 。可认为噪声系数 F 是实际输出噪声功率与假设的无噪声网络的输出噪声功率的比值。从而, 图 8.9 所示级联结构的整体噪声系数可表达为

$$\begin{aligned} F &= \frac{F_1 G_1 N_1 G_2 + (F_2 - 1)N_1 G_2}{N_1 G_1 G_2} \\ &= F_1 + \frac{F_2 - 1}{G_1} \end{aligned} \quad (8.28)$$

这个结果可以扩展至任意多的双端网络的级联结构:

$$F = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \frac{F_4 - 1}{G_1 G_2 G_3} + \dots \quad (8.29)$$

其中, F_1, F_2, F_3, \dots 分别为单个噪声系数, G_1, G_2, G_3, \dots 分别为有效的功率增益。式(8.29)表明, 如果图 8.9 所示级联结构的第一个阶段具有很高的增益, 则此阶段的噪声系数在整个噪声系数 F 中占主导作用。

因此, 可将任意多的双端网络级联结构的整体等价噪声温度表达如下:

$$T_e = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 G_2} + \frac{T_4}{G_1 G_2 G_3} + \dots \quad (8.30)$$

其中, T_1, T_2, T_3, \dots 为单个网络的等价噪声温度, 而 G_1, G_2, G_3, \dots 为有效的功率增益。式(8.30)就是 Friis 公式。不难再次发现: 如果第一个阶段具有很高的增益 G_1 , 则此阶段的等价噪声温度在整个等价噪声温度 T_e 中占主导作用。

例 8.1 地面终端接收机的噪声温度

图 8.10 所示为一个典型的地面终端接收机, 由一个低噪声射频放大器(低噪声放大器 LNA)、下变频器(混频器)、中频(IF)放大器构成。这些组成部分以及接收天线的等价噪声温度为

$$\begin{aligned} T_{\text{antenna}} &= 50 \text{ K} \\ T_{\text{RF}} &= 50 \text{ K} \\ T_{\text{mixer}} &= 500 \text{ K} \\ T_{\text{IF}} &= 1000 \text{ K} \end{aligned}$$

两个放大器的有效功率增益为

$$G_{RF} = 200 = 23 \text{ dB}$$

$$G_{IF} = 1000 = 30 \text{ dB}$$

可利用式(8.30)得到接收机的等价噪声温度, 即

$$\begin{aligned} T_e &= T_{\text{antenna}} + T_{RF} + \frac{T_{\text{mixer}} + T_{IF}}{G_{RF}} \\ &= 50 + 50 + \frac{500 + 1000}{200} \\ &= 107.5 \text{ K} \end{aligned}$$

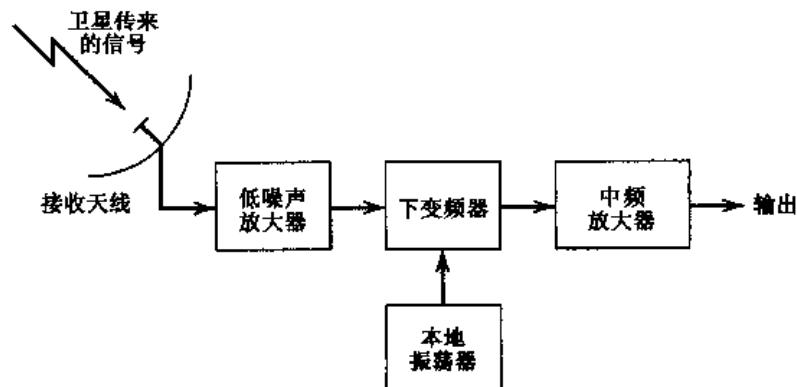


图 8.10 地面终端接收机的框图

例 8.2 数字卫星通信系统的下行链路预算分析

在数字卫星通信系统中, 整体设计和分析的一个关键部分就是下行链路功率预算, 这实际上比上行链路功率预算更加重要, 原因是下行链路功率和卫星天线大小在实际应用中都是相当有限的。本例描述了一个下行链路预算分析过程, 其中假设使系统正常工作所需的一切上行链路功率(在限制范围内)都可实现。

一个需要计算的关键参量就是接收载波功率与噪声功率谱密度之比, 记为 C/N_0 。根据式(8.14)所示的 Friis 自由空间公式, 地面终端的平均接收功率用卫星发送的平均功率 P_t 表示为

$$P_r = P_t G_t G_r \left(\frac{\lambda}{4\pi d} \right)^2$$

其中, G_t 为卫星天线的功率增益, G_r 为地面终端接收天线的功率增益, λ 为下行链路的载波波长, d 为卫星与地面终端之间的距离。假设系统的等价噪声温度为 T_e , 可利用第 1 章中的式(1.94)将噪声频谱密度 N_0 表示为 kT_e , 其中, k 为玻尔兹曼常数。此外, 由式(8.10)可知 $P_t G_t$ 等于卫星的 EIRP。因此用 P_r 除以 N_0 , 得到下行链路的 C/N_0 比值为

$$\left(\frac{C}{N_0} \right)_{\text{downlink}} = (\text{EIRP})_{\text{satellite}} \left(\frac{G_r}{T_e} \right)_{\text{earth terminal}} \left(\frac{\lambda}{4\pi d} \right)^2 \frac{1}{k} \quad (8.31)$$

对于一个给定的卫星系统, 自由空间损耗 $(4\pi d/\lambda)^2$ 为一常数。从每个终端来观察这个系统, 由式(8.31)可知: (C/N_0) 等于 G_r/T_e 。从而可用 G_r/T_e 来评估每个地面终端的“质

量”。一般将其缩写为 G/T , 用来表示地面接收终端的性能参数。将式(8.31)中的 (C/N_0) 用 dB 度量, 可将其表示为下列增益和损耗的总和:

1. $(EIRP)_{\text{satellite}}$, 以 dBW 度量, 此处, dBW 是 1 W 对应的 dB 数, 也就是 0 dBW。
2. $(G/T)_{\text{earth terminal}}$, 以 dB/K 度量, 其中 K 代表开尔文温度。
3. $L_{\text{free space}}$, 表示以 dB 度量的自由空间损耗 $10 \log_{10}(4\pi d/\lambda)^2$ 。
4. $-10 \log_{10} k$, 表示以 dBW/K·Hz 度量的, 除以玻尔兹曼常数 $k = 1.38 \times 10^{-23} \text{ J/K}$ 以后的增益值。

表 8.1 列出了一个典型本地数字卫星通信系统下行链路的增益和损耗值, 系统的其他参数为:

1. 转发器工作在最大输出功率(即无功率损失)下, 得到 EIRP 为 46.5 dBW。
2. 地面接收终端使用功率增益 $G = 45 \text{ dB}$ 的 2 m 长的碟型天线, 接收机的配置与例 8.1 中相同, 等价温度 $T = 107.5 \text{ K}$ 。因此

$$\begin{aligned}\frac{G}{T} &= 45 - 10 \log_{10} 107.5 \\ &= 45 - 20.3 \\ &= 24.7 \text{ dB/K}\end{aligned}$$

3. 自由空间损耗为

$$L_{\text{free-space}} = 92.4 + 20 \log_{10} f + 20 \log_{10} d \text{ dB} \quad (8.32)$$

其中, 下行载波频率 f 的单位为 GHz, 卫星与地面终端距离 d 的单位为 km。对于静地卫星, 卫星与地面终端之间的距离位于 36 000 ~ 41 000 km 之间。因此, 选择 $d = 40 000 \text{ km}$ 并假设 $f = 12 \text{ GHz}$, 利用式(8.32)得到

$$\begin{aligned}L_{\text{free-space}} &= 92.4 + 20 \log_{10} 12 + 20 \log_{10} 40000 \\ &= 92.4 + 21.6 + 92.0 \\ &= 206 \text{ dB}\end{aligned}$$

4. 波尔兹曼常数 $k = 1.38 \times 10^{-23} \text{ J/K}$ 对 C/N_0 的影响为

$$\begin{aligned}-10 \log_{10} k &= 10 \log_{10} 1.38 \times 10^{-23} \\ &= 228.6 \text{ dBW/K·Hz}\end{aligned}$$

综合这些增益和损耗可得

$$\left(\frac{C}{N_0}\right)_{\text{downlink}} = 93.8 \text{ dB·Hz}$$

由式(8.2), “接收”下行链路的 (C/N_0) 值可由地面接收终端的比特能量噪声谱密度比 $(E_b/N_0)_{\text{req}}$ 的“所需值”表示为

$$\left(\frac{C}{N_0}\right)_{\text{downlink}} = \left(\frac{E_b}{N_0}\right)_{\text{req}} + 10 \log_{10} M + 10 \log_{10} R \text{ dB} \quad (8.33)$$

其中, $10 \log_{10} M$ 为以 dB 表示的链路容限, R 是单位为 bps 的数据速率。链路容限允许传播中的降水损耗和其他功率损失。C 频带的链路容限通常为 4 dB, Ku 频带为 6 dB, 由于

更高的降水损耗, K 频带的频率也越高, 因此链路容限也越高。对于工作在 12 GHz 的 Ku 频带对应的频率下, 我们选择链路容限为 6 dB。利用链路预算计算得到的 $C/N_0 = 93.8 \text{ dB}\cdot\text{Hz}$, 链路容限 $10 \log_{10} M = 6 \text{ dB}$, 并假设 $(E_b/N_0)_{\text{req}} = 12.5 \text{ dB}$, 由式(8.33) 可得

$$\begin{aligned} 10 \log_{10} R &= 93.8 - 12.5 - 6 \\ &\approx 75.3 \end{aligned}$$

因此

$$R = 33.9 \text{ Mbps}$$

假设卫星传输数据采用相干 8-PSK 调制, 将 $(E_b/N_0) = 12.5 \text{ dB}$ 代入到第 6 章的式(6.47) 中, 得到符号差错概率 $P_e = 0.6 \times 10^{-3}$ 。

在最差的工作条件下, 此例中的数字卫星通信系统如果采用的是八相 PSK, 则允许下行链路以 $R = 33.9 \text{ Mbps}$ 的速率和 $P_e = 0.6 \times 10^{-3}$ 的符号差错概率传输数据。

表 8.1 例 8.2 的下行链路功率预算

变 量	值
EIRP	+ 46.5 dBW
G/T	+ 24.7 dB/K
自由空间损耗	- 206 dB
波尔兹曼常数	+ 228.6 dBW/K·Hz
C/N_0	93.8 dB·Hz

8.5 无线通信

在此节中, 将介绍多用户无线电通信系统的另一种类型, 称为无线通信, 与移动无线电通信同义。术语移动无线电通信通常包含无线通信的室内或室外的形式。无线电发射机或接收机是能够移动的。移动无线电信道的随机性, 决定了必须同时采用实际测量和统计分析。这种估计方法的目的主要是为了量化以下两个参量:

1. 信号场强中值, 用来预测为了在预定服务区域上提供可接受的覆盖质量, 发射机所需要的最小辐射功率。
2. 信号可变性, 表征了信道的衰落特性。

无线通信最值得关注的就是蜂窝无线电技术^[5], 这种技术能够将移动性引入电话网络中。有了这种功能, 用户就可以在服务区域内自由移动并同时与世界上的任何一个电话用户通信。蜂窝无线电系统的一个理想模型如图 8.11 所示。该模型由一群六边形蜂窝构成, 每个蜂窝内都有一个基站。典型的蜂窝半径为 1 ~ 12 mile (1 mile = 1609.3 m)。基站提供了移动用户与蜂窝无线电系统的接口。基站由专用线路与交换中心连接。

移动交换中心(MSC)具有两个重要的作用。它首先是蜂窝无线系统与 PSTN 的接口。同时, 它还对移动通信系统进行整体管理和控制。后一个功能是通过控制呼叫进程中的信噪比来实现的。该信噪比是在基站与呼叫中的移动用户之间进行通信时测量的。如果信噪比低于

预定的极限，则切换到另一个基站，这在移动用户离开其所处蜂窝或当无线信道减弱时发生。这个切换过程称为切换，它以透明的方式（即在呼叫过程中），无中断地将移动用户从一个基站转移到另一个基站。

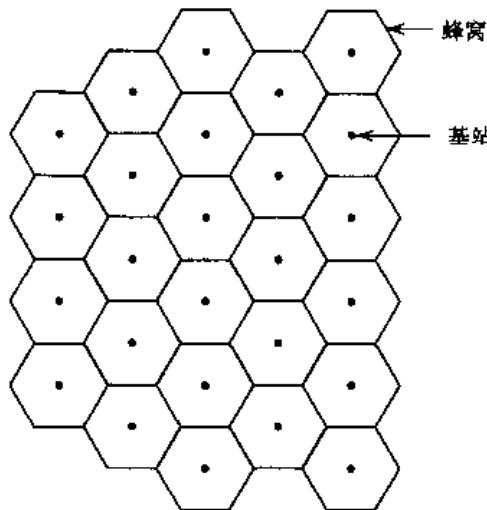


图 8.11 蜂窝无线电的理想模型

蜂窝的概念基于如下的两个重要特征：

1. **频率复用。**频率复用是指使用相同载频的无线电信道覆盖不同的区域，这些区域在物理上的分开程度足够高，因此能够确保不存在共信道干扰。频率复用能够实现两个通常意义上的功能：使每个基站的发射功率保持在最小值，将基站天线定位在刚好能提供对应蜂窝区域覆盖的高度。
2. **蜂窝分裂。**当服务要求超过了分配给某个特定蜂窝的信道数，可采用蜂窝分裂来处理蜂窝通信量的额外增长。蜂窝分裂包括对蜂窝边界的修正，之前被认为是单个蜂窝的本地区域现在包含一些更小的蜂窝，并使用这些新蜂窝的补充信道。新蜂窝的半径小于原先的蜂窝半径，称为微蜂窝。新的基站的发射功率和天线高度也相应减小，同样的频率组按照新规划进行复用。

对于蜂窝无线电系统的六边形模型，可利用六边形蜂窝的几何特性来完成无线电信道的规划方案，以决定哪个信道组分配给哪个蜂窝。从两个整数 i 和 j ($i \geq j$) 开始，这两个整数称为转移参量，是以某种方法事先确定的常量。根据六边形蜂窝几何学，六边形的 6 个链从一个六边形发散出去并在不同方向上向外延伸。因此，以任意一个蜂窝为参考蜂窝，按照如下方法，就能够发现最近的共信道蜂窝：

- 沿六边形的任意一个链移动 i 个蜂窝，逆时针转 60° ，再沿着此新方向的链移动 j 个蜂窝，如此得到的第 j 个蜂窝和参考蜂窝构成了一组共信道蜂窝。

对于不同的参考蜂窝重复这个过程直到覆盖系统中的所有蜂窝。图 8.12 显示了单个参考蜂窝的情况，其中 $i = 2, j = 2$ 。

在北美，分配给蜂窝系统的无线电频率的频带为 800 ~ 900 MHz。其中，子频带 824 ~ 849 MHz 用于接收移动单元的信号，子频带 869 ~ 894 MHz 用于向移动单元发射信号。采

用这些相对较高的频率可提供较好的、能够穿透建筑物的可控覆盖。在欧洲及其他地区,基站 - 移动体和移动体 - 基站子频带是相反的。

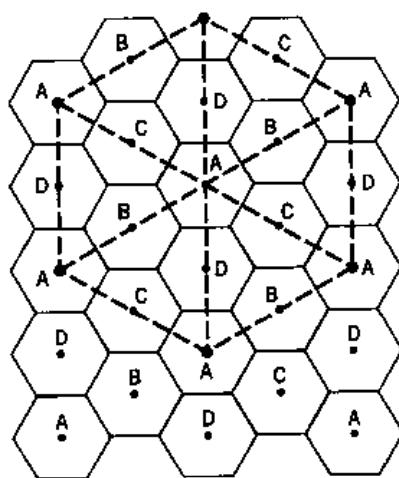


图 8.12 共信道蜂窝的确定

8.5.1 传播影响^[6]

高楼林立地区蜂窝无线电环境中的主要传播问题是由于移动单元的天线可能低于周围的建筑而引起的。简单地说,没有到达基站的“可视”路径。相反,无线电传播主要是通过周围建筑表面的散射和衍射进行的,如图 8.13 所示。由图 8.13 可知,信号功率是通过不同路径到达接收天线的。因此,我们所说的多径现象是指时延不同的多个无线电信号从不同的方向到达目的地。

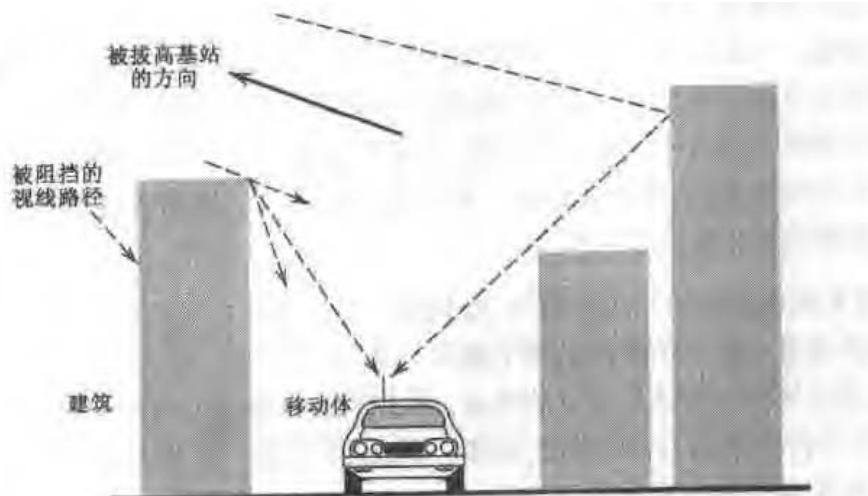


图 8.13 城区的无线电传输机制(经允许摘自 Parsons, 1992)

为了了解多径现象的本质,首先考虑一个“稳态”多径环境,其中包含一个静止的接收机和由窄带信号(例如未调制的正弦载波)构成的发射信号。假设发射信号的两个衰落分量顺序到达接收机,时延不同的结果是在接收信号的两个分量之间引入了相对相位变化。可能发生的极限情况为:

- 相对相位漂移等于 0° 在这种情况下两个分量同向相加,如图 8.14(a)所示。

- 相对相位漂移等于 180° 在这种情况下两个分量反向相加,如图8.14(b)所示。

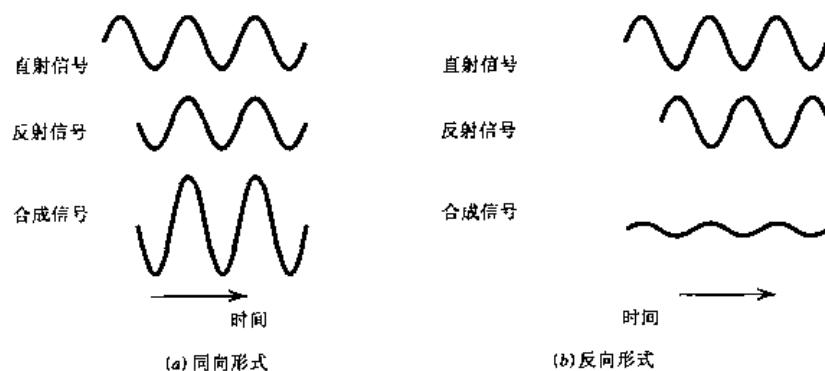


图8.14 正弦信号的多径现象

可用相量图来展示多径的同向和反向作用,分别如图8.15(a)和8.15(b)所示。值得注意的是,在此处的稳态多径环境下,接收信号的幅度不随时间变化。

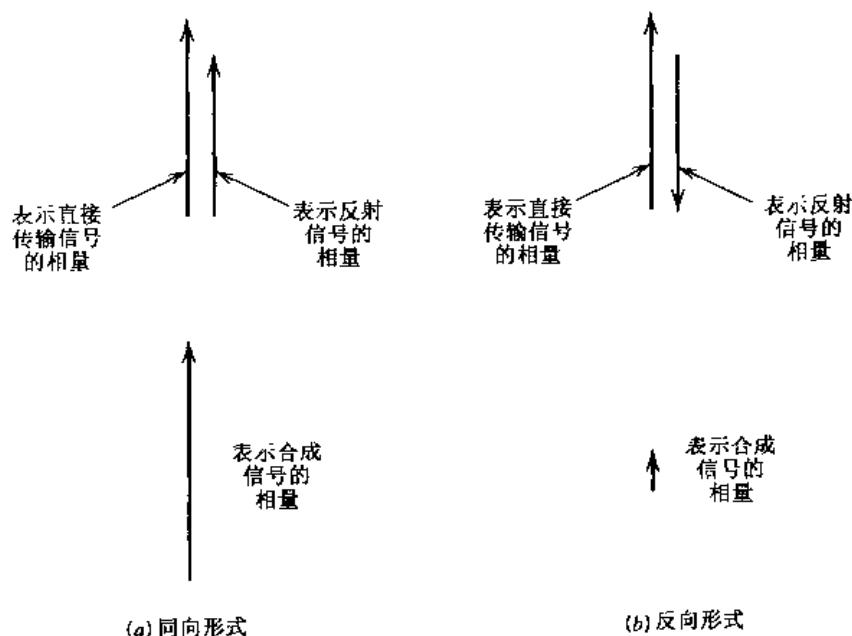


图8.15 多径的相量图表示

接下来考虑“动态”多径环境,其中接收机是运动的,发送窄带信号的两个分量通过不同长度的路径到达接收机。由于接收机的移动,每个传播路径长度会发生连续变化。因此,接收信号两个分量的相对相位漂移是接收机空间位置的函数。当接收机移动时,接收到的幅度(包络)与稳态环境中不同,不是常量,而是随距离而变化,如图8.16所示。在此图的上方,还同时给出了接收机处于不同位置时接收信号的两个分量的向量关系。图8.16表明在某些位置两个分量存在同向相加,而在某些位置上几乎完全抵消(称为信号衰落)。

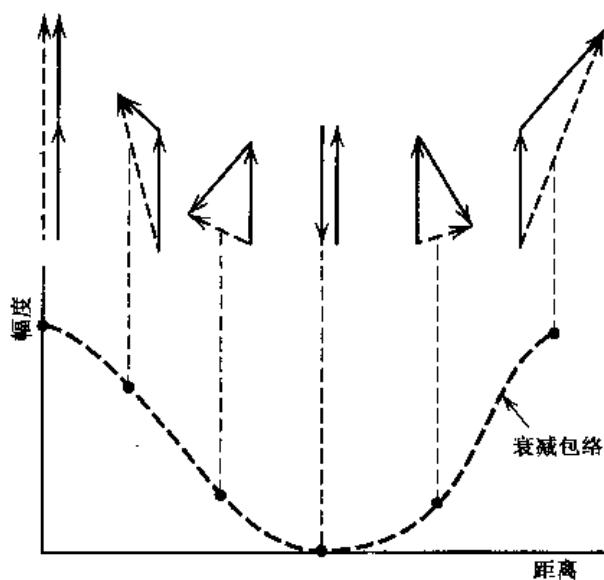


图 8.16 两个具有不同相位的输入信号结合后的包络衰落图(经允许引自 Parsons, 1992)

在一个实际的移动无线电环境中,可能存在许多不同长度的传播路径,这些路径对接收信号的影响可以不同方式进行合并。结果是接收信号的包络以复杂的方式随接收机位置而变化,如图 8.17 所示的城市中接收信号包络的实验记录。图 8.17 中的接收信号包络的单位为 dBm。单位 dBm 的定义为 $10 \log_{10}(P/P_0)$, P 表示所需测量的功率, $P_0 = 1\text{mW}$ 。在图 8.17 的情况下, P 是接收信号包络的瞬时功率。

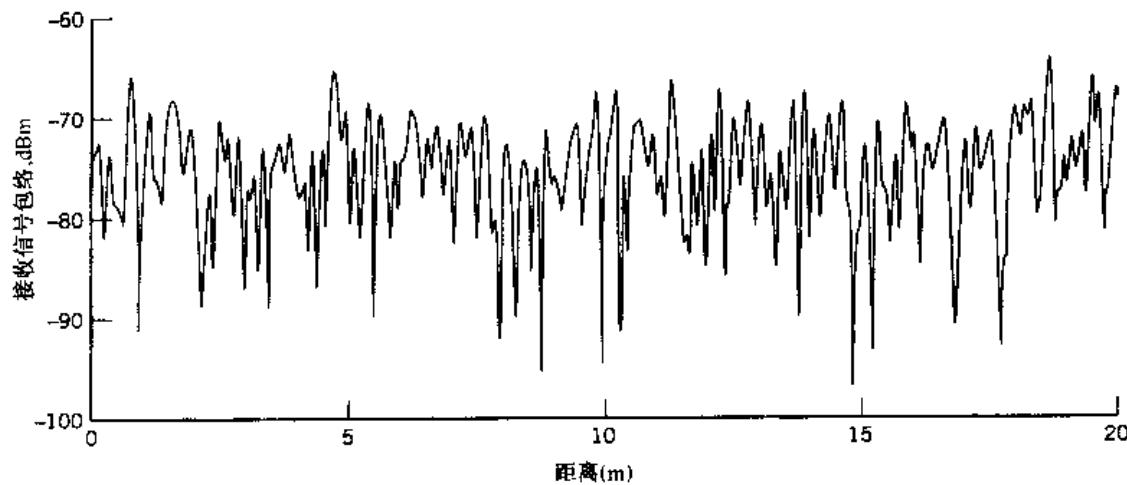


图 8.17 城镇区域内接收信号的实验记录(经允许引自 Parsons, 1992)

信号衰落是一个必然的空间现象,它随着接收机的移动在时域出现。这些变化与接收机的移动之间存在一定的关联。考虑图 8.18 所示的情况,其中,接收机沿着 AA' 以恒定速率 v 移动。假设接收信号由源自散射器 S 的无线电波产生。以 Δt 表示接收机从点 A 移动到 A' 所需的时间。使用图 8.18 中的符号,推导出无线电波路径长度的增量变化为

$$\begin{aligned}\Delta l &= d \cos \alpha \\ &= -v \Delta t \cos \alpha\end{aligned}\quad (8.34)$$

此处 α 为输入无线电波与接收机运动方向的空间夹角。对应地,在点 A' 接收信号的相角相对于点 A 的变化为

$$\begin{aligned}\Delta\phi &= \frac{2\pi}{\lambda} \Delta l \\ &= -\frac{2\pi v \Delta t}{\lambda} \cos \alpha\end{aligned}\quad (8.35)$$

其中, λ 为无线电波长。因此,频率的变化或多普勒频移为

$$\begin{aligned}\nu &= -\frac{1}{2\pi} \frac{\Delta\phi}{\Delta t} \\ &= \frac{v}{\lambda} \cos \alpha\end{aligned}\quad (8.36)$$

当无线电波在移动单元之前到达,则多普勒频移 ν 为正(导致频率的增大),否则为负。

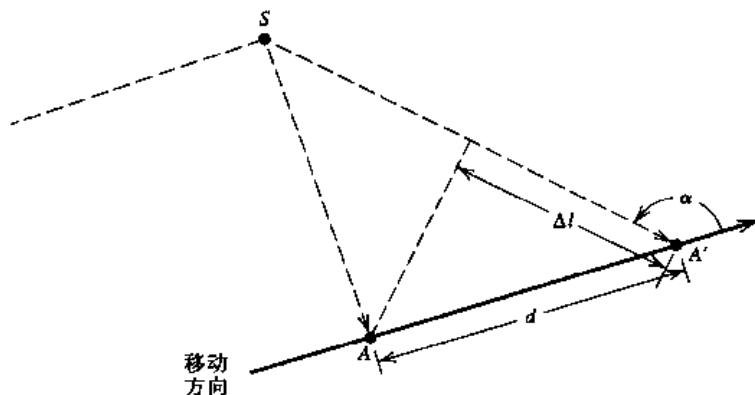


图 8.18 多普勒频移的计算

8.6 多径信道的统计特性

8.5 节所描述的多径环境的窄带特性适合于移动无线电传输,其中,信号带宽相对于传播路径时延的倒数来说是非常小的。这种环境下的多径会产生两个作用:接收信号包络的快衰落以及接收频谱的多普勒频移的传播。但是,在实际移动无线电环境中,信号可能会占据相当大的带宽,因此需要对接收信号多径传播效应进行更多细节上的考虑。本节介绍移动无线电信道的统计特性。^①

考虑一个具有多个传播路径的移动无线电信道。为了与附录 2 中的复数表示相符合,可将发射带通信号表示为

$$s(t) = \operatorname{Re}[\bar{s}(t) \exp(j2\pi f_c t)] \quad (8.37)$$

^① 对多径衰落信道统计特性的数学细节不感兴趣的读者,只需阅读本节最后的关于多径信道分类的一小部分就足够了,可以跳过本节其他内容。

其中, $s(t)$ 为 $s(t)$ 的复数(低通)包络, f_c 为标称载波频率。由于多径信道是随时间变化的, 因此信道的冲激响应与时延有关, 所以是一个时间函数。可将信道的冲激响应表示为

$$h(\tau; t) = \operatorname{Re}[\tilde{h}(\tau; t) \exp(j2\pi f_c t)] \quad (8.38)$$

其中, $\tilde{h}(\tau; t)$ 为信道(低通)复数冲激响应, τ 为时延参量。复合冲激响应 $\tilde{h}(\tau; t)$ 称为信道的输入时延扩展函数。信道输出的(低通)复数包络可由卷积定义为

$$\tilde{s}_o(t) = \frac{1}{2} \int_{-\infty}^{\infty} \tilde{s}(t - \tau) \tilde{h}(\tau; t) d\tau \quad (8.39)$$

其中比例因子 $1/2$ 是使用复数表示的结果。

一般来说, 移动无线电信道的性质仅用统计形式描述就可以了。为了分析的方便, 时延扩展函数 $\tilde{h}(\tau; t)$ 的数学模型可以是零均值的复值高斯过程。在任意时刻 t , 复包络 $|\tilde{h}(\tau; t)|$ 是瑞利分布的, 此时信道被称为瑞利衰落信道。但是, 当移动无线电系统中包含固定散射器时, 就无法再采用零均值模型来描述输入时延扩展函数 $\tilde{h}(\tau; t)$ 。在这种情况下, 采用莱斯分布来描述包络 $\tilde{h}(\tau; t)$ 更为适合, 此时信道被称为莱斯衰落信道。实数随机过程的瑞利和莱斯分布在第 1 章讨论过。在本章的讨论中, 只考虑瑞利衰落信道。

信道的时变转移函数的定义为时延参量 τ 对应的输入时延扩展函数 $\tilde{h}(\tau; t)$ 的傅里叶变换, 如下所示:

$$\bar{H}(f; t) = \int_{-\infty}^{\infty} \tilde{h}(\tau; t) \exp(-j2\pi f\tau) d\tau \quad (8.40)$$

其中, f 表示频率变量, 时变转移函数 $\bar{H}(f; t)$ 可看成是信道的频率传输特性。对于信道的统计特性, 可进行如下的假设:

- 输入时延扩展函数 $\tilde{h}(\tau; t)$ 是零均值的复值高斯过程。这里关心的只是短期的衰落, 因此也认为 $\tilde{h}(\tau; t)$ 是稳态的。由于傅里叶变换是线性的, 所以时变转移函数 $\bar{H}(f; t)$ 具有相似的线性统计特性。
- 信道是非相关性散射信道, 这说明具有不同传播时延的散射器的作用是互不相关的。

考虑输入时延扩展函数 $\tilde{h}(\tau; t)$ 的自相关函数。由于 $\tilde{h}(\tau; t)$ 是复数形式的, 因此用下式来定义自相关函数:

$$R_{\tilde{h}}(\tau_1, t_1; \tau_2, t_2) = E[\tilde{h}^*(\tau_1; t_1) \tilde{h}(\tau_2; t_2)] \quad (8.41)$$

其中, E 为数学期望运算符, $*$ 表示共轭复数, τ_1 和 τ_2 为计算中两条路径传播时延, t_1 和 t_2 为观察两条路径输出的时刻。利用时间变量 t 的稳定性以及时延变量 τ 的非相关散射性, 可重新定义自相关函数 $\tilde{h}(\tau; t)$ 为

$$\begin{aligned} R_{\tilde{h}}(\tau_1, \tau_2; \Delta t) &= E[\tilde{h}^*(\tau_1; t) \tilde{h}(\tau_2; t + \Delta t)] \\ &= r_{\tilde{h}}(\tau_1; \Delta t) \delta(\tau_1 - \tau_2) \end{aligned} \quad (8.42)$$

其中, Δt 为观察时刻之间的差值, $\delta(\tau_1 - \tau_2)$ 为 δ 函数。以 τ 代替 τ_1 , 可将式(8.42)第二行的第一项重新定义为

$$r_b(\tau; \Delta t) = E[\tilde{h}(\tau; t)\tilde{h}^*(\tau; t + \Delta t)] \quad (8.43)$$

函数 $r_b(\tau; \Delta t)$ 称为信道的多径自相关函数。

接着考虑复数形式的信道统计特性, 即时变转移函数 $\tilde{H}(f; t)$ 。由式(8.41)可得, $\tilde{H}(f; t)$ 的自相关函数为

$$R_{\tilde{H}}(f_1, t_1; f_2, t_2) = E[\tilde{H}^*(f_1, t_1)\tilde{H}(f_2, t_2)] \quad (8.44)$$

其中, f_1 和 f_2 表示发射信号频谱中的两个频率。自相关函数 $R_{\tilde{H}}(f_1, t_1; f_2, t_2)$ 提供了一种统计度量, 对于在信道传输过程失真的信号同样适用。由式(8.40)、式(8.41)和式(8.44)不难发现, 自相关函数 $R_{\tilde{H}}(f_1, t_1; f_2, t_2)$ 和 $R_b(\tau_1, t_1; \tau_2, t_2)$ 具有如下的关系:

$$R_{\tilde{H}}(f_1, t_1; f_2, t_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_b(\tau_1, t_1; \tau_2, t_2) \exp[j2\pi(f_1\tau_1 - f_2\tau_2)] d\tau_1 d\tau_2 \quad (8.45)$$

由时域的稳定性, 式(8.44)可表示为

$$R_{\tilde{H}}(f_1, f_2; \Delta t) = E[\tilde{H}^*(f_1; t)\tilde{H}(f_2; t + \Delta t)] \quad (8.46)$$

这个定义表明自相关函数 $R_{\tilde{H}}(f_1, f_2; \Delta t)$ 可通过(对信道输出结果进行互相关运算的)间隔音频对来度量。这种度量是以时域稳定性为前提的。如果同时假设具有频域稳定性, 则可进一步得到

$$\begin{aligned} R_{\tilde{H}}(f, f + \Delta f; \Delta t) &= r_H(\Delta f; \Delta t) \\ &= E[\tilde{H}^*(f; t)\tilde{H}(f + \Delta f; t + \Delta t)] \end{aligned} \quad (8.47)$$

$\tilde{H}(f; t)$ 自相关函数的特殊形式其实是多径自相关分布 $r_b(\tau, \Delta t)$ 对时延变量 τ 的傅里叶变换, 即

$$r_H(\Delta f; \Delta t) = \int_{-\infty}^{\infty} r_b(\tau; \Delta t) \exp(-j2\pi\tau\Delta f) d\tau \quad (8.48)$$

函数 $r_H(\Delta f; \Delta t)$ 称为信道的频率间隔时间间隔相关函数。

最后引入函数 $S(\tau; \nu)$, 它与多径自相关函数 $r_b(\tau; \Delta t)$ 关于变量 Δt 构成了傅里叶变换对, 即

$$S(\tau; \nu) = \int_{-\infty}^{\infty} r_b(\tau; \Delta t) \exp(-j2\pi\nu\Delta t) d(\Delta t) \quad (8.49)$$

和

$$r_b(\tau; \Delta t) = \int_{-\infty}^{\infty} S(\tau; \nu) \exp(j2\pi\nu\Delta t) d\nu \quad (8.50)$$

通过采用如下的双重傅里叶变换函数, $(S(\tau; \nu))$ 也可由 $r_H(\Delta f; \Delta t)$ 定义为关于时间变量 Δt 的傅里叶变换和关于频率变量 Δf 的傅里叶反变换。即

$$S(\tau; \nu) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r_H(\Delta f; \Delta t) \exp(-j2\pi\nu\Delta t) \exp(j2\pi\tau\Delta f) d(\Delta t) d(\Delta f) \quad (8.51)$$

图 8.19 根据傅里叶变换和傅里叶反变换展示了 $r_b(\tau; \Delta t)$, $r_H(\Delta f; \Delta t)$ 和 $S(\tau; \nu)$ 之间的函数关系。

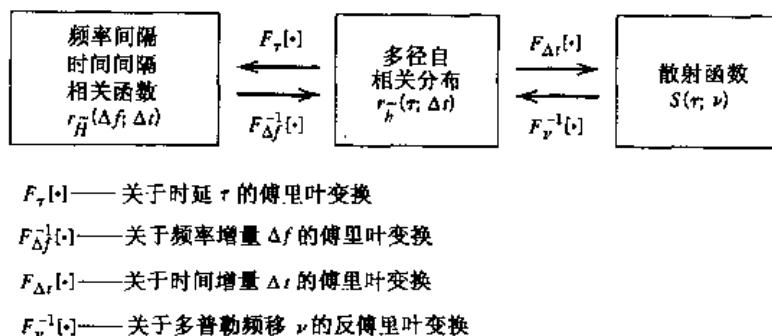


图 8.19 多径自相关分布 $r_b~(\tau; \Delta t)$, 频率间隔时间间隔相关函数 $r_H~(\Delta f; \Delta t)$ 以及散射函数 $S(\tau; \nu)$ 之间的函数关系

函数 $S(\tau; \nu)$ 称为信道的散射函数。为了得到其物理解释, 下面讨论频率 f' (相对于载波) 的单音频传输。滤波器输出的复包络为

$$\tilde{s}_o(t) = \exp(j2\pi f' t) \tilde{H}(f'; t) \quad (8.52)$$

$\tilde{s}_o(t)$ 的自相关函数为

$$\begin{aligned} E[\tilde{s}_o^*(t)\tilde{s}_o(t + \Delta t)] &= \exp(j2\pi f' \Delta t) E[\tilde{H}^*(f'; t)\tilde{H}(f'; t + \Delta t)] \\ &= \exp(j2\pi f' \Delta t) r_H(0; \Delta t) \end{aligned} \quad (8.53)$$

其中, 第二行中用到了式(8.47)。将 $\Delta f = 0$ 代入式(8.48), 并结合式(8.50), 可得

$$\begin{aligned} r_H(0; \Delta t) &= \int_{-\infty}^{\infty} r_b(\tau; \Delta t) d\tau \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} S(\tau; \nu) d\tau \right] \exp(j2\pi \nu \Delta t) d\nu \end{aligned} \quad (8.54)$$

因此, 可将积分

$$\int_{-\infty}^{\infty} S(\tau; \nu) d\tau$$

看成是以多普勒频移 ν 为频率变量的, 对应于发射信号频率 f' 的信道输出的功率谱密度。因此, 可认为散射函数 $S(\tau; \nu)$ 提供了信道输出功率的统计度量, 这个度量可以表示为时延 τ 和多普勒频移 ν 的函数。

8.6.1 时延扩展和多普勒扩展

将 $\Delta t = 0$ 代入式(8.43), 得

$$\begin{aligned} P_b(\tau) &= r_b(\tau; 0) \\ &= E[|\tilde{h}(\tau; t)|^2] \end{aligned} \quad (8.55)$$

函数 $P_b(\tau)$ 描述了传播时延 τ 处散射过程的强度(在衰落波动上进行了平均)。因此, $P_b(\tau)$ 称为信道的时延功率谱或多径强度分布。时延功率谱也可通过在所有多普勒频移上对散射函数 $S(\tau; \nu)$ 进行平均来定义。将 $\Delta t = 0$ 代入式(8.50)并应用式(8.55)的第一行, 可得

$$P_b(\tau) = \int_{-\infty}^{\infty} S(\tau; \nu) d\nu \quad (8.56)$$

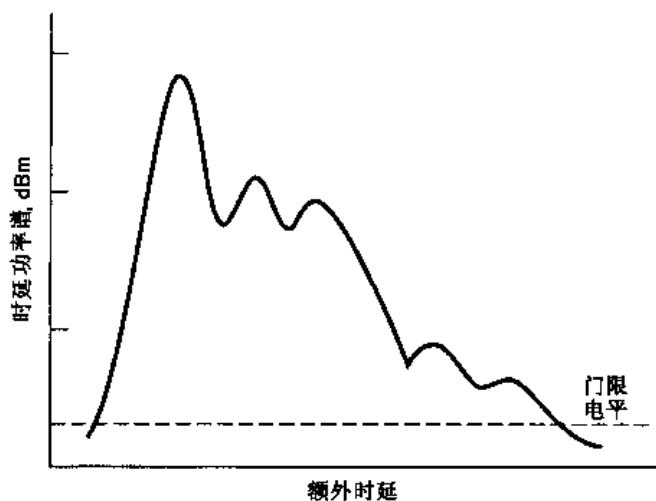


图 8.20 移动无线电信道的功率时延分布举例(经允许引自 Parsons, 1992)

图 8.20 所示为时延功率谱的例子, 图中显示了功率谱密度随额外时延 τ 变化的曲线。额外时延 τ 就是相对于最短反射路径的时延。与图 8.17 中相同, 功率是以 dBm 度量的。图 8.20 中的“门限值”定义了接收机无法正常工作的临界功率值。

我们感兴趣的 $P_b(\tau)$ 的两个统计矩为平均时延 τ_{av} 和时延扩展 σ_τ 。平均时延的定义为 $P_b(\tau)$ 的一阶中心矩(即平均值), 即

$$\tau_{av} = \frac{\int_0^\infty \tau P_b(\tau) d\tau}{\int_0^\infty P_b(\tau) d\tau} \quad (8.57)$$

时延扩展的定义为 $P_b(\tau)$ 的二阶中心矩的平方根, 即

$$\sigma_\tau = \left(\frac{\int_0^\infty (\tau - \tau_{av})^2 P_b(\tau) d\tau}{\int_0^\infty P_b(\tau) d\tau} \right)^{1/2} \quad (8.58)$$

以时延扩展 σ_τ 的倒数作为信道相干带宽的度量, 表示为 B_c 。

接下来考虑多普勒效应与信道时间变化的关系。为此, 首先令 $\Delta f = 0$, 这对应于信道上单频信号(某个适当频率的)的传输。信道的频率间隔时间间隔相关函数可简化为 $r_B(0; \Delta t)$ 。因此, 计算此函数关于时间变量 Δt 的傅里叶变换为

$$S_B(\nu) = \int_{-\infty}^{\infty} r_B(0; \Delta t) \exp(-j2\pi\nu \Delta t) d(\Delta t) \quad (8.59)$$

函数 $S_B(\nu)$ 定义了(表示为多普勒频移 ν 的函数的)信道输出的功率谱, 因此称为信道的多普勒频谱。多普勒频谱也可通过散射函数在所有可能的传播时延上的平均来表示, 即

$$S_B(\nu) = \int_{-\infty}^{\infty} S(\tau; \nu) d\tau \quad (8.60)$$

多普勒频移 ν 为正或为负的可能性是相同的。因此，多普勒频移的平均值为零。多普勒频谱的二阶矩的平方根为

$$\sigma_\nu = \left(\frac{\int_{-\infty}^{\infty} \nu^2 S_H(\nu) d\nu}{\int_{-\infty}^{\infty} S_H(\nu) d\nu} \right)^{1/2} \quad (8.61)$$

参数 σ_ν 提供了多普勒频谱宽度的度量，因此称为信道的多普勒扩展。多普勒扩展的倒数称为信道的相干时间，用 τ_c 表示。

另一个有用的参数是信道的衰落速率。对于一个瑞利衰落信道，平均衰落速率与多普勒传播 σ_ν 的关系为

$$f_e = 1.475 \sigma_\nu \text{ crossings/s} \quad (8.62)$$

顾名思义，衰落速率提供了信道衰落速度的度量。

移动无线电环境中的一些典型参量值为：

- 时延扩展 σ_τ 约等于 $20 \mu s$ 。
- 由于车辆移动造成的多普勒传播 σ_ν 可达到 $40 \sim 80 \text{ Hz}$ 。

8.6.2 多径信道的分类

多径信道可能的衰落形式决定于信道特性是在频域还是在时域上观察得到的。

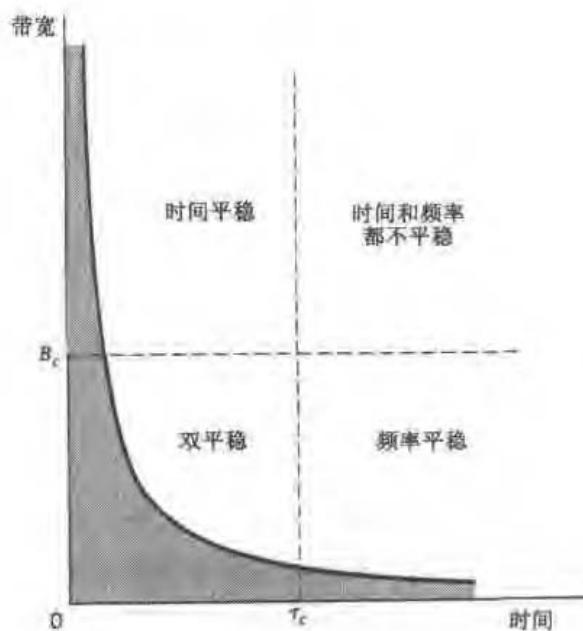
当在频域上观察信道特性时，所关心的参量是信道的相干带宽 B_c ， B_c 是传输带宽的度量，在此带宽上信号通过信道时失真变得很明显。如果信道的相干带宽相对于发射信号带宽来说很小，多径信道可认为是具有频率选择性的。在这种情况下信道具有滤波效果，因为信道对两个频率间隔大于信道相干带宽的正弦分量的处理方式是不同的。但是如果信道的相干带宽大于消息带宽，则可认为衰落是频率非选择性的或是频率平坦的。

当在时域上观察信道特性时，所关心的参量是相干时间 τ_c ， τ_c 是发射信号持续时间的度量，在此时间上信号通过信道时失真变得很明显。如果信道相干时间相对于接收信号的持续时间（即信号的传输时间）来说较小，多径信道可认为是具有时间选择性的。对于数字传输，接收信号的持续时间是符号持续时间加上信道的时延扩展。但是，如果信道的相干时间大于接收信号持续时间，则可认为衰落是时间非选择性的或是时间平坦的，因为此时在发射信号看来，信道具有时间不变性。

根据上述分析，可将多径信道分类如下：

- 平坦信道，频率和时间均是平稳的。
- 频率平坦信道，只有频率是平稳的。
- 时间平坦信道，只有时间是平稳的。
- 非平坦信道，频率和时间均不是平稳的。这种信道有时称为双重弥散信道。

基于这种方法的多径信道的分类如图 8.21 所示。图中阴影部分的禁区是由带宽和持续时间的反比例关系得来的。

图 8.21 复用信道的四种类型; T_c = 相干时间, B_c = 相干带宽

8.7 瑞利衰落信道上的二进制信号

第6章中确定了AWGN信道上的二进制数据传输的平均符号差错概率。在移动无线电环境中，则需要考虑其他因素的影响，即由于多径效应造成的接收信号的幅度和相位波动。考虑瑞利衰落信道上的二进制数据传输，其接收信号的(低通)复数包络的数字模型为：

$$\bar{x}(t) = \alpha \exp(-j\phi)\bar{s}(t) + \bar{w}(t) \quad (8.63)$$

其中， $\bar{s}(t)$ 为发射(带通)信号的复数包络， α 为描述传输衰落的符合瑞利分布的随机变量， ϕ 为描述传输相位变化的符合均匀分布的随机变量， $\bar{w}(t)$ 为复高斯白噪声过程。假设信道对时间和频率都是平坦的，因此可无错误地估计接收信号的相位变化 ϕ 。假设数据传输中使用相干二进制PSK。在这种情况下，在一个比特间隔中 α 为常量，根据第6章的式(6.20)，只由加性高斯白噪声产生的平均符号差错概率(即误比特率)可表示为：

$$P_e(\gamma) = \frac{1}{2} \operatorname{erfc}(\sqrt{\gamma}) \quad (8.64)$$

其中， γ 为发射信号的比特能量与噪声谱密度的比值 E_b/N_0 的衰减因子，即

$$\gamma = \frac{\alpha^2 E_b}{N_0} \quad (8.65)$$

在移动无线电信道中，可认为 $P_e(\gamma)$ 是常量 α 的条件概率。因此，为了计算衰落和噪声均存在的时的平均符号差错概率，就必须将 $P_e(\gamma)$ 对 γ 的所有可能值取平均，即

$$P_e = \int_0^\infty P_e(\gamma) f(\gamma) d\gamma \quad (8.66)$$

其中， $f(\gamma)$ 为 γ 的概率密度函数。由式(8.65)可知 γ 取决于 α 的平方值。由于 α 是瑞利分布的，所以 γ 具有双自由度的 χ^2 (Chi-Square) 分布。^[7] 特别地，将 γ 的概率密度函数表示为

$$f(\gamma) = \frac{1}{\gamma_0} \exp\left(-\frac{\gamma}{\gamma_0}\right), \quad \gamma \geq 0 \quad (8.67)$$

γ_0 为每比特接收信号能量的平均值和噪声功率谱密度之比, 即

$$\begin{aligned} \gamma_0 &= E[\gamma] \\ &= \frac{E_b}{N_0} E[\alpha^2] \end{aligned} \quad (8.68)$$

其中, $E[\alpha^2]$ 为瑞利分布随机变量 α 的均方值。将式(8.64)、式(8.67)代入式(8.66), 然后进行积分, 得到最终的结果为

$$P_e = \frac{1}{2} \left(1 - \sqrt{\frac{\gamma_0}{1 + \gamma_0}} \right) \quad (8.69)$$

式(8.69)定义了平坦瑞利衰落信道的相干二进制 PSK 的误比特率。同理可得相干二进制 FSK、二进制 DPSK 和非相干二进制 FSK 的相应误比特率。表 8.2 对相应结果进行了总结。图 8.22 显示了利用表 8.2 中的公式绘制出的以 dB 表示的误比特率随 γ_0 的变化情况。为进行比较, 也在图 8.22 中绘制了无衰落信道的相干二进制 PSK 和非相干二进制 FSK 的误比特率曲线。可见, 瑞利分布造成数字通带传输系统噪声性能的严重下降, 这可用(相同误比特率下)瑞利分布相对于衰落信道增加的平均信噪比的 dB 值来衡量。特别地, 对于较大的 γ_0 , 可得到表 8.2 最后一列的近似公式, 其中, 误比特率的渐进递减与每比特信噪比 γ_0 的变化成反比例。这与在无衰落信道中有很大的区别。在无衰落信道中, 误比特率随 γ_0 的渐进而发生的递减符合指数律。

表 8.2 平稳瑞利衰落信道中的二进制信号的误比特率

信号类型	误比特率 P_e 的精确公式表示	γ_0 很大时误比特率的近似公式
二进制 PSK	$\frac{1}{2} \left(1 - \sqrt{\frac{\gamma_0}{1 + \gamma_0}} \right)$	$\frac{1}{4\gamma_0}$
相干二进制 FSK	$\frac{1}{2} \left(1 - \sqrt{\frac{\gamma_0}{2 + \gamma_0}} \right)$	$\frac{1}{2\gamma_0}$
二进制 DPSK	$\frac{1}{2(1 + \gamma_0)}$	$\frac{1}{2\gamma_0}$
非相干二进制 FSK	$\frac{1}{2 + \gamma_0}$	$\frac{1}{\gamma_0}$

上述区别的实际应用就是, 在移动无线电环境中, 必须大大提高平均信噪比(相对于无衰落环境), 才能保证足够低的误比特率。为达到这个目的, 就必须增加发射功率、天线尺寸等, 其实施费用都为相当昂贵的。此外, 还可以采用特殊的调制和接收技术, 这些技术应当具有较好的抗衰落能力。这些技术当中, 最著名且应用最广泛的就是采用了分集技术的多接收机, 下面将对其进行简单的讨论。

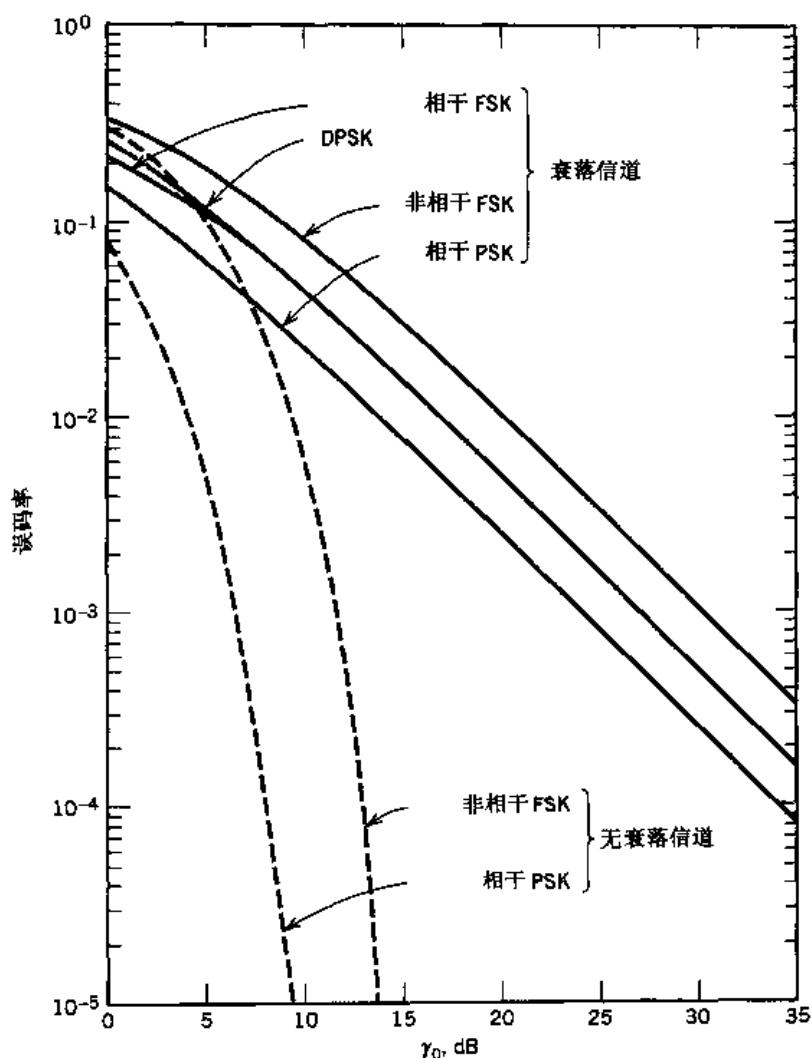


图 8.22 瑞利衰落信道上的二进制调制机制的性能,用连续的曲线表示;虚线表示无衰落信道的情况

8.7.1 分集技术

分集可认为是一种冗余。如果消息信号的多个拷贝可在多个独立的衰落信道中同时传输,则可能性较大的是,至少有一个接收信号没有发生严重衰落。有很多方法可以达到这个效果,下面就介绍其中的一种,称为分集技术。分集技术可分为:

- 频率分集
- 时间(信号接收)分集
- 空间分集

在频率分集中,信号传输时使用多个载波,这些载波在空间上相互分离,足以提供各自独立的衰落信号。选择频率间隔等于或大于信道的相干带宽,就能够实现这一点。

在时间分集中,相同消息信号在不同时隙中传输,相邻时隙之间的间隔等于或大于信道的相干时间。时间分集可看成是用于差错控制编码的重复码。(差错控制编码将在第 10 章中进行介绍。)

在空间分集中,采用了多个发射或接收天线(或同时采用多个发射和接收天线),在选择相邻天线之间的间隔时要保证衰落的独立性,相邻天线间隔至少为无线电波长的 7 倍,才能够满足这个要求。

假设采用上述某种分集方法建立 L 个独立衰落信道。其中可采用包含了 L 个分离的接收机的线性分集合并结构,如图 8.23 所示。这种系统设计只是为了补偿衰落信道的短期效应。此外,假设存在信道衰落因子 $\{\alpha_l\}$ 的无噪声估计值和信道相移 $\{\phi_l\}$ 。要通过线性合并器获得二进制数据传输(仅以此为例,无特别意义)的最佳性能,可将第 l 个接收机中第 k 个匹配滤波器的输出 $v_{lk}(t)$ 与 $\alpha_l \exp(j\phi_l)$ (即第 l 个信道增益的复共轭)相乘。其中 $l = 1, 2, \dots, L$, 且 $k = 0, 1$ 。于是,线性合并器得到两个定义如下的复包络输出:

$$\tilde{v}_k(t) = \sum_{l=1}^L \alpha_l \exp(j\phi_l) v_{lk}(t), \quad k = 0, 1 \quad (8.70)$$

可见, $\alpha_l \exp(j\phi_l)$ 为权重因子。输出复包络 $\tilde{v}_0(t)$ 对应于发送符号 0,另一个输出复包络 $\tilde{v}_1(t)$ 则对应于发送符号 1。 $\tilde{v}_0(t)$ 和 $\tilde{v}_1(t)$ 的实部用于判决过程。此处描述的情况对应于二进制 FSK。在二进制 PSK 的情况下,只需要一个匹配滤波器,线性合并器只提供一个复包络输出。其中仍将合并器输出的实部用于判决过程中。

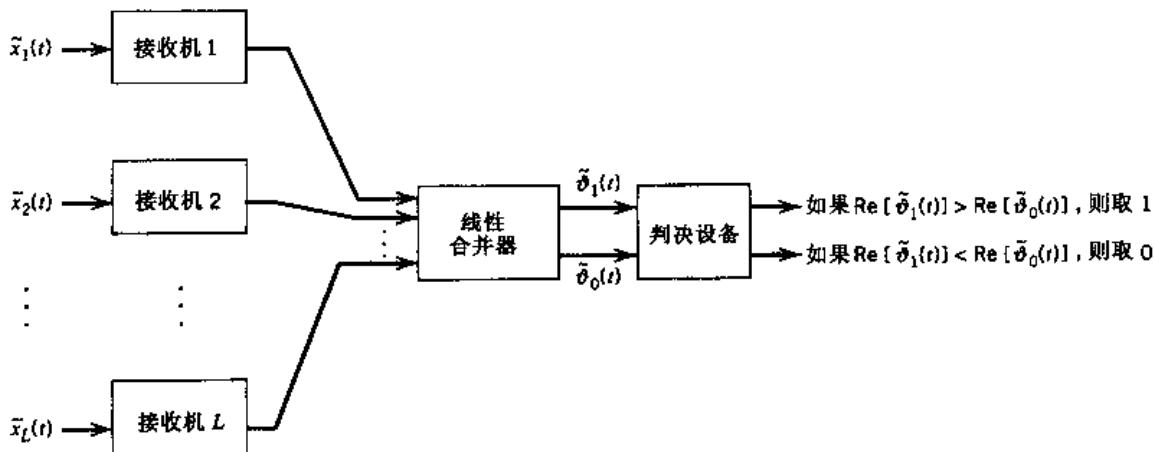


图 8.23 空间分集技术框图

此处描述的线性合并器中,“瞬时”输出信噪比是每个独立分支(信道)上的瞬时信噪比的和。线性合并器的这种优化形式称为最大比合并器(见习题 8.17)。

图 8.24 显示了具有 $L = 2, 4$ 条独立衰落信道的相干二进制 PSK、二进制 DPSK 和非相干二进制FSK 的噪声性能。为了进行比较,图中还包括了未采用分集(即 $L = 1$)的衰落信道的相应曲线。图 8.24 清楚地显示了分集作为一种减轻瑞利分布短期效应的方法,是非常有效的。

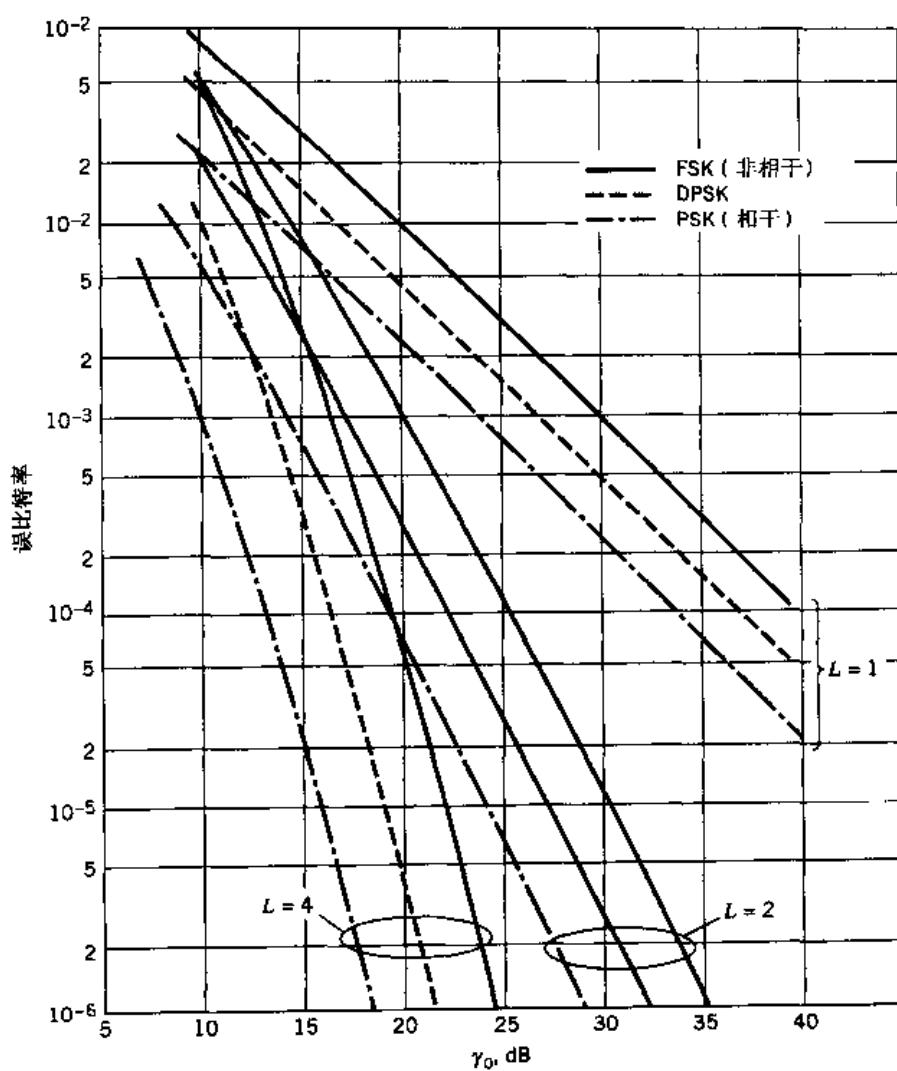


图 8.24 二进制信号分集技术的性能(经 McGraw-Hill 允许摘自 Proakis, 1995)

8.8 TDMA 和 CDMA 无线通信系统^[8]

在无线通信(如通常的电话)中, 用户希望能够同时说话和收听。为满足这一需求, 就需要一定形式的双工通信。能满足这个要求的一个方法就是提供两个频率段, 一个用于从基站到移动台的前向链路, 另一个用于移动台到基站的反向链路。如前所述, 北美的子频带 869~894 MHz 用于前向链路, 子频带 824~849 MHz 用于反向链路。这种双工称为频分双工(FDD)。实际上, 频分双工是表 8.3 中的两种广泛应用的无线系统的组成部分。

第一个 GSM 系统采用的是 TDMA。由 8.2 节可知, TDMA 系统的每个用户可以在预定的时隙访问信道, 且在每个时隙上都有一个用户占用全部信道。数据以脉冲形式在信道上传输, 如图 8.25 的帧结构所示。GSM 的基本框架由 8 个 577 μs 的时隙构成。每个 57 比特数据脉冲前的 1 比特标识位用于区分此数据比特是数字音频还是其他信息承载信号。3 个尾比特都是逻辑 0, 用于信道编码数据比特的卷积译码(卷积码将在第 10 章中讨论)。时隙中间的 26 比特的训练序列用于信道均衡。最后, 保护时间占 8.25 比特, 包含在每个时隙的末端, 以防止从移动台到基站的数据脉冲相互重叠, 在保护时间间隔内不发送任何信号, 即可实现上述目的。每个时隙由 156.25

个比特构成,其中有 40.25 个比特的开销(忽略 2 比特的标识位),GSM 的帧效率为

$$\left(1 - \frac{40.25}{156.25}\right) \times 100 = 74.24\%$$

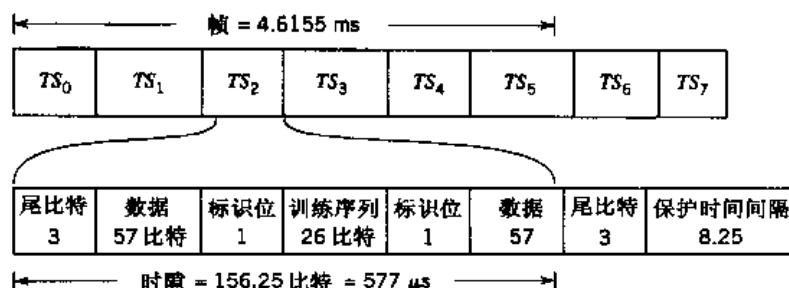


图 8.25 GSM 无线通信系统的帧结构

表 8.3 中的 IS-95 无线通信系统采用的是 CDMA。由 8.2 节可知,CDMA 系统的每个用户被赋予一个独特的扩展码(PN 序列),从而允许每个用户在所有的时间上都能占用全部信道。因此,CDMA 系统中存在一种新的干扰,称为多址干扰(MAI)。多址干扰是由于扩展码偏离了最佳正交性而产生的。另一个需要引起注意的相关现象为远-近效应。远-近效应是在基站接收的由移动台传来的信号功率不相等的情况下发生的。在这种情况下,移动台传来的最强接收信号独占了基站的解调过程,因此将对其他用户造成不利。为了解决远-近效应的问题,通常在基站采用功率控制,即基站对每个移动台的发射信号功率进行控制。功率控制的采用对 CDMA 系统来说尤其重要,其中还有另一个原因。多址系统的目的就是为了系统容量最大化。系统容量的定义为在一定的资源下,系统能够提供可靠服务的最大可能的用户数。显然,如果每个移动台都随意增加其发射功率而不顾及其他用户,系统容量就会受到影响,这是因为,发射功率的增加将导致系统多址干扰的相应增大。为了系统容量的最大化,每个移动台的发射机必须处于基站的控制下,以使信噪比维持在正常服务的最小值。

表 8.3 两个广泛应用的无线通信系统的总结

项 目	GSM*	IS-95**	注 释
双信道的个数	125	20	CDMA 中假设在每个方向上均为 12.5 MHz
信道带宽(kHz)	200	1250	
多址的类型	TDMA	CDMA	
每信道的访问用户	8	20 ~ 35	TDMA 确定性在于每信道的访问用户数由提供的时隙个数定义。另一方面,CDMA 系统为干扰受限的,这是因为它有一个每信道访问用户数的软限制
调制类型(Kbps)	GMSK	BPSK/QPSK	CDMA 中,数据调制为 BPSK,但其传播为 QPSK
数据速率(Kbps)	270.833	9.6 或 14.4	
帧周期(ms)	4.615	20	对 CDMA, 帧周期等于话音编码器/译码器的周期

* GSM 代表全球移动通信系统,最初是作为 Groupe de travail Spéciale pour les services Mobiles 的首字母而引入的。

** IS 代表过渡标准。

8.8.1 RAKE 接收机

RAKE 接收机^[9]是 CDMA 无线通信系统的一个重要部分。RAKE 接收机由 20 世纪 50 年代开始采用的用于均衡多径影响的“分集”接收机发展而来。发射信号的有用信息包含在接收信号的多径分量中。因此,将多径信号近似地看成不同时延的回波的线性组合,然后通过在 RAKE 接收机中采用一种相关方法来单独检测出每个回波信号,最后将这些信号进行代数相加,这样就能有效克服多径的影响。于是,由多径产生的 ISI 就可通过向检测到的回波中重新插入不同的时延来解决。

图 8.26 说明了 RAKE 接收机的基本思想。接收机由多个同时工作的相关器并行相连构成。每个相关器都有两个输入:(1)接收信号的时延形式;(2)PN 序列的拷贝,将作为扩展码以产生发射机的扩频调制信号。PN 序列实质上为“参考信号”。PN 序列的标称带宽为 $W = 1/T_c$,其中, T_c 为码片持续时间。由第 7 章介绍的扩频调制可知,PN 序列的自相关函数在 $1/W$ 处有惟一的峰值,而在 PN 序列一个周期(即一个符号周期)内的其他地方又减小为零。因此,只需使 PN 序列的带宽 W 足够大,就能够“辨认”接收信号中的重要回波。为了确保相关器的所有输出都为同性相加,接收机还需进行另外两个操作。这两个操作由标有“相位和增益调节器”字样的功能块来完成:

1. 每个相关器输出中都加入了一个合适的时延以使相关器输出的相角相互一致。
2. 对相关器输出进行加权,以便使多径环境中强路径上相关器的作用得以增强,使那些与重要路径无关的相关器的作用得以相应抑制。

加权系数 α_k 可根据最大比合并原理^[10]计算得到:

当加权幅度与相应信号强度成比例时,可使加权和的信噪比最大。其中,加权和的每个分量均由信号与固定功率的加性噪声相加而成。

线性合并器输出为

$$y(t) = \sum_{k=1}^M \alpha_k z_k(t) \quad (8.71)$$

其中, $z_k(t)$ 为第 k 个相关器的相位补偿输出, M 为接收机中相关器的个数。假设接收机中采用了足够多的相关器来得到足够大的时延区域,以包含所有可能在多径环境中产生的重要回波。这使得从输出 $y(t)$ 来看,仿佛发射机和接收机之间只有一个传播路径,而不是同时存在多个路径。

为了简化表述,假设发射机在进行扩频调制时,图 8.26 中的接收机采用的是二进制 PSK。因此,图 8.26 中的最后一个步骤就是在比特时隙 T_b 上对线性合并器的输出 $y(t)$ 进行积分,然后判断在此比特时隙内发送的二进制符号为 1 还是 0。

之所以称为 RAKE 接收机,是因为并行相关器的形状看上去类似于一个耙子(rake)。由于扩频调制是 CDMA 无线通信的基础,所以 RAKE 接收机的设计就显得尤为重要。^[11]

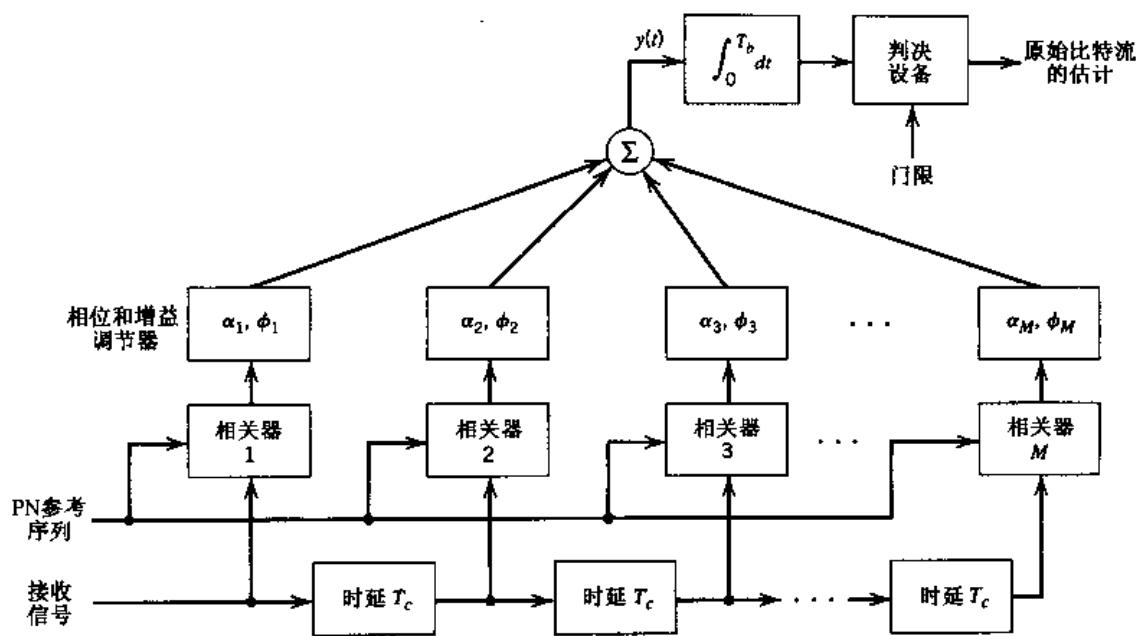


图 8.26 RAKE 接收机框图

8.9 无线通信中的语音信源编码

为有效利用信道带宽，数字无线通信系统（如 TDMA 或 CDMA）都依赖于语音编码来去除语音中几乎所有的冗余，从而维持高质量的译码。通常的方法是采用信源编码，即语音的线性预测编码（LPC）。

此节将介绍语音编码的两种不同技术：多脉冲激励 LPC 与码激励 LPC，它们分别用于 GSM 与 IS-95。本文只对这两个语音编码技术进行概念上的说明。^[12]

8.9.1 多脉冲激励 LPC

这种语音编码技术利用了合成分析原理，即在编码器的设计中包括了译码器的副本。编码器包含三个主要部分，如图 8.27(a)所示：

1. 语音预测模型的合成滤波器。这种滤波器由一个用来模型化语音的短期频谱包络的全极点滤波器（即滤波器的转移函数只含极点）构成。术语短期是指滤波器参量是基于语音信号的当前样值计算得到的。其中，语音信号的当前值由 8~16 个先前的样值预测而得。合成滤波器也可能还包含一个用长期预测器来模型化语音频谱的精细结构。在这种情况下，长期预测器与短期预测器是级联的。在任何情况下，合成滤波器的功能都是为了提供原始语音的高质量合成分本。
2. 产生（用于合成滤波器的）激励的激励发生器。此激励由 5~15 ms 的数量一定的脉冲组成。单个脉冲的幅度和相位都是可调节的。
3. 差错最小化。用来最优化原始语音与合成语音之间可感知的加权错误。进行最小化的目的的是为了最优化激励中采用的脉冲幅度和相位。最小化中通常采用均方差准则。

如图 8.27(a)所示,编码器的三个部分构成了闭环最优化过程。该过程允许编码器在比特率低于 16 Kbps 时仍维持高质量的语音通信。

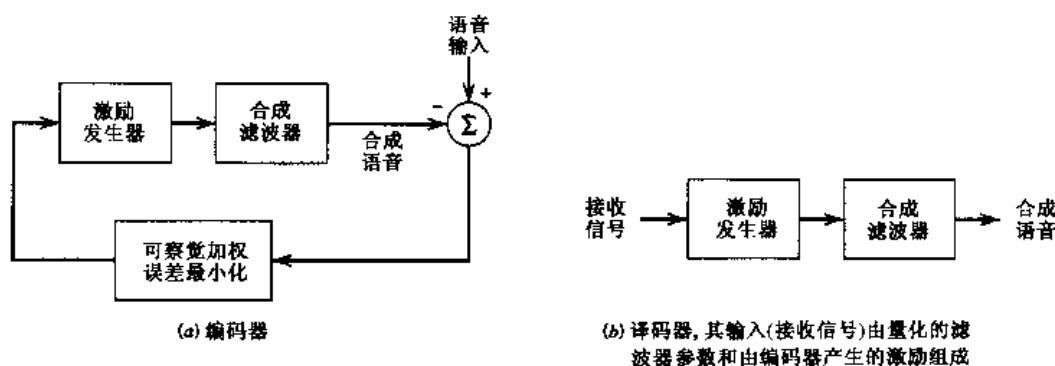


图 8.27 多脉冲激励 LPC

编码过程包括两个主要步骤:

- 合成滤波器的自由参数采用实际语音样本作为输入计算得到。此计算是在最优化回路之外进行的,持续时间为 10~30 ms(这段时间内语音信号被视为是伪静止的)。
- 合成滤波器的最优激励通过最小化图 8.27(a)所示的闭环可察觉加权误差得到。

因此,语音样本分不同帧(10~30 ms)来计算滤波器参数,每个帧又进一步分为子帧(5~15 ms)来最优化激励。量化的滤波器参数和量化的激励构成了发射信号。

这里允许帧与帧之间的滤波器参数各不相同,也允许子帧与子帧之间的激励各不相同,因此编码器能够跟踪语音的非稳态特性,尽管是在分批基础上进行的。

译码器位于接收机内部,由两部分组成:激励发生器和合成滤波器,如图 8.27(b)所示。这两个部分与编码器中的对应部分相同。译码器的作用为利用接收信号来得到原始语音信号的合成信号,这是由译码激励通过(参数设置与编码器中一致的)合成滤波器获得的。

为了减少编译码器(编码器/译码器)的计算复杂度,激励的单个脉冲之间的间隔限制为一个相同值。得到的分析合成编译码器具有规则脉冲激励。

8.9.2 码激励 LPC

图 8.28 为码激励 LPC 的框图,一般称为 CELP。CELP 的突出特点就是采用随机(零均值白色高斯)矢量的预设码本作为合成滤波器的激励源。合成滤波器自身由两个全极点滤波器级联而成,一个用于短期预测,另一个用于长期预测。

和多脉冲激励 LPC 一样,首先采用实际语音样本作为输入,计算出合成滤波器的自由参数。接着将特定矢量(码字)的选择保存到激励码本中。通过最小化原始语音与合成语音(即合成滤波器的输出)之间的可察觉加权误差的平均功率,可将图 8.28 中的增益因子 G 最优化。从码本中选择的随机矢量的地址、相应的量化增益因子和量化滤波器参数构成了发射信号。

将码本的相同副本同时提供给译码器和合成滤波器。因此,给定接收信号,译码器就能够确定合成滤波器的参数,并为合成滤波器确定合适的激励,从而提供原始语音信号的合成结果。

CELP能够在低于8Kbps的速率下提供高质量的语音。但其计算复杂度很高,这是由于需要对激励码本进行彻底搜索。特别地,译码器中的加权合成语音过程中,必须对码本中的每一项进行计算,然后与原始加权语音进行比较。不过,由于数字信号处理和VLSI技术的进步,CELP编译码器的实时应用已成为可能。

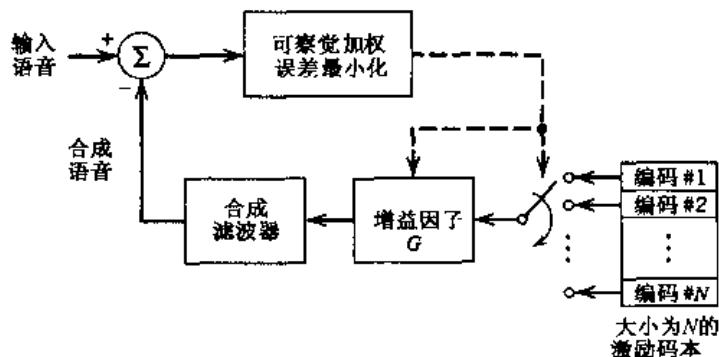


图 8.28 CELP 的编码器。发射信号由选自码本的码地址、量化的 G 和量化的滤波器参数构成

8.10 无线通信中的自适应天线阵列^[13]

无线通信的目的是使尽可能多的用户进行可靠通信,而不需要考虑其位置和移动性。由8.5节和8.6节可知,这个目标的实现受到以下三个主要因素的严重阻碍:

1. 由于不同传播路径之间的相位削减,多径会造成严重的信号衰落。衰落引起有用信号功率的降低,从而使系统抗噪声性能下降。
2. 多个传播路径之间的传播时延差异造成的时延扩展。当时延扩展超过信号持续时间的10%时,接收信号的ISI达到很高的程度,从而引起可用数据速率的降低。
3. 蜂窝系统中的共信道干扰。系统的可用频率信道被分为不同的组,每组被分配给特定的蜂窝,系统中多个蜂窝使用相同的频率组。共信道干扰限制了系统容量(即系统能可靠服务的最大可能的用户数)。

蜂窝系统的每个基站均采用120°的扇区,在给定频率下,只有一个用户访问基站的扇区。可通过使用三个相同但完全分离的天线阵列来克服基站的多径衰落和共信道干扰效应。其中,基站的每个扇区使用一个天线阵列。时延扩展的补偿将在后面进行考虑。图8.29为阵列信号处理器的框图,其中,假设有 N 个用户信号被基站的特定扇区接收,此扇区的阵列包含 M 个相同的天线单元。将一个用户作为观察对象,则其他 $N-1$ 个用户对其产生共信道干扰。除了共信道干扰,阵列信号处理器输入的每个分量还受到加性高斯白噪声的干扰。此处的分析是对基带信号进行的,基带信号通常具有复数值。这就意味着信道和阵列信号处理器必须具有各自的复数特性。图8.29所示的结构显示了与被考察用户相关的一个输出。处于基站其他频率上的用户所对应的阵列信号处理器是相同的。

多径信道由信道矩阵表征,用 \mathbf{C} 表示。矩阵 \mathbf{C} 为 $M \times N$ 的,因此可扩展为 N 个列矢量,即

$$\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_N] \quad (8.72)$$

其中,每个矢量都为 M 维的。

给定图 8.29 所示的结构,目标是设计一个接收机的线性阵列信号处理器,需满足两个条件:

1. $N - 1$ 个干扰用户引起的共信道干扰已被去除。
2. 将被观察用户的输出信噪比最大化。

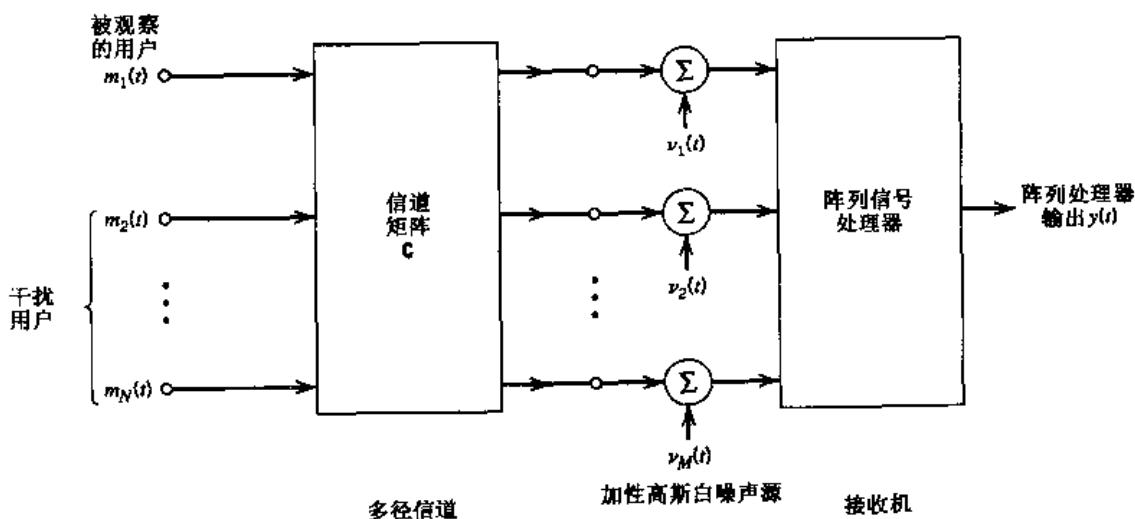


图 8.29 阵列信号处理器框图,包含了由一个复用信道驱动的 M 个天线单元

下文中,这两个要求简称为设计要求 1 和 2。

假设多径信道由平坦瑞利衰落描述。由 8.7 节可知,只要阵列天线单元之间的间隔大到足以提供独立衰落(例如为波长的 7 倍),分集的采用就使得矢量 $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_N$ 线性独立。为简化说明,假设用户 1 是被观察用户,而其余 $N - 1$ 个用户产生共信道干扰,如图 8.29 所示。设计的关键是如何得到加权矢量 \mathbf{w} , \mathbf{w} 对应于阵列信号处理器。 \mathbf{w} 可按如下方法得到:

1. 选择 M 维加权矢量 \mathbf{w} 与对应于干扰用户的矢量 $\mathbf{c}_2, \dots, \mathbf{c}_N$ 正交。这个选择满足了设计要求 1(即共信道干扰的去除)。
2. 为了满足设计要求 2(即最大化信噪比),引入子空间的概念。假设矢量空间(或空间)由一组线性独立的矢量构成,此空间的子空间满足下面两个条件:^[14]
 - (i) 子空间中的两个矢量 \mathbf{z}_1 与 \mathbf{z}_2 的和也在子空间中。
 - (ii) 子空间中的矢量 \mathbf{z} 乘以任一标量 a ,乘积 $a\mathbf{z}$ 仍在子空间中。

返回到如何最大化用户 1 的输出信噪比的问题。我们首先构造一个子空间 \mathcal{W} ,其维数等于天线元素与干扰用户数目之间的差值,即 $M - (N - 1) = M - N + 1$ 。接着,将信道矢量 \mathbf{c}_1 (对应于用户 1)的共轭复数映射到子空间 \mathcal{W} 。此映射定义了加权矢量 \mathbf{w} 。

例 8.3

为了说明确定加权矢量 \mathbf{w} 的两个步骤,假设有一个包含两个用户的系统。这两个用户由

信道矢量 c_1 和 c_2 表征, 天线阵列由三个元素构成, 即 $N = 2, M = 3$ 。本例中, 子空间 W 为二维的, 即

$$M - N + 1 = 3 - 2 + 1 = 2$$

将用户 1 作为观察对象, 则用户 2 为干扰者, 可构造如图 8.30 所示的信号空间。子空间 W (图中的阴影部分) 正交于信道矢量 c_2 。阵列信号处理器的加权矢量 w 由用户 1 的共轭复信道矢量的投影决定, 即将 c_1^* 投影到子空间 W , 如图 8.30 所示。

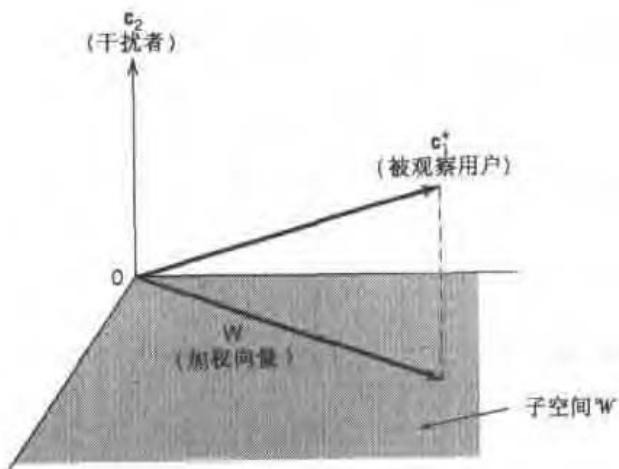


图 8.30 例 8.3 对应的信号空间图, 包含被观察用户、单个干扰者和具有三个元素的天线阵列。子空间 W (如阴影部分所示) 在本例中为二维的

由此可得一个重要的结论, 采用 M 个天线单元的最优合成且包含 $N - 1$ 个干扰用户的线性接收机, 与采用 $M - N + 1$ 个天线单元的无干扰线性接收机具有相同的性能, 这与多径环境无关。为了实现这个等价性, 要求 $M > N - 1$ 。假设这个条件符合, 接收机将分集提高到 $M - N + 1$, 就能够消除共信道干扰, 这代表着系统容量增加了 N 倍。

对应于这个二级子空间过程的阵列信号处理器应当是迫零的。这是因为, 给定 M 个天线单元, 只要 $M > N - 1$, 阵列就具有足够的自由度将由线性独立信道矢量 c_2, \dots, c_M 表示的 $N - 1$ 个干扰用户产生的输出强迫为零。值得注意的是, 此过程也包括了 $N = 1$ 的特殊情况(即只有单用户, 没有干扰用户)。在这种情况下, 信道矩阵为矢量 c_1 , 且存在于子空间 W 中, 迫零解 w 等于 c_1^* 。

上述分析都是在空间的概念中进行的, 没有考虑时延扩展的影响。如果时延扩展与信号持续时间相比很重要, 从而不可忽略呢? 时延扩展会造成 ISI。根据第 4 章中关于电话信道均衡的内容, 在每个天线分支中加入一个线性均衡器就能够补偿时延扩展。改进后的阵列信号处理器如图 8.31 所示。该阵列信号处理器结合了时间和空间的处理过程。空间处理由天线阵列进行, 时间处理由 FIR 滤波器组提供。因此, 这种结构称为空-时处理器。^[15]

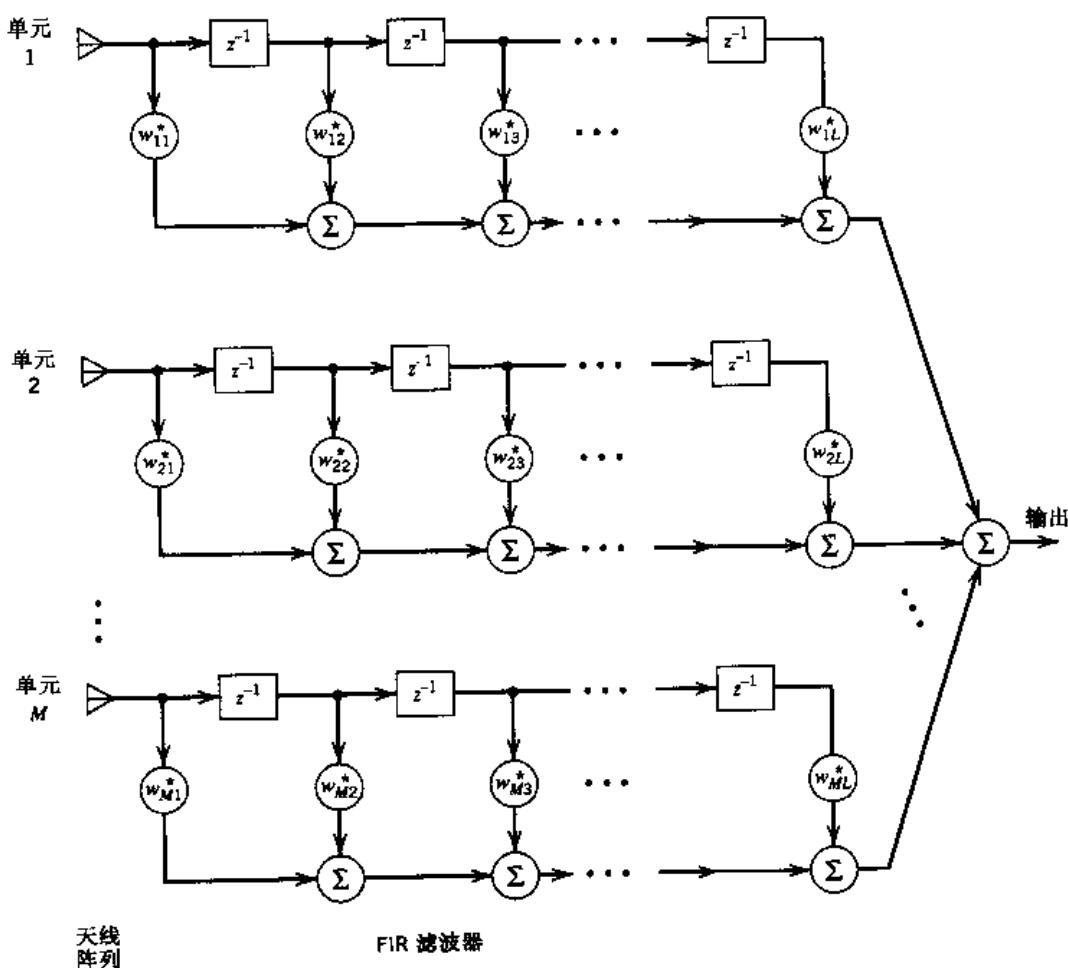


图 8.31 基带空-时处理器。标有 z^{-1} 字样的方框为单位延迟单元，每个延迟单元的时延都等于符号周期，滤波器的系数为复数形式的。假设所有FIR滤波器的长度都是 L

8.10.1 自适应天线阵列

在图 8.29 所示阵列信号处理器的子空间设计过程中，假设信道损耗为稳态的，从而得到信道矩阵 \mathbf{C} 。但在实际情况中多径衰落、时延扩展和共信道干扰都以各自的方式处于非稳态，且信道特性可能是未知的。为了处理这些实际的情况，就需要使图 8.29 所示的阵列信号处理器具有自适应性。本书仅讨论自适应空间处理。假设时延扩展是可忽略的。此外，还假设多径衰落现象足够缓慢，以便可将 LMS 算法用于此自适应过程。

图 8.32 所示为自适应天线阵列的结构，每个天线单元的输出乘以一个可调节的(可控制的)加权值，将这些加权单元的输出相加，就得到阵列输出信号。只要能提供与有用信号相关的参考信号，自适应天线阵列就不需要知道被观察用户发出的有用信号的到达方向。将参考信号减去阵列输出信号，就得到误差信号，误差信号用来对阵列单元加权值进行适当调整。通过这种方法，可将控制加权值的反馈系统加入到天线阵列的操作中，从而使天线阵列对环境变化具有自适应性。值得注意的是，图 8.32 中的框图是基于基带处理的，因此存在单元加权值的共轭复数。在实际的系统中，一般采用的方法是对每个天线单元使用正交混合，将各单元的复值接收信号分离为两个部分：实数和虚数部分。图 8.32 省略了正交混合的过程，以简化框图。

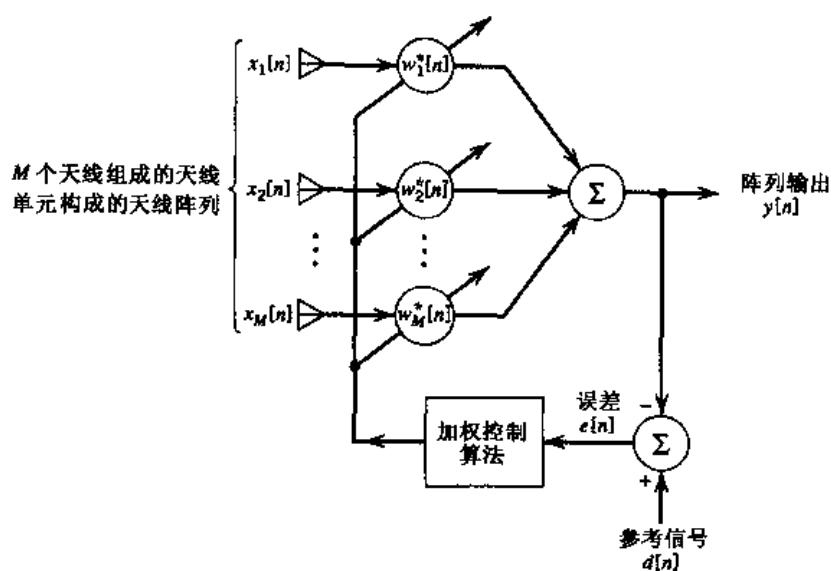


图 8.32 自适应天线阵列框图

为了最优化自适应天线阵列的性能,通常采用均方误差

$$J = E[|e[n]|^2] \quad (8.73)$$

作为用来进行最小化的代价函数。 $e[n]$ 为时刻 $t = nT$ 的误差信号,其中 T 为符号周期, n 为代表离散时间的整数。代价函数 J 的最小化抑制了阵列输出中的干扰信号,并加强了所需信号。但是,LMS 准则使代价函数 J 的瞬时值最小化,并通过连续迭代达到最小均方误差(MMSE)(即单元加权值的最佳解)。将第 4 章提到的时间均衡器扩展到空间域,可认为基于 MMSE 的自适应天线阵列与基于迫零标准,且以两步骤子空间方法实现的自适应天线阵列相比,提供了更好的解决方案。

用 $x_k[n]$ 表示离散时刻 n 的第 k 个天线单元的输出, $w_k[n]$ 表示此单元的相应加权值。天线阵列(由 M 个天线单元构成)的输出信号为

$$y[n] = \sum_{k=1}^M w_k^*[n]x_k[n] \quad (8.74)$$

其中, $w_k^*[n]x_k[n]$ 为复值矢量 $w_k[n]$ 与 $x_k[n]$ 的内积。以 $d[n]$ 表示参考信号,可得到误差信号为

$$e[n] = d[n] - y[n] \quad (8.75)$$

所以,对第 k 个单元加权值的修正为

$$\Delta w_k[n] = \mu e^*[n]x_k[n], \quad k = 1, 2, \dots, M \quad (8.76)$$

其中, μ 为步长参数,则加权值的更新值为

$$w_k[n+1] = w_k[n] + \Delta w_k[n], \quad k = 1, 2, \dots, M \quad (8.77)$$

式(8.74)~式(8.77)构成了复数 LMS 算法。实数信号的 LMS 算法(在第 3 章和第 4 章曾讨论过)可视为其中一个特殊情况。对于所有 k ,设 $w_k[0] = 0$ 以初始化算法。复数 LMS 算法的推导见习题 8.19。

使用复数 LMS 算法的自适应天线阵列的优点有：

- 实现简单
- 复杂度随天线单元数线性增长
- 抗干扰性能强

但是，上述系统也存在如下缺陷：

- 收敛速度缓慢，通常为权值的 10 倍。这限制了复数 LMS 算法在慢衰落环境中的应用，因为多普勒频移相对于观察间隔时间的倒数来说较小。
- 对参考信号和共信道干扰功率变化收敛性能的敏感性。

复数 LMS 算法的这些限制，通过采用直接矩阵倒置(DMI)算法就能够得到解决。DMI 算法对应于第 4 章讨论的 Wiener 滤波器(见习题 8.21)。不同于 LMS 算法，DMI 算法以分批方式进行操作，单元加权值的计算基于一批 K 个抽样点。分批的大小 K 是在两个矛盾要求之间进行折中得到的：

- K 应该足够小，以使得计算中的一批抽样点可认为是伪稳态的。
- K 应该足够大，以使单元加权的计算值接近于 MMSE 的解。

DMI 算法是当前应用于许多基站阵列天线的最优合并技术。如果需要的话，DMI 算法可由迭代计算^[16]重新表示。

对于电话通信很繁忙的情况，基站一般构造为微蜂窝，如办公室地板或通过定向天线沿高速公路延伸的工作站。在这种构造方式中采用的是相互之间很接近的多个造价相对便宜的基站。自适应天线阵列提供了一种替代结构，其中采用较少(但较昂贵)的基站，基站相互之间的间隔大于相应的微蜂窝系统。

8.11 总结与讨论

本章讨论了多用户通信的两个重要类型：卫星通信和无线通信。卫星通信系统提供了全球范围的覆盖，而无线通信系统则提供了移动性。由这两个通信系统提供的全球覆盖和移动性深刻改变了我们的通信方式，包括本地和全球的。

虽然卫星通信与无线通信系统以完全不同的方式工作，但它们都依靠无线电传播来连接接收机和发射机。在卫星通信中，存在从地面终端至卫星转发器的上行链路，以及从卫星到另一个地面终端的下行链路。卫星的作用如同一个中继器。此外，由于卫星位于对地同步轨道，因此上行链路和下行链路均工作在固定长度的可视路径上。卫星通信信道可近似地视为 AWGN 信道。

无线通信系统中同样存在两个链路：一个上行链路或反向链路，用于移动台至基站的传输；一个下行链路或前向链路，用于基站至移动台的传输。基站为固定的，位于覆盖区域的中心或边缘，由无线电信道、发射机和安置在塔上的接收天线构成。无线通信衰落的三个主要来源为共信道干扰、衰落以及时延扩展，其中后两种为多径的副产品。这些信道损耗的共同特性是它们都是信号独立现象。不同于普遍存在的信道噪声，干扰和多径的衰落效应无法通过简单地增加发射信号功率来解决。干扰和多径都需要采用特殊技术来进行处理。这些特殊技术

包括分集、自适应天线阵列以及 RAKE 接收机。

下面以对无线通信与有线通信的比较来结束讨论。由第 3 章可知, 有线通信系统中最关心的问题是噪声。有线通信系统具有足够的信道带宽, 因此允许采用 PCM 作为将语音转换为 64 Kbps 信号流的标准方法。因此, 能够以几乎无噪声的性能提供基本数据。另一方面, 在无线通信中, 信道带宽是宝贵的资源, 为了节省这种资源, 需要采用有效的语音编码技术来提供速率远小于 PCM 的高质量数字化语音通信。遗憾的是, 以(第 3 章中讨论过的)ADPCM 为例的波形编码器不满足这样的要求。较好的解决方法是采用有效的信源编码技术, 如多脉冲激励 LPC 或其规则脉冲激励方式, 以及 CELP。这些信源编码技术通过去除语音中几乎所有的冗余, 以低于 16 Kbps 的比特率维持了高质量的合成语音。为了抵抗噪声, 采用了信道编码, 有控制地将冗余比特插入到发送数据流中。信道编码的采用同时还扩展了低功率手机的使用范围, 并延长了电池的使用寿命。信道编码将在第 10 章中讨论。

注释与参考

- [1] 卫星通信及相关内容的详细讨论见下列书籍:Sklar(1988), Pratt 和 Bostian(1986), Wu (1984), Bhargava 等人(1981)以及 Spilker, Jr. (1977)。上述书籍中, 除了第二本以外, 都总结了数字通信中卫星的使用。而 Pratt 和 Bostian 的书则较广泛地介绍了卫星通信, 强调了无线电波传播、天线、轨道力学、信号处理、无线电电子学等不同主题。
- [2] Sklar(1988)和 Anderson(1999)讨论了链路预算分析。而卫星通信在 Bhargava 等人(1981)的著作中进行了讨论。
- [3] 天线的基础知识见 Kraus(1950)以及 Jordan 和 Balmain(1973)所编著作的第 11 章。
- [4] 自由空间公式(式 8.14)是为了纪念 Friis(1946)而命名的。式(8.30)的 Friis 公式的原始形式见 Friis(1944)。
- [5] 蜂窝无线电通信的原始论述见 MacDonald(1979)。
- [6] 移动无线电传播信道的全面论述见 Parsons(1992)。该书介绍了 VHF 和 UHF 传播的基础知识, 它们是在不规则陆地和高楼林立地区的传播。书中还介绍了移动无线电信道的统计特性。移动无线电信道的统计特性在 Proakis(1995)中也有讨论。该书提供了对瑞利衰落信道中衰落对差错性能的影响的可靠论述以及分集技术的讨论。关于该主题的全面论述见 Schwartz, Bennett, Stein(1966)所编著作的第 9~11 章。
- [7] χ^2 分布为 Γ 分布的一个特例。 Γ 分布随机变量 X 的概率密度函数有两个参数: $\alpha > 0, \lambda > 0$, 其中

$$f_X(x) = \frac{\lambda(\lambda x)^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, \quad 0 < x < \infty$$

这里, $\Gamma(\alpha)$ 为 Γ 函数, 其定义为

$$\Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z} dz, \quad \alpha > 0$$

Γ 函数有以下两个特性:

$$\begin{aligned}\Gamma(1/2) &= \sqrt{\pi} \\ \Gamma(\alpha + 1) &= \alpha\Gamma(\alpha), \quad \alpha > 0\end{aligned}$$

假设 $\lambda = 1/2$, $\alpha = k/2$, k 为正整数, 可得到具有 $2k$ 个自由度的 χ^2 分布

$$f_X(x) = \frac{x^{(k-2)/2} e^{-x/2}}{2^{k/2} \Gamma(k/2)}, \quad 0 < x < \infty$$

- [8] 无线通信发展过程的概述文章见 Oliphant(1999)。介绍无线通信系统基础知识的书有 Steele 和 Hanzo(1999), 以及 Stüber(1996) 和 Rappaport(1996)。
- 有关 GSM 的详细描述见 Steele 和 Hanzo(1999) 所编著作的第 8 章。关于 IS-95 的详细描述见 Lee 和 Miller(1998) 的手写本。
- [9] RAKE 接收机的经典论文源自 Price 和 Green(1958)。
- [10] 关于如何根据多个噪声信号的和来进行信噪比最大化的原始论文见 Brennan(1955) 的经典文章。
- [11] 关于 CDMA 无线通信系统中 RAKE 接收机的详细讨论见 Viterbi(1995)。
- [12] 语音编码的多脉冲激励的概念来源于 Atal 和 Remde(1982)。语音的 CELP 首先由 Atal 和 Schroeder(1984) 提出。语音编码的多脉冲激励、规则脉冲激励以及码激励类型的详细的数学讨论, 以及与无线通信相关的内容, 见 Steele 和 Hanzo(1999) 所编著作的第 3 章。
- [13] 在无线通信文献中, 自适应天线阵列通常称为智能天线。无线通信的自适应天线阵列的相关描述见 Winters(1998) 的文章和 Winters(1999) 的实验笔记。设计图 8.29 中的阵列信号处理器的二级子空间程序基于 Winters(1999) 提供的资料。Rappaport(1999) 收集了有关自适应天线阵列的论文, 并按照算法、体系结构、硬件使用、信道模型和性能估计进行了分类。
- [14] 子空间的概念是基于矩阵代数学的。有关的讨论见 Strang(1980) 和 Stewart(1973)。统计信号处理的子空间分解的讨论见 Scharf(1991)。
- [15] 无线通信的空 - 时处理的指示性讨论见 Paulraj 和 Ng(1998), Paulraj 和 Papadias(1997) 以及 Kohno(1998)。
- [16] DMI 算法的迭代使用引出了一个新的算法, 称为迭代最小平方(RLS)算法。RLS 算法的推导及其参数见 Haykin(1996)。

习题

自由空间传播

8.1 一无线电链路采用一对各自效率均为 60% 的 2 m 碟型天线, 分别作为发射和接收天线。链路的其他参数为:

发射功率	= 1 dBW
载波频率	= 4 GHz
接收机与发射机之间的距离	= 150 m

试计算:

(a) 自由空间损耗

(b) 每个天线的功率增益

(c) 单位为 dBW 的接收功率

8.2 载波频率为 12 GHz 时, 重复习题 8.1 的计算。

8.3 式(8.14)为 Friis 自由空间公式的一种形式。证明此公式也可由以下的等价形式表示:

$$(a) P_r = \frac{P_t A_t A_r}{\lambda^2 d^2}$$

$$(b) P_r = \frac{P_t A_t G_r}{4\pi d^2}$$

其中, P_t 为发射功率, A_t 为发射天线的有效面积, λ 为载波波长, d 为接收机到发射机之间的距离, G_r 为接收天线的功率增益, A_r 为接收天线的有效面积, P_r 为接收功率。

讨论这些公式分别更适用于何种情况。

8.4 由自由空间损耗的数学定义

$$L_{\text{free space}} = \left(\frac{4\pi d}{\lambda} \right)^2$$

可知, 它由载波波长 λ 或频率 f 决定。试从物理意义上证明这种波长或频率的依赖性是正确的。

8.5 在卫星通信系统中, 上行链路的载波频率一般高于下行链路。说明该选择的合理性。

8.6 连续波信标发射机位于静地同步轨道卫星上。此发射机的 12 GHz 输出被 40 000 km 以外的地面站监控。卫星发射天线为 1 m 长的碟型天线, 其孔径效率为 70%, 地面接收天线为 10 m 长的碟型天线, 孔径效率为 55%。假设发射机输出功率为 100 mW, 计算接收功率。

噪声系数

8.7 考虑一个维持在“室温”(290 K)的 75Ω 电阻。假设带宽为 1 MHz, 计算:

(a) 终端与电阻之间的由温度噪声引起的电压的均方根(RMS)值。

(b) 传递到匹配负载的最大有用噪声功率。

8.8 回顾例 8.1, 该例基于图 8.10 的接收机配置。假设有损耗的波导被插入到接收天线与低噪声放大器之间。波导损耗为 1 dB, 其物理温度为 290 K。重新计算接收机的有效噪声温度。

8.9 图 P8.9 的接收机由有损耗波导、低噪声射频放大器、下变频器(混合器)以及中频放大器构成。图中包括了这四个部分的噪声系数和功率增益。天线温度为 50 K。

(a) 计算图 P8.9 中四个部分的等效噪声温度, 假设室温 $T = 290$ K。

(b) 计算整个接收机的有效噪声温度。

链路预算计算

8.10 考虑例 8.2 中的数字通信系统的上行链路功率预算。链路参数如下:

- 载波频率 = 14 GHz
 饱和状态下的 TWT 放大器的功率密度 = -81 dBW/m²
 卫星品质因数 G/T = 1.9 dB/K
 地面发射终端与卫星之间的距离 = 40 000 km
 (a) 假设 TWT 无功率损耗, 计算卫星的 C/N_0 比值。
 (b) 假设上行链路的数据速率和例 8.2 中下行链路相同, 计算链路容限为 6 dB 的上行链路的符号差错概率。并与例 8.2 中的结果做比较。

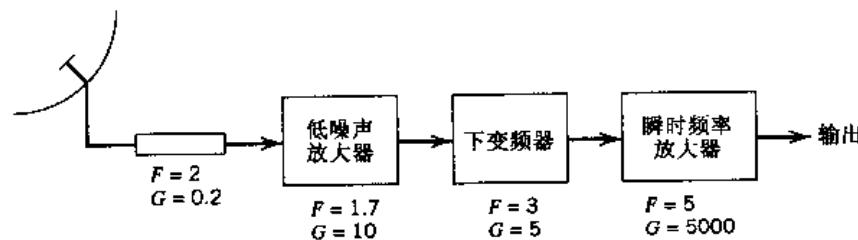


图 P8.9

8.11 直接广播卫星(DBS)系统的下行链路 C/N_0 为 85 dB·Hz。链路的参数为:

- 卫星 EIRP = 57 dBW
 下行链路载波频率 = 12.5 GHz
 数据速率 = 10 Mbps
 地面接收终端所需的 C/N_0 = 10 dB
 地面接收终端与卫星之间的距离 = 41 000 km
 假设碟型天线的效率为 55%, 且所处位置周围的温度为 310 K, 计算碟型天线为提供满意的电视接收所需的最小直径。假设 DBS 系统的操作本质上为下行链路受限的。

无线通信

8.12 无线通信和卫星通信的工作都依赖于无线电传播。试总结:

- (a) 这两个多用户通信系统的相似之处
 (b) 区分两者的主要差别

8.13 在无线通信系统中, 上行链路(反向链路)的载波频率小于下行链路(前向链路)。证明这样选择的正确性。

8.14 图 P8.14 显示了空地上的无线电链路的直接(可视)和非直接(反射)路径。基站的发射天线高度和移动单元的接收天线高度分别为 h_b 和 h_m 。假设有:

- 地面的反射系数为 -1。
- 两天线间的距离 d 足够长, 使得反射和直接路径之间的相位差 ϕ 小于 1 弧度, 因此可认为 $\sin\phi \approx \phi$ 。

证明接收功率 P_r 可近似表示为

$$P_r \approx P_t G_b G_m \left(\frac{h_b h_m}{d^2} \right)^2$$

其中, P_t 为发射功率, G_b 和 G_m 分别为发射基站和移动天线的功率增益。并将此结

果与 Friis 自由空间公式比较。

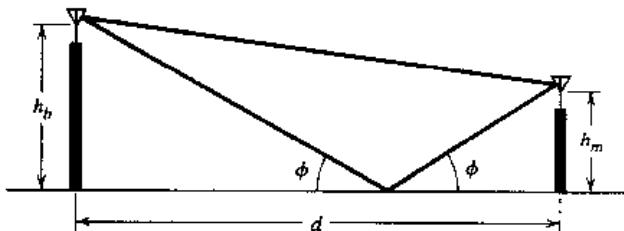


图 P8.14

8.15 由如下冲激响应定义的双路模型：

$$h(t) = a_1 \delta(t - \tau_1) + a_2 \exp(-j\theta) \delta(t - \tau_2)$$

经常被用于无线通信系统的分析中。模型参数有：时延 \$\tau_1\$ 和 \$\tau_2\$，均匀分布的相位 \$\theta\$，以及实系数 \$a_1\$ 和 \$a_2\$。

- (a) 求(i)模型的转移函数；(ii)功率时延函数。
- (b) 证明模型具有频率选择衰落，是由于系数 \$a_1\$ 和 \$a_2\$ 的变化

8.16 在图 8.26 所示的 RAKE 接收机中，通过向接收信号插入适当的时延，使每个相关器都同步的。

- (a) 证明在参考信号(即 PN 序列)中插入适当的时延，在理论上能产生同样的结果
- (b) 实际情况中，较好的方法是使用图 8.26 所示的程序，试给出相应的理由。

8.17 本题研究最大比分集合成机制。考虑一组噪声信号 \$\{x_j(t)\}_{j=1}^N\$，其中，\$x_j(t)\$ 的定义为

$$x_j(t) = s_j(t) + n_j(t), \quad j = 1, 2, \dots, N$$

假设：

- 信号分量 \$s_j(t)\$ 是本地相关的，即

$$s_j(t) = z_j m(t), \quad j = 1, 2, \dots, N$$

其中，\$z_j\$ 为正实数，\$m(t)\$ 表示单位功率的消息信号。

- 噪声分量 \$n_j(t)\$ 具有零均值且统计独立，即

$$E[n_j(t)n_k(t)] = \begin{cases} \sigma_j^2 & k = j \\ 0 & \text{其他} \end{cases}$$

线性合并器的输出定义为

$$x(t) = \sum_{j=1}^N \alpha_j x_j(t)$$

其中，\$\alpha_j\$ 为需要确定的参量。

- (a) 证明输出信噪比为

$$(\text{SNR})_O = \frac{\left(\sum_{j=1}^N \alpha_j z_j \right)^2}{\sum_{j=1}^N \alpha_j^2 \sigma_j^2}$$

(b) 设

$$u_i = \alpha_i \sigma_i$$

$$v_i = \frac{z_i}{\sigma_i}$$

并重新表示 $(\text{SNR})_o$ 。将施瓦茨不等式应用到此新定义中, 证明:

$$(i) \quad (\text{SNR})_o \leq \sum_{i=1}^N (\text{SNR})_i$$

其中 $(\text{SNR})_j = z_j^2 / \sigma_j^2$

(ii) 合并器系数的最优值定义为

$$\alpha_i = \frac{z_i}{\sigma_i^2}$$

此时, 施瓦茨不等式的等号成立。

施瓦茨不等式在 5.2 节中曾有讨论。

自适应天线阵列

8.19 在图 8.29 的阵列信号处理器中, 只有两个用户 ($N = 2$), 天线阵列有两个单元 ($M = 2$)。构造子空间 \mathcal{W} , 并利用信号空间图来说明表征阵列信号处理器的加权值的计算过程。

8.19 本题推导复数 LMS 算法。参考图 8.32, 从瞬时代价函数开始:

$$J = \frac{1}{2} |e[n]|^2$$

其中, $e[n]$ 为误差信号, M 为天线单元的个数, 推导按如下步骤进行:

(a) 求代价函数 J 关于第 k 个单元的加权值 $w_k[n]$ 的导数。

(b) 使用瞬时导数 $\partial J / \partial w_k[n]$, 并将其记为 $\hat{\nabla}J[k]$, 按如下规则, 求第 k 个单元加权值的调整量 $\Delta w_k[n]$:

$$\Delta w_k[n] = -\mu \hat{\nabla}J[k]$$

(c) 证明式(8.75)~式(8.77)所示复数 LMS 算法的合成公式。

需注意的是, $w_k[n]$ 具有复数值, 因此需要分别考虑实数和虚数部分。

8.20 采用 LMS 算法的自适应天线阵列存在一个实际限制, 就是天线阵列工作的动态范围。这个限制是由于 LMS 算法的加权值的反应速度正比于天线阵列输入的平均信号功率而产生的。

(a) 证明天线阵列输入的平均信号功率的动态范围正比于 R_b/f_{\max} , 其中, R_b 为以 bps 表示的数据速率, f_{\max} 为以 Hz 表示的最大衰落速率。

(b) 假设比例因子为 0.2(用于 R_b/f_{\max} 中)。当 $R_b = 32$ Kbps, $f_{\max} = 70$ Hz 时, 计算使用 LMS 算法的自适应天线阵列的动态范围, 并给出结论。(比例因子 0.2 是采

用 PSK 系统时较为合理的选择)。

8.21 本题将推导用于调整自适应天线阵列加权值的直接矩阵转置算法。为此,先回顾一下第 3 章提到的 Wiener 滤波器。

(a) 证明

$$\hat{\mathbf{R}}_{\mathbf{x}} \mathbf{w} = \hat{\mathbf{r}}_{\mathbf{x}d}$$

其中, $\hat{\mathbf{R}}_{\mathbf{x}}$ 为输入矢量 $\mathbf{x}[k]$ 的相关矩阵的一个估计:

$$\hat{\mathbf{R}}_{\mathbf{x}} = \frac{1}{K} \sum_{k=1}^K \mathbf{x}[k] \mathbf{x}^H[k]$$

$\hat{\mathbf{r}}_{\mathbf{x}d}$ 为 $\mathbf{x}[k]$ 与参考信号 $d[k]$ 之间的互相关矢量的一个估计:

$$\hat{\mathbf{r}}_{\mathbf{x}d} = \frac{1}{K} \sum_{k=1}^K \mathbf{x}[k] d^*[k]$$

$\hat{\mathbf{R}}_{\mathbf{x}}$ 的公式中上标 H 表示埃尔米特共轭变换(即转置和共轭复数),因此, $\mathbf{x}[k]$ $\mathbf{x}^H[k]$ 表示 $\mathbf{x}[k]$ 与其自身的内积。在全部 K 个抽样点上对 $\hat{\mathbf{R}}_{\mathbf{x}}$ 和 $\hat{\mathbf{r}}_{\mathbf{x}d}$ 求和,其中每个抽样点由 $\{\mathbf{x}[k], d[k]\}$ 表示。

(b) 利用(a)中的公式,描述计算加权矢量 \mathbf{w} 的算法,假设数据集由 K 个抽样点组成。证明当加权矢量 \mathbf{w} 大小用 M 表示时,算法复杂度以 M^3 增长。

第9章 信息论基础

香农(1948)关于信息论的划时代论述以及其他研究人员对该理论的完善,是电子工程师设计有效而可靠的通信系统的基础。从信源到信宿的有效通信通过信源编码获得,噪声信道上的可靠通信通过差错控制编码获得。本章主要讨论以下几个重要问题:

- 作为信息基本度量的熵。
- 信源编码定理和无失真数据压缩算法。
- 互信息及其与信息传输的通信信道容量的关系。
- 作为可靠通信基础的信道编码定理。
- 作为信道带宽与信噪比折中基础的信息容量定理。
- 符合保真度标准的信源编码率失真理论。

9.1 引言

如“背景与预览”中所述,通信系统的目标就是将承载信息的基带信号通过通信信道从一个地方传输到另一个地方。前面的章节中描述了很多用来实现这个目标的调制方法。但是,何谓信息?为了说明这个概念,就需要用到信息论。^[1]这门基于数学的学科不仅对通信、计算机科学、统计物理、统计推论以及概率与统计都有很大的贡献。

对通信而言,信息论仅对通信系统进行数学模型化和分析,而不是针对物理信源和物理信道。特别地,信息论提供了以下两个基本问题的答案:

- 压缩信号所需的最低复杂度至少为多少?
- 噪声信道上的可靠通信的最终传输速率是多少?

这些问题的答案分别是信源的熵和信道的容量。熵是根据信源的统计特性来定义的,这个名称来源于热力学。信道容量的定义为信道传输信息的能力。信道容量与信道的噪声特性密切相关。由信息论可得到一个重要结论,即如果信源的熵小于信道容量,则信道上的无差错通信就可能实现。下面就以讨论不确定性、信息和熵之间的关系来开始对信息论的学习。

9.2 不确定性、信息和熵

假设概率实验包括每个单位时刻(信号间隔时间)对离散信源输出的观察。信源输出被调制为离散随机变量 S 。 S 从固定的有限字符集

$$\mathcal{S} = \{s_0, s_1, \dots, s_{K-1}\} \quad (9.1)$$

中选取符号的概率为

$$P(S = s_k) = p_k, \quad k = 0, 1, \dots, K - 1 \quad (9.2)$$

当然,这组概率值必须满足如下的条件:

$$\sum_{k=0}^{K-1} p_k = 1 \quad (9.3)$$

假设相邻信号间隔内,信源输出的符号是统计独立的。具有以上特性的信源称为离散无记忆信源,无记忆的意思是,任何时刻的输出符号都与前面的选择无关。

能否找到对信源信息的度量?为了回答这个问题,首先应注意到,信息的概念是与不确定性或意外紧密关联的,下面对此进行讨论。

设事件 $S = s_k$,信源输出为 s_k 的概率为 p_k ,如式(9.2)所定义。显然,如果 $p_k = 1$,对所有 $i \neq k$, $p_i = 0$,则当发射符号 s_k 时,不存在“意外”,也就不存在“信息”,因为我们知道信源发出的符号必然是什么。如果信源发出的符号具有不同的概率,且概率 p_k 较低,则信源发出信号 s_k 时存在的意外以及信息的量,要多于发出具有较高概率的符号 s_i , $i \neq k$ 时的相应值。这样,不确定性、意外和信息都被联系起来。在事件 $S = s_k$ 发生之前,具有许多的不确定性。在事件 $S = s_k$ 发生的时候,会有许多意外存在。在事件 $S = s_k$ 发生以后,信息量增加,其实质可看成是不确定性的分解。信息量与事件发生概率的倒数相关。

以对数函数^[2]定义事件 $S = s_k$ 以概率 p_k 发生后的信息增量,即

$$I(s_k) = \log\left(\frac{1}{p_k}\right) \quad (9.4)$$

这个定义直观上应满足以下重要特性:

$$1. \quad I(s_k) = 0 \quad p_k = 1 \quad (9.5)$$

显然,如果在事件发生之前,就可以绝对地确定其结果,则不存在信息增量。

$$2. \quad I(s_k) \geq 0 \quad 0 < p_k \leq 1 \quad (9.6)$$

也就是说,事件 $S = s_k$ 的发生会提供一些(或零)信息,但决不会造成信息的丢失。

$$3. \quad I(s_k) > I(s_i) \quad p_k < p_i \quad (9.7)$$

也就是说,一个事件发生的概率越小,当它发生时我们得到的信息就越多。

$$4. \text{ 如果 } s_l \text{ 和 } s_k \text{ 统计独立, 则 } I(s_k s_l) = I(s_k) + I(s_l)。$$

式(9.4)中对数的底是任意的。但是,现在的标准方法是采用以 2 为底的对数。得到的信息单位相应地称为比特(binary digit 的缩写)。即

$$\begin{aligned} I(s_k) &= \log_2\left(\frac{1}{p_k}\right) \\ &= -\log_2 p_k \quad k = 0, 1, \dots, K-1 \end{aligned} \quad (9.8)$$

当 $p_k = 1/2$ 时,有 $I(s_k) = 1$ 比特。因此,1 比特是当两个等可能(即等概率)事件之一发生时获得的信息增量。注意,信息 $I(s_k)$ 是正的,这是因为一个小于 1 的数字(如概率)的对数值是负的。

在任意符号间隔上,信源所提供的信息量 $I(s_k)$ 取决于当前时刻信源发送的符号 s_k 。实际上, $I(s_k)$ 是一个离散随机变量,分别以概率 p_0, p_1, \dots, p_{K-1} 取值 $I(s_0), I(s_1), \dots, I(s_{K-1})$ 。在信源字符集 \mathcal{S} 上, $I(s_k)$ 的平均值为

$$\begin{aligned}
 H(\mathcal{S}) &= E[I(s_k)] \\
 &= \sum_{k=0}^{K-1} p_k I(s_k) \\
 &= \sum_{k=0}^{K-1} p_k \log_2 \left(\frac{1}{p_k} \right)
 \end{aligned} \tag{9.9}$$

$H(\mathcal{S})$ 是一个重要的量, 称为具有信源字符集 \mathcal{S} 的离散无记忆信源的熵^[3]。熵是每个信源符号包含的平均信息的度量。注意熵 $H(\mathcal{S})$ 只取决于信源字符集 \mathcal{S} 中符号的发生概率。所以, $H(\mathcal{S})$ 中的符号 \mathcal{S} 不是一个函数变量, 而是信源的一个标识。

9.2.1 熵的一些性质

考虑一个离散无记忆信源, 其数学模型如式(9.1)和式(9.2)所定义。此信源的熵 $H(\mathcal{S})$ 取值有如下的限制:

$$0 \leq H(\mathcal{S}) \leq \log_2 K \tag{9.10}$$

其中, K 为信源字符集 \mathcal{S} 的基数(符号个数)。我们可以给出如下的结论:

1. 当且仅当某些 k 值对应的概率 $p_k = 1$ 而其他概率均为 0 时, $H(\mathcal{S}) = 0$ 。此下限对应着无不确定性。
2. 当且仅当对于所有 k 都有 $p_k = 1/K$ (即字符集 \mathcal{S} 中的符号都是等概率的)时, $H(\mathcal{S}) = \log_2 K$ 。此上限对应着最大不确定性。

下面证明 $H(\mathcal{S})$ 的这些特性。首先, 由于每个概率 p_k 不是小于 1 就是等于 1, 则式(9.9)中的 $p_k \log_2 (1/p_k)$ 是非负的, 因此 $H(\mathcal{S}) \geq 0$ 。此外, 当且仅当 $p_k = 0$ 或 1 时, $p_k \log_2 (1/p_k)$ 等于 0。从而可推导出, 当且仅当 $p_k = 0$ 或 1, 即某些 k 值对应的概率 $p_k = 1$, 其他概率均为 0 时, $H(\mathcal{S}) = 0$ 。

这就完成了对式(9.10)中的下限和结论 1 的证明。

为了证明式(9.10)中的上限和结论 2, 需用到自然对数的一个特性:

$$\log x \leq x - 1, \quad x \geq 0 \tag{9.11}$$

这个不等式可通过函数 $\log x$ 和 $(x - 1)$ 关于 x 的曲线来证明, 如图 9.1 所示。图中, 直线 $y = x - 1$ 都在曲线 $y = \log x$ 之下。两个函数只在点 $x = 1$ 取等号, 此时, 此直线与曲线相切。

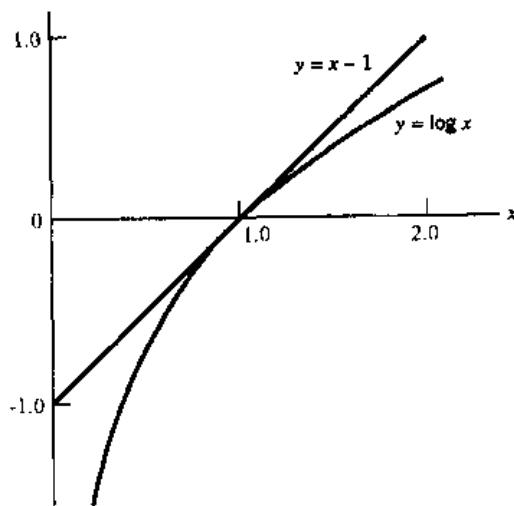


图 9.1 函数 $\log x$ 和 $x - 1$ 对 x 的变化曲线

下面考虑离散无记忆信源的字符集 $\mathcal{S} = \{s_0, s_1, \dots, s_{K-1}\}$ 上的两个概率分布 $\{p_0, p_1, \dots, p_{K-1}\}$ 和 $\{q_0, q_1, \dots, q_{K-1}\}$ 。将其转换为自然对数形式，即

$$\sum_{k=0}^{K-1} p_k \log_2 \left(\frac{q_k}{p_k} \right) = \frac{1}{\log 2} \sum_{k=0}^{K-1} p_k \log \left(\frac{q_k}{p_k} \right)$$

由式(9.11)所示的不等式，可得

$$\begin{aligned} \sum_{k=0}^{K-1} p_k \log_2 \left(\frac{q_k}{p_k} \right) &\leq \frac{1}{\log 2} \sum_{k=0}^{K-1} p_k \left(\frac{q_k}{p_k} - 1 \right) \\ &\leq \frac{1}{\log 2} \sum_{k=0}^{K-1} (q_k - p_k) \\ &\leq \frac{1}{\log 2} \left(\sum_{k=0}^{K-1} q_k - \sum_{k=0}^{K-1} p_k \right) = 0 \end{aligned}$$

从而得到基本不等式

$$\sum_{k=0}^{K-1} p_k \log_2 \left(\frac{q_k}{p_k} \right) \leq 0 \quad (9.12)$$

当且仅当对于所有 k 都有 $q_k = p_k$ 时等号成立。

假设

$$q_k = \frac{1}{K}, \quad k = 0, 1, \dots, K-1 \quad (9.13)$$

上式对应于字符集 \mathcal{S} 中的符号具有相等概率的情况。具有此特性的离散无记忆信源的熵为

$$\sum_{k=0}^{K-1} q_k \log_2 \left(\frac{1}{q_k} \right) = \log_2 K \quad (9.14)$$

将式(9.13)代入式(9.12)，得到

$$\sum_{k=0}^{K-1} p_k \log_2 \left(\frac{1}{p_k} \right) \leq \log_2 K$$

等价地，字符集 \mathcal{S} 中的符号具有任意概率分布时的离散无记忆信源的熵的上限为

$$H(\mathcal{S}) \leq \log_2 K$$

当且仅当字符集 \mathcal{S} 中的符号具有相等概率时取等号，如式(9.13)所示。这就完成了对式(9.10)、结论1和结论2的证明。

例 9.1 二进制无记忆信源的熵

为了说明 $H(\mathcal{S})$ 的特性，考虑一个二进制信源，其中符号 0 以概率 p_0 发生，符号 1 以概率 $p_1 = 1 - p_0$ 发生。假设信源是无记忆的，因此信源发出的相邻符号是统计独立的。

信源的熵为

$$\begin{aligned} H(\mathcal{S}) &= -p_0 \log_2 p_0 - p_1 \log_2 p_1 \\ &= -p_0 \log_2 p_0 - (1 - p_0) \log_2 (1 - p_0) \text{ 比特} \end{aligned} \quad (9.15)$$

因此得到如下结论：

1. 当 $p_0 = 0$ 时, 熵 $H(\mathcal{S}) = 0$ 。这是因为 $x \rightarrow 0$ 时, $x \log x \rightarrow 0$ 。
2. 当 $p_0 = 1$ 时, 熵 $H(\mathcal{S}) = 0$ 。
3. 当 $p_1 = p_0 = 1/2$, 即符号 1 和 0 等概出现时, 熵 $H(\mathcal{S})$ 达到最大值, $H_{\max} = 1$ 比特。

式(9.15)第二行给出的 p_0 的函数, 将会经常在信息论习题中遇到。因此定义

$$\mathcal{H}(p_0) = -p_0 \log_2 p_0 - (1 - p_0) \log_2 (1 - p_0) \quad (9.16)$$

称 $\mathcal{H}(p_0)$ 为熵函数。式(9.15)与式(9.16)之间的区别需要特别注意。式(9.15)中的 $H(\mathcal{S})$ 给出了具有信源字符集 \mathcal{S} 的离散无记忆信源的熵。而式(9.16)中的 $\mathcal{H}(p_0)$ 是(定义在区间 $[0, 1]$ 上的)先验概率 p_0 的函数。熵函数 $\mathcal{H}(p_0)$ 在区间 $[0, 1]$ 上对 p_0 的曲线如图 9.2 所示。图 9.2 的曲线证实了上述 3 个结论。

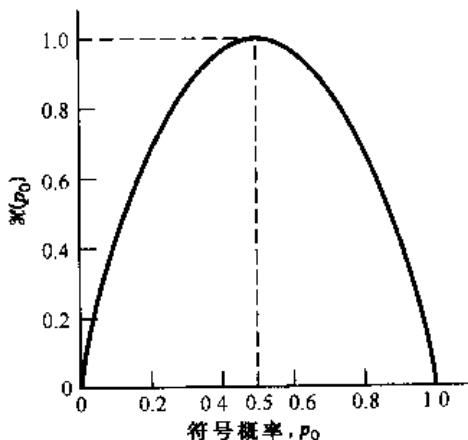


图 9.2 熵函数 $\mathcal{H}(p_0)$

9.2.2 离散无记忆信源的扩展

在讨论信息论时, 研究块往往比研究单个符号更加有用。每个块由 n 个连续的信源符号构成。可认为每个这样的块是由扩展信源提供的, 信源字符集 \mathcal{S}^n 中含有 K^n 个不同的块, K 为原始信源的字符集 \mathcal{S} 中不同符号的个数。在离散无记忆信源的情况下, 信源符号是统计独立的。因此, \mathcal{S}^n 中每个信源符号的概率等于 \mathcal{S} 中构成 \mathcal{S}^n 的 n 个信源符号的概率乘积。直观上可这样理解扩展: 信源的熵 $H(\mathcal{S}^n)$ 等于原始信源的熵 $H(\mathcal{S})$ 的 n 倍。即

$$H(\mathcal{S}^n) = nH(\mathcal{S}) \quad (9.17)$$

例 9.2 扩展信源的熵

考虑一个离散无记忆信源, 其字符集 $\mathcal{S} = \{s_0, s_1, s_2\}$ 分别具有概率

$$\begin{aligned} p_0 &= \frac{1}{4} \\ p_1 &= \frac{1}{4} \\ p_2 &= \frac{1}{2} \end{aligned}$$

应用式(9.9)得到信源的熵为

$$\begin{aligned}
 H(\mathcal{S}) &= p_0 \log_2\left(\frac{1}{p_0}\right) + p_1 \log_2\left(\frac{1}{p_1}\right) + p_2 \log_2\left(\frac{1}{p_2}\right) \\
 &= \frac{1}{4} \log_2(4) + \frac{1}{4} \log_2(4) + \frac{1}{2} \log_2(2) \\
 &= \frac{3}{2} \text{ 比特}
 \end{aligned}$$

接着考虑信源的二阶扩展。信源字符集 \mathcal{S} 由 3 个符号构成，则扩展信源的信源字符集 \mathcal{S}^2 中有 9 个符号。表 9.1 的第一行给出了 \mathcal{S}^2 的 9 个符号，分别表示为 $\sigma_0, \sigma_1, \dots, \sigma_8$ 。表的第二行给出了这 9 个符号以信源符号 s_0, s_1, s_2 的相应序列表示的形式，每次取两个信源符号。扩展信源中 9 个符号的概率在最后一行中给出。应用式(9.9)得到扩展信源的熵为

$$\begin{aligned}
 H(\mathcal{S}^2) &= \sum_{i=0}^8 p(\sigma_i) \log_2 \frac{1}{p(\sigma_i)} \\
 &= \frac{1}{16} \log_2(16) + \frac{1}{16} \log_2(16) + \frac{1}{8} \log_2(8) + \frac{1}{16} \log_2(16) \\
 &\quad + \frac{1}{16} \log_2(16) + \frac{1}{8} \log_2(8) + \frac{1}{8} \log_2(8) + \frac{1}{8} \log_2(8) + \frac{1}{4} \log_2(4) \\
 &= 3 \text{ 比特}
 \end{aligned}$$

不难注意到 $H(\mathcal{S}^2) = 2H(\mathcal{S})$ ，这是符合式(9.17)的。

表 9.1 离散无记忆信源的二阶扩展的字符集

\mathcal{S}^2 的符号	σ_0	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8
\mathcal{S} 的相应符号序列	$s_0 s_0$	$s_0 s_1$	$s_0 s_2$	$s_1 s_0$	$s_1 s_1$	$s_1 s_2$	$s_2 s_0$	$s_2 s_1$	$s_2 s_2$
概率 $p(\sigma_i), i=0,1,\dots,8$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$

9.3 信源编码定理

通信中的一个重要问题就是离散信源产生的数据的有效表示方法。这个可由信源编码来完成。完成信源编码的设备称为信源编码器。为了信源编码器的有效性，需要对信源统计特性有所了解。特别地，可通过为常出现的信源符号分配短码字，为不常出现的信源符号分配长码字，将信源的统计特性应用到信源编码的产生中。这种信源编码称为可变长度编码。莫尔斯(Morse)码就是可变长度编码的一种。在莫尔斯码中，字母表中的字母和数字都被编码为传号流和空格流，分别由点“.”和“-”表示。在英语中，字母 E 比字母 Q 更经常出现，因此莫尔斯码将 E 编码为单个点“.”(这是最短的码字)，而将 Q 编码为“- - . -”(这是最长的码字)。

我们的主要目的是开发一个有效的信源编码，该信源编码必须满足以下两个基本要求：

1. 编码器产生的码字是二进制形式的。
2. 信源编码具有唯一的译码，这样才能够从编码后的二进制序列中恢复出原始信源序列。

图9.3所示为一个离散无记忆信源,其输出 s_k 由信源编码器转换为0和1组成的块,由 b_k 表示。假设信源字符集具有 K 个不同的符号,第 k 个信号 s_k 发生的概率为 p_k , $k=0,1,\dots,K-1$ 。编码器分配给符号 s_k 的二进制码字的长度为 l_k ,以比特度量。定义信源编码器的平均码长 \bar{L} 为

$$\bar{L} = \sum_{k=0}^{K-1} p_k l_k \quad (9.18)$$

在物理术语中,参数 \bar{L} 表示信源编码过程中每个信源符号的平均比特数。用 L_{\min} 表示 \bar{L} 的最小可能值,可定义信源编码器的码率为

$$\eta = \frac{L_{\min}}{\bar{L}} \quad (9.19)$$

由于 $\bar{L} \geq L_{\min}$,显然有 $\eta \leq 1$ 。当 η 接近于1,可认为信源编码器是有效的。

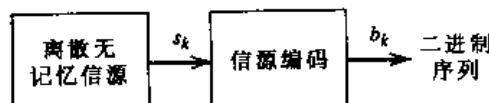


图9.3 信源编码

但是,如何确定最小值 L_{\min} ?这个基本问题的答案就包含在香农第一定理,即信源编码定理^[4]中:

给定熵为 $H(\mathcal{S})$ 的离散无记忆信源,任意无失真信源编码方式的平均字长 \bar{L} 的上限为

$$\bar{L} \geq H(\mathcal{S}) \quad (9.20)$$

针对一类特殊信源编码的相应证明见下一节。根据信源编码定理,熵 $H(\mathcal{S})$ 表明了一个离散无记忆信源中每个信源符号平均比特数的基本限制,即该比特数不能小于熵 $H(\mathcal{S})$ 。当 $L_{\min} = H(\mathcal{S})$ 时,可将信源编码器效率以 $H(\mathcal{S})$ 表示为

$$\eta = \frac{H(\mathcal{S})}{\bar{L}} \quad (9.21)$$

9.4 无失真数据压缩

物理信源产生的信号都有一个共性:它们的原始形式包含大量的冗余信息,这些冗余信息的传输是对主要通信资源的浪费。为了进行有效的信号传输,冗余信息必须在传输之前就从信号中去除掉。这个操作对信息没有损伤,且通常对数字信号进行,因此称为无失真数据压缩或是无损数据压缩。用压缩后的码字来表示信源的输出,可确保每个符号的平均比特数都是高效而准确的,因为原始数据可以无损地被恢复。信源的熵确定了从数据中去除冗余的基本限制。基本上,无失真数据压缩均是通过对信源输出中经常出现的符号赋予短记号、对不经常出现的符号赋予长记号来实现。

本节将讨论用于无失真数据压缩的一些信源编码方案。下面以一种称为前缀码的信源编码作为讨论的开始。前缀码不仅是可译码的,而且有可能使平均码长无限接近于信源熵。

9.4.1 前缀码

设一个离散无记忆信源,其字符集为 $\{s_0, s_1, \dots, s_{K-1}\}$,统计量为 $\{p_0, p_1, \dots, p_{K-1}\}$,由于表

示信源输出的信源编码具有实际应用性,码字必须是可惟一译码的。这个限制可确保对于信源发出符号的每个有限序列来说,其码字对应的序列与任何其他信源序列对应码字的序列是各不相同的。前缀码就是符合这个限制的一类特殊的编码。为了定义前缀条件,将分配给信源信号 s_k 的码字表示为 $|m_{k_1}, m_{k_2}, \dots, m_{k_n}|$, 其中, 单个变量 $m_{k_1}, m_{k_2}, \dots, m_{k_n}$ 的取值非 0 即 1, n 为码字长度。码字的前一部分由 $m_{k_1}, m_{k_2}, \dots, m_{k_i}$ 表示, $i \leq n$ 。由码字前一部分构成的序列称为此码字的前缀。在前缀码中, 没有一个码字可作为其他任何码字的前缀。

为了说明前缀码的含义, 考虑表 9.2 中的三个信源编码。编码 I 不是前缀码, 这是因为比特 0(即 s_0 的码字)是 00 的前缀, 而 00 是 s_2 的码字。同样的, 比特 1(即 s_1 的码字)是 11 的前缀, 而 11 是 s_3 的码字。类似地, 可以证明编码 III 也不是前缀码, 但码 II 是前缀码。

表 9.2 前缀码的定义说明

信源符号	出现概率	码 I	码 II	码 III
s_0	0.5	0	0	0
s_1	0.25	1	10	01
s_2	0.125	00	110	011
s_3	0.125	11	111	0111

在对前缀信源编码产生的码字序列进行译码时, 信源译码器从序列的起始端开始, 每次译一个码字。译码器建立起一个判决树, 判决树是对特定信源编码中码字的图形描述。图 9.4 所示为表 9.2 中编码 II 的判决树。图中的判决树包含一个初始状态和(对应于信源符号码 s_0, s_1, s_2, s_3 的)四个最终状态。译码通常从初始状态开始。如果第一个接收到的比特是 0, 则将译码器的状态变为最终状态 s_0 , 否则, 转向第二个判决点。在后一种情况下, 第二个比特将译码器状态沿判决树进一步移动, 如果是 0 则最终状态为 s_1 , 否则, 转向第三个判决点, 以此类推。一旦每个最终状态都发送了对应的符号, 译码器就重置为初始状态。注意, 接收到的编码序列的每个比特只被检验一次。例如编码序列 1011111000…就逐步译码为信源序列 $s_1 s_3 s_2 s_0 s_0 \dots$ 。读者可以自己执行一下这个译码过程。

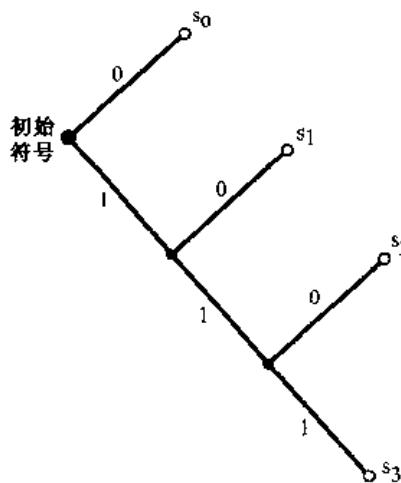


图 9.4 表 9.2 中编码 II 的判决树

前缀码的重要特性是它总可以被惟一地译码。但是,反过来说就不一定正确了。例如,表9.2中的编码III就不满足前缀码的条件,但它也能够被惟一地译码,这是因为比特0标明了每个码字的起始。

如果前缀码被用于构造一个离散无记忆信源,其信源字符集为 $|s_0, s_1, \dots, s_{K-1}|$,信源统计值为 $|p_0, p_1, \dots, p_{K-1}|$,符号 s_k 的码字长度为 l_k , $k = 0, 1, \dots, K-1$,则码字长度通常满足一个称为Kraft-McMillan的不等式:^[5]

$$\sum_{k=0}^{K-1} 2^{-l_k} \leq 1 \quad (9.22)$$

其中,因子2为二进制字符集的基数(符号个数)。但需要注意的是,Kraft-McMillan不等式并不能告诉我们一个信源编码是否是前缀码。它只是码字长度必须满足的一个条件,而不是判断码字本身是否为前缀码的条件。例如表9.2中的三种码:

- 码I不符合Kraft-McMillan不等式,因此不可能是前缀码。
- 码II和III都符合Kraft-McMillan不等式,但只有码II是前缀码。

前缀码不同于其他可惟一译码的码,就是因为其每个码字的末尾都能被识别出来。所以,一旦表示信源符号的二进制序列被全部接收,前缀码的译码也就完成了。正因如此,前缀码又称为瞬时码。

假设离散无记忆信源的熵为 $H(\mathcal{S})$,前缀码的平均码长为 \bar{L} ,则有如下的限制:

$$H(\mathcal{S}) \leq \bar{L} < H(\mathcal{S}) + 1 \quad (9.23)$$

当信源发出符号 s_k 的概率如下时,式(9.23)的左半部分取等号:

$$p_k = 2^{-l_k} \quad (9.24)$$

其中, l_k 为分配给信源符号 s_k 的码字长度。从而可得

$$\sum_{k=0}^{K-1} 2^{-l_k} = \sum_{k=0}^{K-1} p_k = 1$$

在这个条件下,由式(9.22)所示的Kraft-McMillan不等式可知,可以构造一个前缀码,且分配给信源符号 s_k 的码字长度为 $-\log_2 p_k$ 。对于这样的编码,平均码长为

$$\bar{L} = \sum_{k=0}^{K-1} \frac{l_k}{2^{-l_k}} \quad (9.25)$$

相应的信源熵为

$$\begin{aligned} H(\mathcal{S}) &= \sum_{k=0}^{K-1} \left(\frac{1}{2^{-l_k}} \right) \log_2(2^{-l_k}) \\ &= \sum_{k=0}^{K-1} \frac{l_k}{2^{-l_k}} \end{aligned} \quad (9.26)$$

这是一种特殊(不常见)的情况。由式(9.25)和式(9.26)可知,前缀码与信源匹配,这是因为 $\bar{L} = H(\mathcal{S})$ 。

但是,如何将前缀码匹配于任意离散无记忆信源呢?答案就是采用扩展码。用 \bar{L}_n 代表扩

展前缀码的平均长度。对于一个可惟一译码的编码, \bar{L}_n 代表了最小的可能性。由式(9.23)可推导出

$$H(\mathcal{G}^n) \leq \bar{L}_n < H(\mathcal{G}^n) + 1 \quad (9.27)$$

将有关扩展信源的式(9.17)代入式(9.27), 可得

$$nH(\mathcal{S}) \leq \bar{L}_n < nH(\mathcal{S}) + 1$$

或等价地得到

$$H(\mathcal{S}) \leq \frac{\bar{L}_n}{n} < H(\mathcal{S}) + \frac{1}{n} \quad (9.28)$$

当 n 趋于无穷时, 式(9.28)的上限和下限重合, 于是有

$$\lim_{n \rightarrow \infty} \frac{1}{n} \bar{L}_n = H(\mathcal{S}) \quad (9.29)$$

因此, 可以认为, 通过使扩展前缀信源编码器的阶数 n 足够大, 可以使码尽可能如期望地如实代表离散无记忆信源 \mathcal{S} 。换言之, 按照信源编码定理, 只要扩展码具有足够高的阶数, 扩展前缀码的平均码长就可以和信源熵一样小。但是, 降低平均码长的代价是提高了译码的复杂度, 因为译码复杂度是随着扩展前缀码阶数的增加而提高的。

9.4.2 霍夫曼编码

接下来讨论一类重要的前缀码, 霍夫曼编码。^[6]的基本思想是为字符集的每个符号分配一个比特序列, 其长度与按符号传递的信息量大致相等。最后得到一个信源编码, 其平均码长接近于离散无记忆信道熵的基本限制值, 即 $H(\mathcal{S})$ 。用于合成霍夫曼码的算法, 本质上是将预设的离散无记忆信源的统计值替换为更简单的形式。简化过程是一步步进行的, 直到最后只剩下两个信源统计量(符号), 其中(0,1)是最佳的编码。从这个码出发, 可反过来构造给定信源的霍夫曼编码。

霍夫曼编码算法可按如下方式进行:

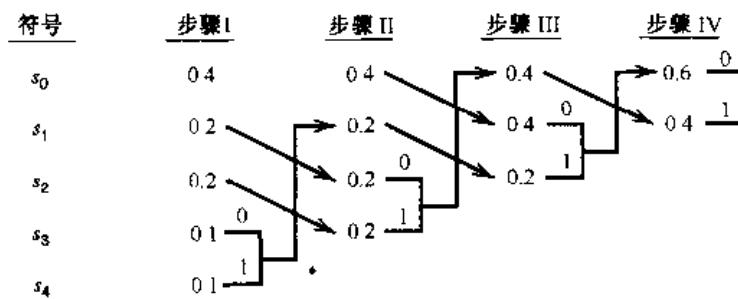
1. 将信源符号按概率递减的顺序排列。概率最低的两个信源符号被赋以 0 和 1。这个步骤称为分裂阶段。
2. 将这两个信源符号结合为新的信源符号, 其概率等于两个原始概率的和。(因此信源符号的列表以及信源统计量相应减少 1 个。)新符号的概率置于与其值相应的列表位置上。
3. 重复这个过程, 直到信源统计量(符号)的最终列表只剩下两个, 分别被赋以 0 和 1。

通过反向追溯此符号及其后续符号的 0 和 1 的序列即可得到每个(原始)信源符号的编码。

例 9.3 霍夫曼树

离散无记忆信源字符集的五个符号及其概率如图 9.5(a)的最左端两列所示。按照霍夫曼算法, 经过四步可得到运算结果, 如图 9.5(a)所示的霍夫曼树。将该信源的霍夫曼编码的码字制成表格, 如图 9.5(b)所示。平均码长为

$$\bar{L} = 0.4 \times (2) + 0.2 \times (2) + 0.2 \times (2) + 0.1 \times (3) + 0.1 \times (3) \\ = 2.2$$



(a) 霍夫曼编码

符号	概率	码字
s_0	0.4	00
s_1	0.2	10
s_2	0.2	11
s_3	0.1	010
s_4	0.1	011

(b) 信源编码

图 9.5 霍夫曼编码示例

如式(9.9), 离散无记忆信源的熵为

$$\begin{aligned} H(\mathcal{S}) &= 0.4 \log_2\left(\frac{1}{0.4}\right) + 0.2 \log_2\left(\frac{1}{0.2}\right) + 0.2 \log_2\left(\frac{1}{0.2}\right) \\ &\quad + 0.1 \log_2\left(\frac{1}{0.1}\right) + 0.1 \log_2\left(\frac{1}{0.1}\right) \\ &= 0.52877 + 0.46439 + 0.46439 + 0.33219 + 0.33219 \\ &= 2.12193 \end{aligned}$$

由本例可做如下论断:

1. 平均码长 \bar{L} 仅比熵 $H(\mathcal{S})$ 大 3.67%。
2. 平均码长 \bar{L} 满足式(9.23)。

值得注意的是,霍夫曼编码过程(即霍夫曼树)不是惟一的。造成霍夫曼编码非惟一性的原因有两个。首先,在构造霍夫曼编码的每个分裂过程中,给最后两个信源符号赋予 0 或 1 具有任意性。但不管如何赋值,最终的区别都是微小的。第二点,合成符号的概率(通过将指定操作步骤的最末两个概率相加得到)可能与列表中的另一个概率相同,从而造成模棱两可的情况。可将新符号的概率放置得尽量高,如例 9.3。相反地,也可放置得尽量低。预先设定是高放置还是低放置,这在编码过程中固定不变。但这个原因造成了码字差异很大,因为最终相同信源编码的码字可能具有不同长度。但无论如何,平均码长是不变的。

作为对信源编码的码字长度差异的度量,在整个信源符号上定义平均码长 \bar{L} 的方差为

$$\sigma^2 = \sum_{k=0}^{K-1} p_k (l_k - \bar{L})^2 \quad (9.30)$$

其中, p_0, p_1, \dots, p_{K-1} 为信源统计量, l_k 为分配给符号 s_k 的码字长度。通常,尽可能高地放置合成符号得到的霍夫曼编码的方差 σ^2 比尽可能低地放置要小得多。基于此,选择前一种霍夫曼编码是较为合理的。

9.4.3 Lempel-Ziv 编码

霍夫曼编码的一个缺点在于,它要求知道信源的概率模型。遗憾的是,在实际应用中,信源统计值不是总能预先知道的。此外,在模型文本中,存储要求阻碍了霍夫曼编码捕获码字与短语之间的高阶关系,从而降低了码率。为了克服这些实际限制,可采用 Lempel-Ziv 算法^[7],这种算法比霍夫曼编码更具有自适应性,且更易实现。

Lempel-Ziv 算法的编码是通过将信源数据流解析为以前未遇到过的最短的子序列来完成的。为了说明这个简单但相当严格的概念,考虑如下二进制输入序列:

000101110010100101…

假设二进制符号 0 和 1 已经按此顺序存储在码本中。即

存储的子序列: 0, 1

被解析的数据: 000101110010100101…

编码过程从左端开始。由于 0 和 1 已经被存储,第一次遇到的、之前未见过的数据流的最短子序列是 00,即

存储的子序列: 0, 1, 00

被解析的数据: 0101110010100101…

第二个之前未见过的最短子序列是 01,可相应地写为

存储的子序列: 0, 1, 00, 01

被解析的数据: 01110010100101…

下一个之前未见过的最短子序列是 011,可相应地写为

存储的子序列: 0, 1, 00, 01, 011

被解析的数据: 10010100101

继续这个过程,直到所给的数据流被全部解析。根据这个例子,可得到二进制子序列的码本,如图 9.6 中第二行所示。

位置的数字表示	1	2	3	4	5	6	7	8	9
子序列	0	1	00	01	011	10	010	100	101
数字化的表示			11	12	42	21	41	61	62
二进制编码块			0010	0011	1001	0100	1000	1100	1101

图 9.6 序列 000101110010100101… 的 Lempel-Ziv 算法编码过程

图中的第一行只是说明了码本中的单个子序列的数字位置。不难发现,数据流的第一个子序列 00 是码本的第一项 0 与其自身的级联,因此由数字 11 表示。数据流的第二个子序列 01 是码本的第一项 0 与第二项 1 的级联,因此由数字 12 表示。剩下的子序列也按照这种方式处理。码本中不同子序列的数字表示的完整集见图 9.6 的第三行。其中,子序列 010 由位置 4 的子序列 01 和位置 1 的符号 0 构成,因此其数字表示为 41。图 9.6 的最后一行是数据流中不同子序列的二进制编码表示。

码本中的每个子序列的最后一个符号(即图 9.6 的第二行)都是创新符号。称其为创新符号的原因是,这个符号附在特定子序列的后面,使自己与所有先前保存在码本中的子序列不相

同。因此,数据流(即图 9.6 的第四行)二进制编码形式中的每个相同块的最后一个比特代表了此子序列的创新符号。剩下的比特提供了根子序列“指针”的等价二进制表示,根子序列与除创新符号以外的符号相匹配。

译码器与编码器一样简单。其中采用指针来识别根子序列,然后将创新符号附在后面。例如处于位置 9 的二进制编码块 1101,最后一个比特 1 为创新符号,剩下的比特 110 指向位置 6 的根子序列 10。因此,1101 译码为 101 才是正确的。

由此例可知,不同于霍夫曼编码,Lempel-Ziv 算法采用固定长度的码字表示不同长度的信源符号。这个特性使得 Lempel-Ziv 编码适合于进行同步传输。实际应用中采用 12 比特的固定块,这暗示着码本有 4096 项。

长期以来,霍夫曼编码一直作为首选的无失真数据压缩算法,这种地位未受到过任何挑战。但是,Lempel-Ziv 算法几乎完全取代了霍夫曼编码的地位。Lempel-Ziv 算法是文件压缩的标准算法。当应用于普通的英语文章时,Lempel-Ziv 算法能获得 55% 的压缩率。这显然优于霍夫曼编码所获得的 43% 的压缩率。如前所述,这种情况发生的原因是,霍夫曼编码在语言字符间冗余上不占有任何优势。而 Lempel-Ziv 算法能够以较高的效率尽最大可能压缩文本(在某些限制下)。

9.5 离散无记忆信道

到现在为止,本章一直在讨论产生信息的离散无记忆信源。接下来将介绍信息传输的概念,其中特别强调信息传输的可靠性。下面以一个离散无记忆信道,即离散无记忆信源的配对来开始我们的讨论。

离散无记忆信道是具有输入 X 和输出 Y 的统计模型,其中, Y 是 X 的有噪声形式, X 和 Y 都是随机变量。在每个单位时间内,信道接收到选自字符集 \mathcal{X} 中的一个输入符号 X 。作为回应,从字符集 \mathcal{Y} 中输出符号 Y 。当字符集 \mathcal{X} 和 \mathcal{Y} 均为有限大时,这种信道被认为是“离散”的。如果当前输出符号只取决于当前输入符号,而与先前输入的无关,信道则被认为是“无记忆”的。

图 9.7 所示为一个离散无记忆信道。信道被描述为一个输入字符集

$$\mathcal{X} = \{x_0, x_1, \dots, x_{j-1}\}, \quad (9.31)$$

一个输出字符集

$$\mathcal{Y} = \{y_0, y_1, \dots, y_{k-1}\}, \quad (9.32)$$

和一组转移概率

$$p(y_k | x_j) = P(Y = y_k | X = x_j) \quad \text{对于所有 } j \text{ 和 } k \quad (9.33)$$

显然有

$$0 \leq p(y_k | x_j) \leq 1 \quad \text{对于所有 } j \text{ 和 } k \quad (9.34)$$

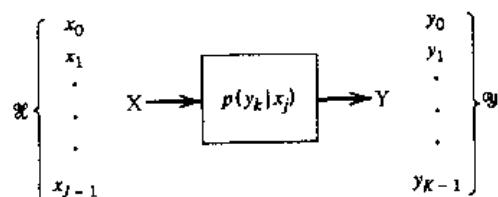


图 9.7 离散无记忆信道

输入字符集 \mathcal{X} 和输出字符集 \mathcal{Y} 不必具有相同大小。例如,在信道编码中,输出字符集 \mathcal{Y} 的大小 K 可能大于输入字符集 \mathcal{X} 的大小 J (即 $K \geq J$)。有时也可能遇到这样的情况:两个输入符号中的任意一个被发送时,信道发出的符号是一样的,此时 $K \leq J$ 。

描述离散无记忆信道的一个简便方法,是将信道的各种转移概率置于如下矩阵中:

$$\mathbf{P} = \begin{bmatrix} p(y_0|x_0) & p(y_1|x_0) & \cdots & p(y_{K-1}|x_0) \\ p(y_0|x_1) & p(y_1|x_1) & \cdots & p(y_{K-1}|x_1) \\ \vdots & \vdots & \ddots & \vdots \\ p(y_0|x_{J-1}) & p(y_1|x_{J-1}) & \cdots & p(y_{K-1}|x_{J-1}) \end{bmatrix} \quad (9.35)$$

$J \times K$ 矩阵 \mathbf{P} 称为信道矩阵或传输矩阵。值得注意的是,信道矩阵 \mathbf{P} 的每一行对应于一个固定的信道输入,而矩阵的每一列则对应着一个固定的信道输出。此外,还需注意信道矩阵 \mathbf{P} 的基本特性:矩阵每一行数据的和总等于1。即

$$\sum_{k=0}^{K-1} p(y_k|x_j) = 1 \quad \text{对于所有 } j \quad (9.36)$$

假设离散无记忆信道的输入概率分布为 $\{p(x_j), j = 0, 1, \dots, J-1\}$ 。换言之,信道输入 $X = x_j$ 的发生概率为

$$p(x_j) = P(X = x_j) \quad j = 0, 1, \dots, J-1 \quad (9.37)$$

以随机变量 X 表示信道输入,随机变量 Y 表示信道输出,则随机变量 X 和 Y 的联合概率分布为

$$\begin{aligned} p(x_j, y_k) &= P(X = x_j, Y = y_k) \\ &= P(Y = y_k | X = x_j)P(X = x_j) \\ &= p(y_k|x_j)p(x_j) \end{aligned} \quad (9.38)$$

输出随机变量 Y 的边缘概率分布由 $p(x_j, y_k)$ 在 x_j 上的平均来获得:

$$\begin{aligned} p(y_k) &= P(Y = y_k) \\ &= \sum_{j=0}^{J-1} P(Y = y_k | X = x_j)P(X = x_j) \\ &= \sum_{j=0}^{J-1} p(y_k|x_j)p(x_j) \quad k = 0, 1, \dots, K-1 \end{aligned} \quad (9.39)$$

概率 $p(x_j), j = 0, 1, \dots, J-1$ 被认为是输入符号的先验概率。式(9.39)说明,如果给定输入先验概率 $p(x_j)$ 和信道矩阵(即转移概率 $p(y_k|x_j)$ 的矩阵),就能够计算输出符号的概率 $p(y_k)$ 。

例 9.4 二进制对称信道

二进制对称信道具有重要的理论和实用价值,它是离散无记忆信道在 $J = K = 2$ 时的特例。信道有两个输入符号($x_0 = 0, x_1 = 1$)和两个输出符号($y_0 = 0, y_1 = 1$)。发送0而接收到的是1,与发送1而接收到的是0的概率相同,则信道是对称的。此时条件差错概率由 p 表示。二进制对称信道的转移概率如图9.8所示。

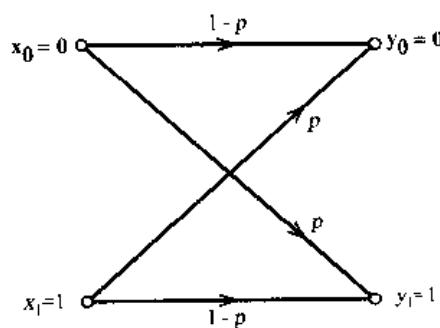


图 9.8 二进制对称信道的转移概率图

将图 9.8 所示的转移概率与由 3.3 节中的 PCM 接收机确定的条件差错概率 p_{10} 和 p_{01} 相联系是很有意义的。当二进制符号 0 和 1 等概出现时, 上述两个差错概率的最优值是相等的。回顾以下定义(使用图 9.8 中的符号):

$$p_{10} = P(y = 1 | x = 0)$$

和

$$p_{01} = P(y = 0 | x = 1)$$

可立即由图 3.4 的 PCM 接收机得出

$$p_{10} = p_{01} = p$$

9.6 互信息

假设信道输出 Y (选自字符集 \mathcal{Y})为信道输入 X (选自字符集 \mathcal{X})的有噪声形式, 熵 $H(\mathcal{X})$ 为 X 的先验不确定性的度量, 如何度量得到 Y 以后的 X 的不确定性? 为了回答这个问题, 下面将通过定义给定 $Y = y_k$ 时, 选自字符集 \mathcal{X} 的 X 的条件熵, 将 9.2 节中讨论的概念进行扩展。该条件熵的定义为

$$H(\mathcal{X} | Y = y_k) = \sum_{j=0}^{J-1} p(x_j | y_k) \log_2 \left[\frac{1}{p(x_j | y_k)} \right] \quad (9.40)$$

这个量本身是一个随机变量, 分别以概率 $p(y_0), \dots, p(y_{K-1})$ 取值 $H(\mathcal{X} | Y = y_0), \dots, H(\mathcal{X} | Y = y_{K-1})$ 。在整个输出字符集 \mathcal{Y} 上, $H(\mathcal{X} | Y = y_k)$ 的平均为

$$\begin{aligned} H(\mathcal{X} | \mathcal{Y}) &= \sum_{k=0}^{K-1} H(\mathcal{X} | Y = y_k) p(y_k) \\ &= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j | y_k) p(y_k) \log_2 \left[\frac{1}{p(x_j | y_k)} \right] \\ &= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j, y_k) \log_2 \left[\frac{1}{p(x_j | y_k)} \right] \end{aligned} \quad (9.41)$$

上式最后一行中,用到了关系式

$$p(x_j, y_k) = p(x_j | y_k)p(y_k) \quad (9.42)$$

$H(\mathcal{X} | \mathcal{Y})$ 称为条件熵。它表示信道输出确定时,信道输入不确定性的度量。

由于 $H(\mathcal{X})$ 表示在确定信道输出之前的信道输入不确定性,条件 $H(\mathcal{X} | \mathcal{Y})$ 表示信道输出确定以后的信道输入不确定性,因此差 $H(\mathcal{X}) - H(\mathcal{X} | \mathcal{Y})$ 必然是表示通过观察信道输出而消除的信道输入的不确定性。这个重要的量称为信道的互信息。互信息可由 $I(\mathcal{X}; \mathcal{Y})$ 表示为

$$I(\mathcal{X}; \mathcal{Y}) = H(\mathcal{X}) - H(\mathcal{X} | \mathcal{Y}) \quad (9.43)$$

相应地,有

$$I(\mathcal{Y}; \mathcal{X}) = H(\mathcal{Y}) - H(\mathcal{Y} | \mathcal{X}) \quad (9.44)$$

其中, $H(\mathcal{Y})$ 为信道输出的熵, $H(\mathcal{Y} | \mathcal{X})$ 为给定信道输入时信道输出的条件熵。

9.6.1 互信息的性质

互信息 $I(\mathcal{X}; \mathcal{Y})$ 具有以下重要性质:

性质 1

信道的互信息是对称的,即

$$I(\mathcal{X}; \mathcal{Y}) = I(\mathcal{Y}; \mathcal{X}) \quad (9.45)$$

其中,互信息 $I(\mathcal{X}; \mathcal{Y})$ 为通过观察信道输出而消除的信道输入不确定性的度量,互信息 $I(\mathcal{Y}; \mathcal{X})$ 为通过信道输入消除的信道输出不确定性的度量。

为了证明这个性质,可利用熵的公式,并结合式(9.36)和式(9.38),将 $H(\mathcal{X})$ 表述为

$$\begin{aligned} H(\mathcal{X}) &= \sum_{j=0}^{J-1} p(x_j) \log_2 \left[\frac{1}{p(x_j)} \right] \\ &= \sum_{j=0}^{J-1} p(x_j) \log_2 \left[\frac{1}{p(x_j)} \right] \sum_{k=0}^{K-1} p(y_k | x_j) \\ &= \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(y_k | x_j) p(x_j) \log_2 \left[\frac{1}{p(x_j)} \right] \\ &= \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left[\frac{1}{p(x_j)} \right] \end{aligned} \quad (9.46)$$

将式(9.41)和式(9.46)代入式(9.43)并进行合并可得

$$I(\mathcal{X}; \mathcal{Y}) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left[\frac{p(x_j | y_k)}{p(x_j)} \right] \quad (9.47)$$

如式(9.38)和式(9.42),由条件概率的贝叶斯准则,有

$$\frac{p(x_j | y_k)}{p(x_j)} = \frac{p(y_k | x_j)}{p(y_k)} \quad (9.48)$$

将式(9.48)代入式(9.47)并互换加法的顺序,得

$$\begin{aligned} I(\mathcal{X}; \mathcal{Y}) &= \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j, y_k) \log_2 \left[\frac{p(y_k | x_j)}{p(y_k)} \right] \\ &= I(\mathcal{Y}; \mathcal{X}) \end{aligned} \quad (9.49)$$

这正是我们想要的结果。

性质 2

互信息总是非负的, 即

$$I(\mathcal{X}; \mathcal{Y}) \geq 0 \quad (9.50)$$

为了证明这个性质, 首先由式(9.42)得到

$$p(x_j | y_k) = \frac{p(x_j, y_k)}{p(y_k)} \quad (9.51)$$

将式(9.51)代入式(9.47), 可将信道的互信息表示为

$$I(\mathcal{X}; \mathcal{Y}) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left(\frac{p(x_j, y_k)}{p(x_j)p(y_k)} \right) \quad (9.52)$$

接着, 直接应用由式(9.12)定义的基本不等式可得

$$I(\mathcal{X}; \mathcal{Y}) \geq 0$$

当且仅当

$$p(x_j, y_k) = p(x_j)p(y_k) \quad \text{对于所有 } j \text{ 和 } k \quad (9.53)$$

时取等号。

性质 2 说明: 平均来说, 不会由于观察信道输出而丢失信息。此外, 当且仅当信道的输入和输出符号统计独立时, 互信息才等于 0, 见式(9.53)。

性质 3

信道的互信息与信道输入输出联合熵的关系如下:

$$I(\mathcal{X}; \mathcal{Y}) = H(\mathcal{X}) + H(\mathcal{Y}) - H(\mathcal{X}, \mathcal{Y}) \quad (9.54)$$

其中, 联合熵 $H(\mathcal{X}, \mathcal{Y})$ 的定义为

$$H(\mathcal{X}, \mathcal{Y}) = \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left(\frac{1}{p(x_j, y_k)} \right) \quad (9.55)$$

为了证明式(9.54), 首先将 $H(\mathcal{X}, \mathcal{Y})$ 写为

$$\begin{aligned} H(\mathcal{X}, \mathcal{Y}) &= \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left[\frac{p(x_j)p(y_k)}{p(x_j, y_k)} \right] \\ &\quad + \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left[\frac{1}{p(x_j)p(y_k)} \right] \end{aligned} \quad (9.56)$$

式(9.56)中等式右边的第一项可认为是信道互信息 $I(\mathcal{X}; \mathcal{Y})$ 的取负, 见式(9.52)。对于第二项, 有

$$\begin{aligned}
 \sum_{j=0}^{J-1} \sum_{k=0}^{K-1} p(x_j, y_k) \log_2 \left[\frac{1}{p(x_j)p(y_k)} \right] &= \sum_{j=0}^{J-1} \log_2 \left[\frac{1}{p(x_j)} \right] \sum_{k=0}^{K-1} p(x_j, y_k) \\
 &\quad + \sum_{k=0}^{K-1} \log_2 \left[\frac{1}{p(y_k)} \right] \sum_{j=0}^{J-1} p(x_j, y_k) \\
 &= \sum_{j=0}^{J-1} p(x_j) \log_2 \left[\frac{1}{p(x_j)} \right] \\
 &\quad + \sum_{k=0}^{K-1} p(y_k) \log_2 \left[\frac{1}{p(y_k)} \right] \\
 &= H(\mathcal{X}) + H(\mathcal{Y})
 \end{aligned} \tag{9.57}$$

将式(9.52)和式(9.57)代入式(9.56), 得

$$H(\mathcal{X}, \mathcal{Y}) = -I(\mathcal{X}; \mathcal{Y}) + H(\mathcal{X}) + H(\mathcal{Y}) \tag{9.58}$$

重排各项, 得到式(9.54)所示的结果, 从而证明了性质 3。

由式(9.43), 式(9.44), 式(9.54)的图表解释, 可得到信道互信息的相关结论(见图 9.9)。在图9.9中, 信道输入 X 的熵由左边的圆圈表示。信道输出 Y 的熵由右边的圆圈表示。信道的互信息由两个圆圈的重叠部分表示。

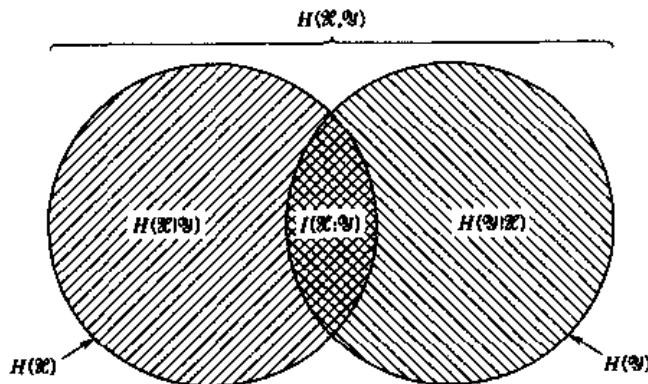


图 9.9 信道输入输出熵的关系图

9.7 信道容量

考虑一个离散无记忆信道, 其输入字符集为 \mathcal{X} , 输出字符集为 \mathcal{Y} , 转移概率为 $p(y_k|x_j)$, 其中, $j = 0, 1, \dots, J-1$ 且 $k = 0, 1, \dots, K-1$ 。信道的互信息由式(9.49)的第一行定义, 即

$$I(\mathcal{X}; \mathcal{Y}) = \sum_{k=0}^{K-1} \sum_{j=0}^{J-1} p(x_j, y_k) \log_2 \left[\frac{p(y_k|x_j)}{p(y_k)} \right]$$

由式(9.38), 有

$$p(x_j, y_k) = p(y_k|x_j)p(x_j)$$

同样, 由式(9.39), 有

$$p(y_k) = \sum_{j=0}^{J-1} p(y_k|x_j)p(x_j)$$

显然,如果已知输入概率分布 $\{p(x_j)\mid j=0,1,\dots,J-1\}$,就能够计算互信息 $I(\mathcal{X};\mathcal{Y})$ 。信道的互信息不仅取决于信道,还取决于如何使用信道。

输入概率分布 $\{p(x_j)\}$ 显然是独立于信道的。因此可以使得互信息 $I(\mathcal{X};\mathcal{Y})$ 关于 $\{p(x_j)\}$ 最大化。定义离散无记忆信道的信道容量为在信道的任何单个使用中(即信号发送间隔)的最大互信息 $I(\mathcal{X};\mathcal{Y})$ 。其中,最大化是在 \mathcal{X} 所有可能的输入概率分布 $\{p(x_j)\}$ 上进行的。信道容量通常由 C 表示,于是有

$$C = \max_{\{p(x_j)\}} I(\mathcal{X};\mathcal{Y}) \quad (9.59)$$

信道容量 C 由比特/信道使用或比特/传输来度量。

值得注意的是,信道容量 C 只是(用于定义信道的)转移概率 $\{p(y_k|x_j)\}$ 的函数。 C 就是 J 个变量即输入概率 $p(x_0), \dots, p(x_{J-1})$ 上的互信息 $I(\mathcal{X};\mathcal{Y})$ 的最大化,满足如下两个限制:

$$p(x_j) \geq 0 \text{ 对于所有 } j$$

和

$$\sum_{j=0}^{J-1} p(x_j) = 1$$

总的来说,真正的难题在于确定 C 是如何变化的。

例 9.5 二进制对称信道(回顾)

再次考虑由图 9.8 的转移概率图描述的二进制对称信道。该信道由条件差错概率 p 惟一定义。

当信道输入概率 $p(x_0) = p(x_1) = 1/2$ 时,熵 $H(X)$ 取最大值,其中 x_0 和 x_1 不是 0 就是 1。互信息 $I(\mathcal{X};\mathcal{Y})$ 同样取最大值。因此有

$$C = I(\mathcal{X};\mathcal{Y}) \Big|_{p(x_0)=p(x_1)=1/2}$$

由图 9.8,有

$$p(y_0|x_1) = p(y_1|x_0) = p$$

和

$$p(y_0|x_0) = p(y_1|x_1) = 1 - p$$

将这些信道转移概率和 $J = K = 2$ 代入式(9.49),并对应于式(9.59)将输入概率设为 $p(x_0) = p(x_1)$,则二进制对称信道的容量为

$$C = 1 + p \log_2 p + (1 - p) \log_2(1 - p) \quad (9.60)$$

使用式(9.16)所示的熵函数的定义,可将式(9.60)简化为

$$C = 1 - H(p)$$

信道容量 C 随差错概率(转移概率) p 成凹状分布,如图 9.10 所示,该图关于 $p = 1/2$ 对称。将此图中的曲线与图 9.2 进行比较,可得如下结论:

- 当信道是无噪声的,允许 $p = 0$,此时信道容量 C 获得最大值(以比特/信道使用表示),这个值其实也就是每个信道输入的信息。在 p 点,熵 $H(p)$ 具有最小值 0。
- 当由噪声产生的条件差错概率 $p = 1/2$ 时,信道容量 C 获得最小值 0,而熵函数 $H(p)$ 获得最大值 1。这种情况下信道被认为是无用的。

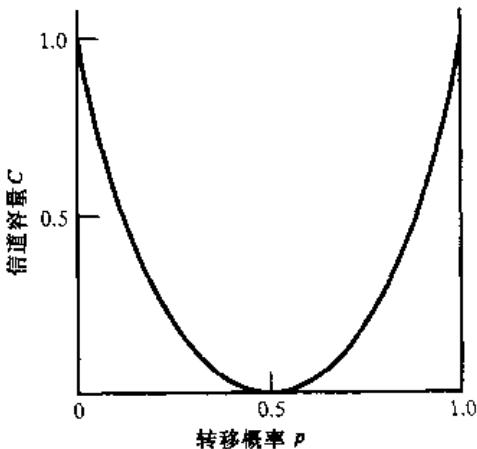


图 9.10 转移概率为 p 的二进制对称信道的信道容量变化曲线

9.8 信道编码定理

信道中的噪声是不可避免的,这就造成了数字通信系统中的输出和输入序列之间的差异(错误)。对于一个相对多噪声信道(例如无线通信信道),差错概率可能达到 10^{-1} ,这表明每 10 个发送比特(平均的)中只有 9 个是正确接收的。在很多应用中,这种可靠性等级是无法接受的。实际上,通常都要求差错概率等于 10^{-6} 或更小。为此,可采用信道编码技术。

信道编码的设计目标是增强数字通信系统对信道噪声的抵抗能力。信道编码就是将输入数据序列映射到信道输入序列,将信道输出序列逆映射到输出数据序列,这样使得信道噪声对系统的整体影响最小化。第一个映射操作由发射机中的信道编码器完成,而第二个映射(即逆映射)操作由接收机中的信道译码器完成,如图 9.11 的框图所示。为了简化说明,图 9.11 中没有包含信源编码(在信道编码之前)和信源译码(在信道译码之后)。

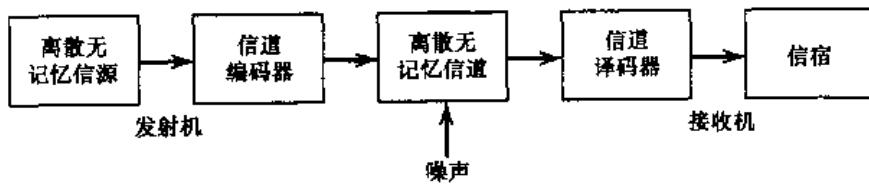


图 9.11 数字通信系统框图

图 9.11 中的信道编码器和译码器都是由设计者控制的,其作用是使得通信系统的整体可靠性最优化。采取的方法是在信道编码器中引入冗余,以尽可能精确地重构原始信源序列。更广义地讲,可认为信道编码是信源编码的重复,前者引入受控制的冗余以提高可靠性,而后者则减少冗余以提高有效性。

信道编码将在第 10 章中详细讨论。现在仅着重讨论分组码。在这类编码中,消息序列被

分解为连续的分组,每一个分组有 k 个比特,将每个 k 比特分组映射到一个 n 比特分组上,其中 $n > k$ 。编码器为每个传输分组加入的冗余为 $n - k$ 比特。 k/n 称为码率,用 r 表示。即

$$r = \frac{k}{n}$$

显然, r 小于 1。对于一个预设的 k ,当分组长度 n 趋近于无穷时,码率 r (以及系统码率)趋近于 0。

信宿端的原始信源序列精确重构要求平均符号差错概率任意小。这就引出了下述重要问题:是否存在一种信道编码机制,使得消息比特发生差错的概率小于任意一个正数(即按需要确定的任意小的一个正数),这种信道编码机制又能否在码率毋需太小的情况下有效?答案是肯定的。实际上,关于信道容量的香农第二定理就回答了这个问题。下面就介绍这个定理。到目前为止,在我们对信道容量的讨论中,并没有过多地涉及到时间。假设图 9.11 中的离散无记忆信源具有信源字符集 \mathcal{S} ,且每个信源符号包含 $H(\mathcal{S})$ 个比特。并假设信源每 T_s 秒发送一次信号。因此,信源的平均信息速率为 $H(\mathcal{S})/T_s$ bps。译码器从信源字符集 \mathcal{S} 中将译码符号传送到目的地,且速率与信源速率相同。离散无记忆信道的信道容量等于 C 比特/信道使用。假设信道能够每 T_c 秒被使用一次,则单位时间的信道容量为 C/T_c bps,这表示信道中信息传输的最大速率。下面就给出香农第二定理的定义,也就是通常所说的信道编码定理。

离散无记忆信道的信道编码定理^[8]可表述为如下两个部分:

(i) 令一个具有字符集 \mathcal{S} 的离散无记忆信源的熵为 $H(\mathcal{S})$,且信源每 T_s 秒发送一次信号。

假设离散无记忆信道具有容量 C ,且每 T_c 秒被使用一次。如果有

$$\frac{H(\mathcal{S})}{T_s} \leq \frac{C}{T_c} \quad (9.61)$$

则存在一个编码方式,使得信源输出能够在信道中进行传输,且能在任意小的差错概率下进行重构。参数 C/T_c 称为临界速率。当式(9.61)取等号时,一般认为系统是在临界速率上发射信号。

(ii) 相反地,如果有

$$\frac{H(\mathcal{S})}{T_s} > \frac{C}{T_c}$$

则无法在信道上传输信息,也不能在任意小的差错概率下重构信息。

信道编码定理是信息论的一个最重要的结论。定理指出,信道容量 C 是在离散无记忆信道上可靠无误地传输消息的基本速率限制。但是,需要注意以下两点:

- 信道编码定理没有指出如何构造一个好的编码。因此该定理只是一个存在性的证明,即如果满足式(9.61)的条件,就存在较好的编码。(在第 10 章中我们将讨论若干离散无记忆信道的较好编码。)
- 该定理没有给出信道输出译码后的符号差错概率的精确结果,而仅指出,在满足式(9.61)的情况下,随着码长的增加,符号差错概率将趋近于 0。

值得注意的是,该定理其实隐含了功率和带宽限制。这两个系统限制实际上就出现在离

散无记忆信道的信道矩阵 \mathbf{P} 中。将例 9.5 的二进制信道的结果与 5.3 节中 PCM 接收机的噪声分析相结合, 可以很容易地验证这个论断。

9.8.1 信道编码定理在二进制对称信道中的应用

考虑一个离散无记忆信源, 该信源每 T_s 秒等概地发送二进制符号 0 或 1。由于信源的熵等于 1 比特/信源符号(见例 9.1), 信源的信息速率为 $(1/T_s)$ bps。将信源序列作为码率为 r 的信道编码器的输入。信道编码器每 T_c 秒发送一个符号。所以, 编码符号的传输速率是 $1/T_c$ 符号/秒。信道编码器每 T_c 秒将编码符号送入二进制对称信道一次。故单位时间的信道容量为 (C/T_c) bps, 其中, C 是根据式(9.60), 由预设的信道转移概率 p 确定的。信道编码定理的(i)部分表明, 如果

$$\frac{1}{T_s} \leq \frac{C}{T_c} \quad (9.62)$$

则通过使用适合的信道编码方式, 可使得差错概率任意小。而 T_c/T_s 等于信道编码器的码率, 即

$$r = \frac{T_c}{T_s} \quad (9.63)$$

因此, 可将式(9.62)简单表示为

$$r \leq C \quad (9.64)$$

对于 $r \leq C$, 存在一种编码(码率小于或等于 C), 使系统能够获得任意小的差错概率。

例 9.6 重复码

本例将给出信道编码定理的图解。通过一种简单的编码方式, 我们可对该定理有更多的发现。

首先考虑一个转移概率为 $p = 10^{-2}$ 的二进制对称信道。根据这个 p 值, 由式(9.60), 可得到信道容量 $C = 0.9192$ 。由信道编码定理, 可认为: 对于任何 $\epsilon > 0$ 和 $r \leq 0.9192$, 存在一个具有足够大的长度为 n 且码率为 r 的编码, 以及一个相应的译码算法, 使得编码比特流在给定信道上进行传送时, 信道译码的平均差错概率小于 ϵ 。此结果如图 9.12 所示, 其中极限值 $\epsilon = 10^{-8}$ 。

为了说明这个结果的重要性, 接下来讨论一个简单的编码方式, 其中包含了重复码的使用, 即消息的每个比特重复出现多次。假设每个比特(0 或 1)重复 n 次, 其中, $n = 2m + 1$, 且 n 为奇数。例如, 对于 $n = 3$, 0 和 1 对应的编码分别为 000 和 111。直观上不难看出, 采用多数准则译码较为合理。其步骤如下: 如果 n 个接收比特(即消息的一个比特)构成的分组中, 0 的个数大于 1 的个数, 译码器就判决为 0, 否则判决为 1。所以, 当 $n = 2m + 1$ 比特中的 $m + 1$ 或更多的比特接收不正确时, 就会出错。假设信道具有对称性, 平均差错概率 P_e 独立于 0 和 1 的先验概率。因此有(见习题 9.24):

$$P_e = \sum_{i=m+1}^n \binom{n}{i} p^i (1-p)^{n-i} \quad (9.65)$$

其中, p 为信道的转移概率。

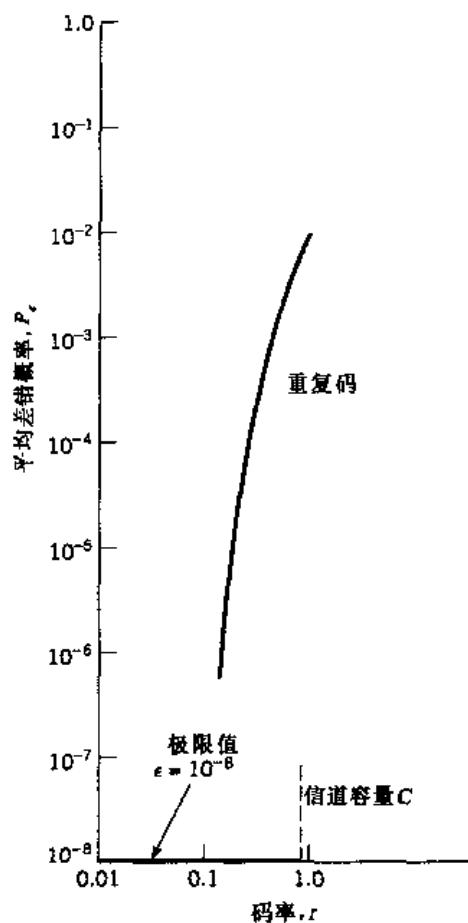


图 9.12 信道编码定理的重要性

表 9.3 给出了重复码的平均差错概率 P_e ，这是利用式(9.65)对不同的码率 r 进行计算得到的。假设信道为转移概率 $p = 10^{-2}$ 的二进制对称信道。表 9.3 所示的可靠性的提高是以降低码率为代价的。此结论也可由图 9.12 所示的标有“重复码”的曲线得出。该曲线表明，以码率换取消息可靠性是重复码的一个特征。

表 9.3 重复码的平均差错概率

码率, $r = 1/n$	平均差错概率, P_e
1	10^{-2}
$\frac{1}{3}$	3×10^{-4}
$\frac{1}{5}$	10^{-6}
$\frac{1}{7}$	4×10^{-7}
$\frac{1}{9}$	10^{-8}
$\frac{1}{11}$	5×10^{-10}

本例使我们对信道编码定理有了更深的认识。即该定理只要求码率小于信道容量 C ，而不要求通过使码率 r 趋近于 0 (如重复码的例子中) 来获得越来越高的通信链路的可靠性。

9.9 连续信号的相对熵和互信息

以上对信息论的讨论中所涉及的信源和信道包括所有具有离散幅度的随机变量。本节将把某些概念扩展至连续随机变量和随机矢量的情况。这样做的原因是找到了信息论中另一个基本限制的描述方法，这个问题将在 9.10 节中讨论。

考虑一个概率密度函数为 $f_X(x)$ 的连续随机变量 X 。类似于离散随机变量的熵，我们引入如下的定义：

$$b(X) = \int_{-\infty}^{\infty} f_X(x) \log_2 \left[\frac{1}{f_X(x)} \right] dx \quad (9.66)$$

我们将 $b(X)$ 称为 X 的相对熵，以区别于普通的熵（或绝对熵）。这样做的原因是，虽然 $b(X)$ 是一个有用的数学量，但却并不是 X 随机性的度量。将连续随机变量 X 看成是离散随机变量的极限形式，假设 $x_k = k\Delta x$ ，其中， $k = 0, \pm 1, \pm 2, \dots$ ，且 Δx 趋近于 0。根据定义，连续随机变量 X 以概率 $f_X(x_k)\Delta x$ 在区间 $[x_k, x_k + \Delta x]$ 上取值。令 Δx 趋近于 0，连续随机变量 X 的普通熵可用极限形式表示如下：

$$\begin{aligned} H(X) &= \lim_{\Delta x \rightarrow 0} \sum_{k=-\infty}^{\infty} f_X(x_k) \Delta x \log_2 \left(\frac{1}{f_X(x_k) \Delta x} \right) \\ &= \lim_{\Delta x \rightarrow 0} \left[\sum_{k=-\infty}^{\infty} f_X(x_k) \log_2 \left(\frac{1}{f_X(x_k)} \right) \Delta x - \log_2 \Delta x \sum_{k=-\infty}^{\infty} f_X(x_k) \Delta x \right] \\ &= \int_{-\infty}^{\infty} f_X(x) \log_2 \left(\frac{1}{f_X(x)} \right) dx - \lim_{\Delta x \rightarrow 0} \log_2 \Delta x \int_{-\infty}^{\infty} f_X(x) dx \\ &= b(X) - \lim_{\Delta x \rightarrow 0} \log_2 \Delta x \end{aligned} \quad (9.67)$$

其中最后一行应用了式(9.66)以及概率密度函数 $f_X(x)$ 曲线下的总面积等于 1 的事实。当 Δx 趋近于 0 时， $-\log_2 \Delta x$ 趋近于无穷。这意味着连续随机变量的熵为无穷大。直观上，此结论有可能是正确的，因为连续随机变量可以在区间 $(-\infty, \infty)$ 任意取值，变量的不确定性趋近于无穷。为避免涉及到 $-\log_2 \Delta x$ ，我们将 b 作为相对熵， $-\log_2 \Delta x$ 作为参考值。此外，由于信道上传输的信息实际上是两个具有共同参考的熵之差，因此信息等于相应相对熵之差，这就很好地证明，将式(9.66)定义的 $b(X)$ 作为连续随机变量 X 的相对熵是合理的。

若连续随机矢量 \mathbf{X} 由 n 个随机变量 X_1, X_2, \dots, X_n 构成，则可定义 \mathbf{X} 的相对熵为如下的 n 次积分：

$$b(\mathbf{X}) = \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x}) \log_2 \left[\frac{1}{f_{\mathbf{X}}(\mathbf{x})} \right] d\mathbf{x} \quad (9.68)$$

其中， $f_{\mathbf{X}}(\mathbf{x})$ 为 \mathbf{X} 的联合概率密度函数。

例 9.7 均匀分布

假设随机变量 X 在区间 $(0, a)$ 上均匀分布，其概率密度函数为

$$f_X(x) = \begin{cases} \frac{1}{a}, & 0 < x < a \\ 0, & \text{其他} \end{cases}$$

应用式(9.66)可得

$$\begin{aligned} h(X) &= \int_0^a \frac{1}{a} \log(a) dx \\ &= \log a \end{aligned} \quad (9.69)$$

注意, $a < 1, \log a < 0$ 。这个例子说明, 不同于离散随机变量, 连续随机变量的相对熵可以为负。

例 9.8 高斯分布

考虑随机变量 X 和 Y , 其概率密度函数分别由 $f_Y(x)$ 和 $f_X(x)$ 表示, 其中, x 只是一个虚构的变量。应用式(9.12)的基本不等式, 有^[9]

$$\int_{-\infty}^{\infty} f_Y(x) \log_2 \left(\frac{f_X(x)}{f_Y(x)} \right) dx \leq 0 \quad (9.70)$$

或等价地有

$$-\int_{-\infty}^{\infty} f_Y(x) \log_2 f_Y(x) dx \leq -\int_{-\infty}^{\infty} f_Y(x) \log_2 f_X(x) dx \quad (9.71)$$

式(9.71)中不等号左边的参量是随机变量 Y 的相对熵, 因此有

$$h(Y) \leq -\int_{-\infty}^{\infty} f_Y(x) \log_2 f_X(x) dx \quad (9.72)$$

假设随机变量 X 和 Y 如下所述:

- 随机变量 X 和 Y 具有相同的均值 μ 和方差 σ^2 。
- 随机变量 X 是高斯分布的, 即

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (9.73)$$

将式(9.73)代入式(9.72), 并将对数运算的底由 2 换为 $e = 2.7183$, 可得

$$h(Y) \leq -\log_2 e \int_{-\infty}^{\infty} f_Y(x) \left(-\frac{(x-\mu)^2}{2\sigma^2} - \log(\sqrt{2\pi}\sigma) \right) dx \quad (9.74)$$

不难发现, 随机变量 Y 具有如下性质(假设其均值为 μ , 方差为 σ^2):

$$\begin{aligned} \int_{-\infty}^{\infty} f_Y(x) dx &= 1 \\ \int_{-\infty}^{\infty} (x-\mu)^2 f_Y(x) dx &= \sigma^2 \end{aligned}$$

从而可将式(9.74)简化为

$$h(Y) \leq \frac{1}{2} \log_2(2\pi e \sigma^2) \quad (9.75)$$

式(9.75)中不等式右边的部分实际上是高斯随机变量 X 的相对熵

$$h(X) = \frac{1}{2} \log_2(2\pi e \sigma^2) \quad (9.76)$$

最后,结合式(9.75)和式(9.76)可得

$$h(Y) \leq h(X), \quad \begin{cases} X: \text{高斯随机变量} \\ Y: \text{另一随机变量} \end{cases} \quad (9.77)$$

当且仅当 $Y = X$ 时等号成立。

可将这个重要例子的结果概括为高斯随机变量熵的两个性质:

1. 对于有限方差 σ^2 , 高斯随机变量可获得最大的相对熵。
2. 高斯随机变量 X 的熵由 X 的方差唯一确定(即与 X 的均值无关)。

实际上,由于具有性质 1, 高斯信道模型在数字通信系统的研究中被广泛用为保守模型。

9.9.1 互信息

接下来讨论连续随机变量对 X 和 Y 。类似于式(9.47), 定义随机变量 X 和 Y 之间的互信息如下:

$$I(X; Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \log_2 \left[\frac{f_{X,Y}(x, y)}{f_X(x) f_Y(y)} \right] dx dy \quad (9.78)$$

其中, $f_{X,Y}(x, y)$ 为 X 和 Y 的联合概率密度函数, $f_X(x|y)$ 为给定 $Y = y$ 时, X 的条件概率密度函数。同样, 类似于式(9.45), 式(9.50)和式(9.44), 不难发现互信息 $I(X; Y)$ 具有以下的性质:

$$1. I(X; Y) = I(Y; X) \quad (9.79)$$

$$2. I(X; Y) \geq 0 \quad (9.80)$$

$$\begin{aligned} 3. I(X; Y) &= h(X) - h(X|Y) \\ &= h(Y) - h(Y|X) \end{aligned} \quad (9.81)$$

$h(X)$ 为 X 的相对熵, $h(Y)$ 则为 Y 的相对熵。 $h(X|Y)$ 为给定 Y 时, X 的条件相对熵。如式(9.41), 该条件相对熵由二次积分定义为

$$h(X|Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) \log_2 \left[\frac{1}{f_X(x|y)} \right] dx dy \quad (9.82)$$

$h(Y|X)$ 是给定 X 时, Y 的条件相对熵, 其定义与 $h(X|Y)$ 相同。

9.10 信息容量定理

在本节中, 我们将利用互信息的概念来阐明带宽受限、功率受限的高斯信道的信息容量定理。设一个带限为 B 赫的零均值平稳过程 $X(t)$ 。用 X_k 表示以每秒 $2B$ 个抽样点的奈奎斯特速率对随机过程 $X(t)$ 进行均匀抽样得到的连续随机变量, 其中, $k = 1, 2, \dots, K$ 。这些抽样点在 T 秒内通过一个同样带限为 B 赫的噪声信道。因此, 样点的个数 K 可由下式给出:

$$K = 2BT \quad (9.83)$$

X_k 称为发射信号的样本。信道输出受到均值为零、功率谱密度为 $N_0/2$ 的加性高斯白噪声的

干扰。噪声的带限为 B 赫。用连续随机变量 $Y_k, k = 1, 2, \dots, K$ 表示接收信号的抽样点：

$$Y_k = X_k + N_k, \quad k = 1, 2, \dots, K \quad (9.84)$$

噪声样本 N_k 为零均值且具有如下方差的高斯噪声：

$$\sigma^2 = N_0 B \quad (9.85)$$

假设抽样点 $Y_k, k = 1, 2, \dots, K$ 是统计独立的。

噪声和接收信号分别如式(9.84)和式(9.85)所述的信道称为离散时间无记忆高斯信道。其模型如图 9.13 所示。首先为每个信道输入分配一个代价值。发射机通常是功率受限的，因此此代价值的合理定义为

$$E[X_k^2] = P, \quad k = 1, 2, \dots, K \quad (9.86)$$

其中， P 为平均发射功率。此处描述的功率受限的高斯信道在理论和实际应用中都是非常重要的，因为它是许多通信信道的模型，其中包括可视无线电和卫星链路。

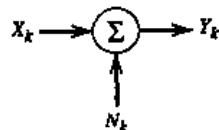


图 9.13 离散时间无记忆高斯信道的模型

信道信息容量的定义为：在输入 X_k 的所有分布都满足式(9.86)的功率限制的情况下，信道输入 X_k 和信道输出 Y_k 之间的互信息的最大值。用 $I(X_k; Y_k)$ 表示 X_k 与 Y_k 之间的互信息，可将信道的信息容量定义为

$$C = \max_{f_{X_k}(x)} \{I(X_k; Y_k) : E[X_k^2] = P\} \quad (9.87)$$

其中，最大化是关于 X_k 的概率密度函数 $f_{X_k}(x)$ 进行的。

互信息 $I(X_k; Y_k)$ 可表示为式(9.81)所示的两种等价形式中的任意一种。此处采用第二行的表示方法，即

$$I(X_k; Y_k) = h(Y_k) - h(Y_k | X_k) \quad (9.88)$$

由于 X_k 和 N_k 是独立随机变量，两者之和为 Y_k ，如式(9.84)所示。当给定 X_k 时， Y_k 的条件相对熵等于 N_k 的相对熵（见习题 9.28），即

$$h(Y_k | X_k) = h(N_k) \quad (9.89)$$

因此，式(9.88)可写为

$$I(X_k; Y_k) = h(Y_k) - h(N_k) \quad (9.90)$$

由于 $h(N_k)$ 独立于 X_k ，由式(9.87)可知，要使 $I(X_k; Y_k)$ 最大化，就必须使接收信号样本 Y_k 的相对熵 $h(Y_k)$ 最大化。因此 Y_k 必须是高斯随机变量（见例 9.8）。即接收信号的样本表示一个噪声型的过程。由于 N_k 为高斯过程，发射信号的样本 X_k 也必须是高斯过程。因此，式(9.87)所描述的最大化，可通过从平均功率为 P 的噪声型过程中选取发射信号的样点而得到。相应地，可将式(9.87)表示为

$$C = I(X_k; Y_k) : X_k \text{ 高斯}, \quad E[X_k^2] = P \quad (9.91)$$

其中,互信息 $I(X_k; Y_k)$ 是按照式(9.90)定义的。

信息容量 C 的计算可按照下面三个步骤进行:

1. 接收信号样本 Y_k 的方差等于 $P + \sigma^2$ 。因此,利用式(9.76)可得 Y_k 的相对熵为

$$h(Y_k) = \frac{1}{2} \log_2[2\pi e(P + \sigma^2)] \quad (9.92)$$

2. 噪声样本 N_k 的方差等于 σ^2 。因此,利用式(9.76)可得 N_k 的相对熵为

$$h(N_k) = \frac{1}{2} \log_2(2\pi e\sigma^2) \quad (9.93)$$

3. 将式(9.92)和式(9.93)代入式(9.90),使用式(9.91)给出的信息容量的定义,可得到所需的结果

$$C = \frac{1}{2} \log_2\left(1 + \frac{P}{\sigma^2}\right) \quad \text{比特/传输} \quad (9.94)$$

在 T 秒内 K 次使用信道以传输过程 $X(t)$ 的 K 个样点,可发现每单位时间的信息容量为式(9.94)所示结果的 K/T 倍。 K 等于 $2BT$,如式(9.83)所示。因此,可将信息容量表示为如下的等价形式:

$$C = B \log_2\left(1 + \frac{P}{N_0 B}\right) \text{ bps} \quad (9.95)$$

此处应用了式(9.85),其中,噪声方差为 σ^2 。

基于式(9.95),可将最著名的香农第三定理,即信息容量定理^[10]表述如下:

受到噪声功率谱密度为 $N_0/2$ 且带限为 B 的加性高斯白噪声干扰的、带宽为 B 赫的连续信道的信息容量由下式给出:

$$C = B \log_2\left(1 + \frac{P}{N_0 B}\right) \text{ bps}$$

其中, P 为平均发射功率。

信息容量定理是信息论中最重要的结论,因为它用一个公式高度总结了三个关键的系统参量:信道带宽、平均发射功率(或平均接收信号功率)和信道输出的噪声功率谱密度之间的关系。信息容量 C 与信道带宽 B 呈线性关系,而与信噪比 $P/N_0 B$ 成对数关系。因此,在指定的噪声方差下,扩大信道带宽比提高发射功率更容易提高通信信道的信息容量。

如式(9.95)所定义,上述定理表明,在给定的平均发射功率 P 和信道带宽 B 下,能够以每秒 C 比特的速率发送信息,并通过采用足够复杂的编码系统来获得任意小的差错概率。编码系统不可能以超过 C bps 的速率发送信息而没有一定的差错概率。所以,信道容量理论定义了一个功率受限和带宽受限的高斯信道的无误传输速率的基本极限。为了达到这个极限,发射信号必须具有与高斯白噪声相近的统计特性。

9.10.1 球体打包^[11]

下面提供一个理论以验证信息容量定理。假设采用某编码方式得到 K 个码字,每个码字

对应于发射信号的一个样本。用 n 表示每个码字的长度(即比特数)。假设选定的编码方式能够提供一个可接受的符号差错概率。并且码字满足如下功率限制:每个 n 比特码字的传输所包含的平均功率等于 nP ,其中 P 为每比特的平均功率。

假设编码中的任一码字被发送。 n 比特的接收矢量是高斯分布的,其均值与发送码字相同,方差等于 $n\sigma^2$,其中 σ^2 为噪声方差。接收矢量以很高的概率位于以 $\sqrt{n\sigma^2}$ 为半径、发送码字为中心的球体内。这个球体本身包含于半径为 $\sqrt{n(P + \sigma^2)}$ 的更大的球体中,其中 $n(P + \sigma^2)$ 为接收矢量的平均功率。

图 9.14 就是此图的形象化表示。由于半径为 $\sqrt{n\sigma^2}$ 的球体全部都分配给球心处的码字,因此当这个特殊码字被发送时,接收矢量将位于正确“译码”球体内的概率是很大的。关键问题是,有多少译码球体可以被包含在接收矢量的大球体内?换言之,实际可选择的码字有多少?下面我们就来回答这个问题。首先将半径为 r 的 n 维球体的体积表示为 $A_n r^n$,其中 A_n 为比例因子。因此可得出如下论断:

- 接收矢量球体的体积为 $A_n [n(P + \sigma^2)]^{n/2}$ 。
- 译码球体的体积为 $A_n (n\sigma^2)^{n/2}$ 。

因此,能够被包含到接收矢量球体内的非交叉译码球体的最大个数为

$$\frac{A_n [n(P + \sigma^2)]^{n/2}}{A_n (n\sigma^2)^{n/2}} = \left(1 + \frac{P}{\sigma^2}\right)^{n/2} = 2^{(n/2) \log_2(1 + P/\sigma^2)} \quad (9.96)$$

将此结果对 2 取对数后,不难发现,低差错概率下每传输的最大比特数实际上是由前面的式(9.94)定义的。

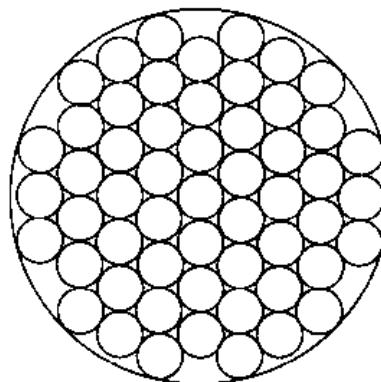


图 9.14 球体打包问题

例 9.9 衰减功率星座图的重构

为了说明球体打包的思想,考虑图 9.15(a)所示的 64-QAM 正方形星座图。此图显示了星座图中以消息点为中心的二维非交叉译码球体。在维持消息点间欧氏距离的同时,尽可能紧密地包含这些译码球体,就可以得到图 9.15(b)所示的另一个星座图。由于消息点之间具有相同的欧式距离,假设在 AWGN 信道上维持足够高的信噪比,则图 9.15 所示的两种星座图将提供相同的误比特率,见式(5.95)。比较这两种星座图,不难发现,图 9.15(b)中消息点到原点距离的平方和比图 9.15(a)中要小。所以,图 9.15(b)所示的紧

密压缩的星座图与图 9.15(a)所示的星座图相比,具有更大的优势,即在 AWGN 信道中,同样的误比特率对应的发送符号平均能量更小。

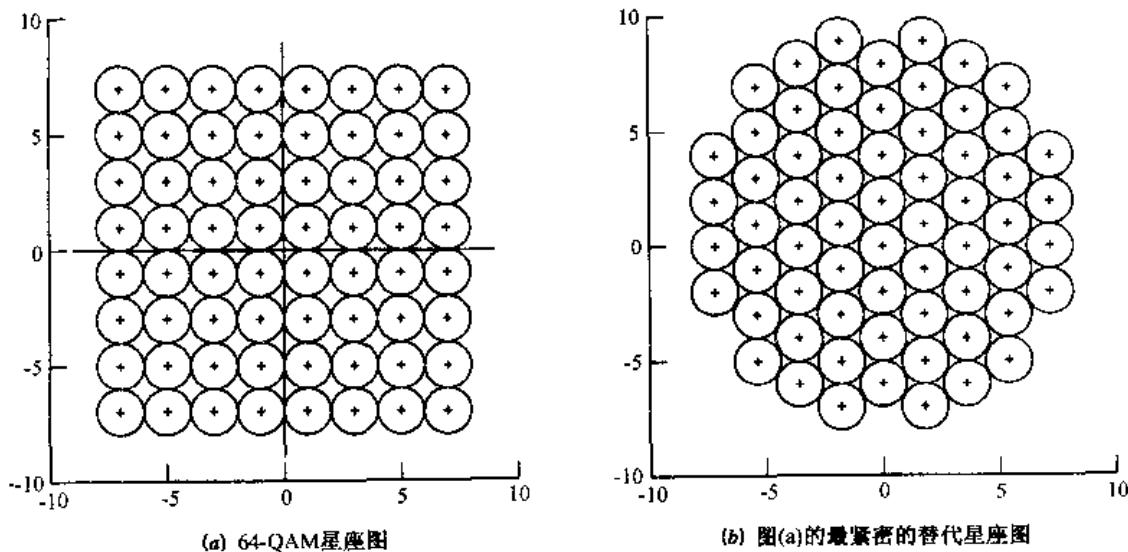


图 9.15 衰减功率星座图

9.11 信息容量定理的含义

对信息容量定理有了直观的了解,下面就可以继续讨论它在功率和带宽受限的高斯信道上的应用了。因此,首先需要一个可以用来评估实际通信系统性能的理想框架。为此,引入一个理想系统,即以等于信息容量 C 的比特速率 R_b 传输数据的系统。这个系统的平均发射功率可定义:

$$P = E_b C \quad (9.97)$$

其中, E_b 为每比特信号能量。因此,理想系统可由下式定义

$$\frac{C}{B} = \log_2 \left(1 + \frac{E_b}{N_0} \frac{C}{B} \right) \quad (9.98)$$

等价地,可根据理想系统的 C/B 值来定义每比特信号能量与噪声功率谱密度 E_b/N_0 的比值:

$$\frac{E_b}{N_0} = \frac{2^{C/B} - 1}{C/B} \quad (9.99)$$

带宽效率 R_b/B 对 E_b/N_0 的曲线称为带宽效率图。这种曲线的大致形式如图 9.16 所示,其中,标有“容量边界”的曲线对应于 $R_b = C$ 的理想系统。基于图 9.16,可得到如下结论:

- 对于无限带宽, E_b/N_0 趋近于极限值

$$\begin{aligned} \left(\frac{E_b}{N_0} \right)_\infty &= \lim_{B \rightarrow \infty} \left(\frac{E_b}{N_0} \right) \\ &= \log 2 = 0.693 \end{aligned} \quad (9.100)$$

该极限值称为码率为零时 AWGN 信道的香农极限值, 即 -1.6 dB 。相应的信道容量极限值可在式(9.95)中令信道带宽 B 趋近于无穷得到。因此有

$$\begin{aligned} C_{\infty} &= \lim_{B \rightarrow \infty} C \\ &= \frac{P}{N_0} \log_2 e \end{aligned} \quad (9.101)$$

其中, e 为自然对数的底。

2. 容量边界由临界比特率 $R_b = C$ 对应的曲线定义。该边界的作用是区分 $R_b < C$ (支持无误传输) 和 $R_b > C$ (不支持无误传输) 两种情况。后一种情况对应的区域如图 9.16 中的阴影部分所示。
3. 此图间接地强调了 E_b/N_0 , R_b/B 和符号差错概率 P_e 之间的权衡。特别地, 对于固定的 R_b/B , 可将工作点沿着水平线的移动视为 P_e 与 E_b/N_0 之间的折中。另一方面, 对于固定的 E_b/N_0 , 可将工作点沿垂直线的移动视为 P_e 与 R_b/B 之间的折中。

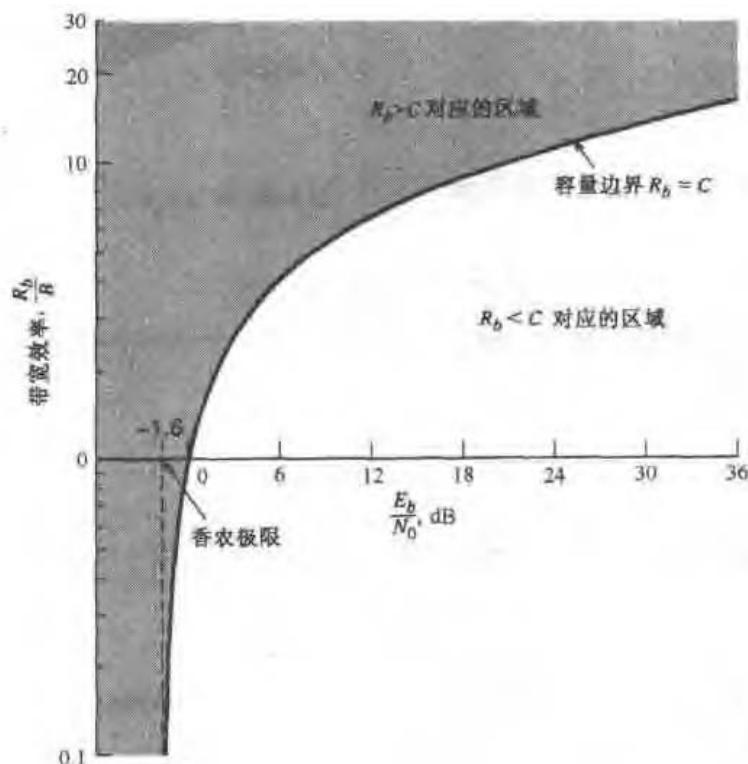


图 9.16 带宽效率图

例 9.10 M 进制 PCM

本例中将根据信道容量定理研究 M 进制 PCM 系统。假设系统工作在错误门限之上, 也就是说, 由信道噪声产生的平均差错概率可以忽略。

假设 M 进制 PCM 系统采用由 n 个码元构成的码字, 其中每个码元的取值为 M 个可能的离散幅度电平之一, 因此称为“ M 进制”。第 3 章曾提到, 为使 PCM 系统工作在差错门限之上, 必须预留足够大的噪声容限, 从而使信道噪声引起的差错概率可以忽略。这就意味着 M 个离散幅度电平之间必须有一定间隔。令该间隔为 $k\sigma$, 其中, k 为常数, $\sigma^2 = N_0 B$

是信道带宽为 B 的噪声方差。幅度电平个数 M 通常是 2 的整数次方。如果幅度区间关于 0 对称，则平均发射功率为最小。关于间隔归一化的离散幅度电平将取值 $\pm 1/2, \pm 3/2, \dots, \pm (M-1)/2$ 。假设这 M 个不同的幅度电平是等概的。由此可得平均发射功率为

$$\begin{aligned} P &= \frac{2}{M} \left[\left(\frac{1}{2}\right)^2 + \left(\frac{3}{2}\right)^2 + \dots + \left(\frac{M-1}{2}\right)^2 \right] (k\sigma)^2 \\ &= k^2 \sigma^2 \left(\frac{M^2 - 1}{12} \right) \end{aligned} \quad (9.102)$$

假设将此处描述的 M 进制 PCM 系统用于传输最高频率分量等于 W 赫的消息信号。信号以每秒 $2W$ 个抽样点的奈奎斯特速率进行抽样。假设系统采用中升型的量化器，具有 L 个等概的量化级。因此， L 个量化级中任何一个的发生概率均为 $1/L$ 。相应地，信号的单个样点所携带的信息量为 $\log_2 L$ 比特。在每秒 $2W$ 个抽样点的最大抽样速率下，PCM 系统信息传输的最大速率（以 bps 度量）为

$$R_b = 2W \log_2 L \text{ bps} \quad (9.103)$$

由于 PCM 系统采用由 n 个码元构成的码字，其中每个码元都有 M 个可能的离散幅度值，因此得到 M^n 个可能的不同码字。对于一个编码过程，我们要求

$$L = M^n \quad (9.104)$$

很显然，系统的信息传输速率不受所选编码过程的影响。将式(9.103)和式(9.104)联立并约掉 L 可得

$$R_b = 2Wn \log_2 M \text{ bps} \quad (9.105)$$

式(9.102)定义了将 M 进制 PCM 系统维持在差错门限之上所需的平均发射功率。由此可得离散幅度电平的个数 M 为

$$M = \left(1 + \frac{12P}{k^2 N_0 B} \right)^{1/2} \quad (9.106)$$

其中， $\sigma^2 = N_0 B$ 为以信道带宽 B 衡量的噪声方差。将式(9.106)代入式(9.105)可得

$$R_b = Wn \log_2 \left(1 + \frac{12P}{k^2 N_0 B} \right) \quad (9.107)$$

传输持续时间为 $1/2nW$ （表示码字中的一个码元）的矩形脉冲信号所需要的信道带宽由下式给出（参见第 3 章）：

$$B = \kappa n W$$

其中 κ 为 1 到 2 之间的一个常量。当 κ 取最小可能的值 1 时，信道带宽 $B = nW$ 。于是式(9.107)可写为

$$R_b = B \log_2 \left(1 + \frac{12P}{k^2 N_0 B} \right) \quad (9.108)$$

理想系统由式(9.95)所示的香农信道容量定理描述。比较式(9.108)与式(9.95)可知，理想系统的平均发射功率是 PCM 系统的 $k^2/12$ 倍。由式(9.108)不难看出，PCM 系统的功率和带宽是以对数为基础进行交换的，信息容量 C 与信道带宽 B 成正比例。

例 9.11 M 进制 PSK 和 M 进制 FSK

本例将根据香农信息容量定理比较 M 进制 PSK 和 M 进制 FSK 信号的带宽 - 功率交换能力。首先考虑连续 M 进制 PSK 系统。为了进行二进制数据的传输, 该系统采用了一组 M 个非正交的 PSK 信号。每个信号表示一个 $\log_2 M$ 比特的符号。如式(6.51), 应用零 - 零带宽的定义, 可将 M 进制 PSK 的带宽效率表示如下:

$$\frac{R_b}{B} = \frac{\log_2 M}{2}$$

图 9.17(a)所示为相位等级个数 $M = 2, 4, 8, 16, 32, 64$ 的不同工作点。每个工作点对应的平均符号差错概率 $P_e = 10^{-5}$ 。图中还显示了理想系统的容量边界。由图 9.17(a)可知, 随着 M 的增加, 带宽效率也随之提高, 但无误传输所需要的 E_b/N_0 值就会偏离香农极限值。

接下来考虑连续 M 进制 FSK 系统。为了进行二进制数据的传输, 该系统采用了一组 M 个正交的 FSK 信号, 相邻信号频率之间的间隔等于 $1/2T$, 其中, T 为符号周期。与 M 进制 PSK 中类似, 此处每个信号表示一个 $\log_2 M$ 比特的符号。如式(6.143), M 进制 FSK 的带宽效率为

$$\frac{R_b}{B} = \frac{2 \log_2 M}{M}$$

图 9.17(b)所示为频率等级个数 $M = 2, 4, 8, 16, 32, 64$ 的不同工作点, 每个工作点对应的平均符号差错概率 $P_e = 10^{-5}$ 。图中还显示了理想系统的容量边界。由图 9.17(b)可知, M 的增加对(正交) M 进制 FSK 的影响与对(非正交) M 进制 PSK 的影响相反。随着 M 的增加(相当于带宽要求的提高), 工作点更加接近于香农极限值。

例 9.12 二进制输入加性高斯白噪声信道的容量

本例研究采用二进制正反信号编码(即电平 -1 和 +1 分别表示二进制符号 0 和 1)的 AWGN 信道的容量。其中, 对于不同的码率 r , 系统可达到的最小误比特率为 E_b/N_0 的函数。假设二进制符号 0 和 1 是等概出现的。

用随机变量 X 和 Y 分别表示信道输入和输出。 X 为离散变量, 而 Y 为连续变量。根据式(9.81)的第二行, 可将信道输入与输出之间的互信息表示为

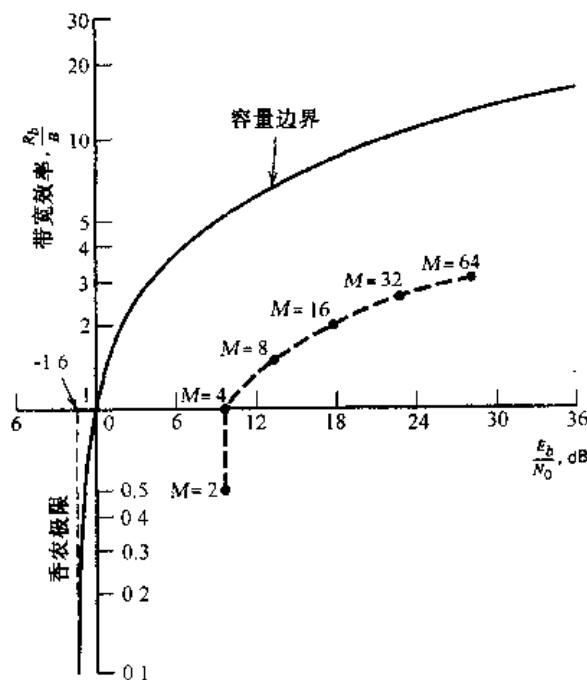
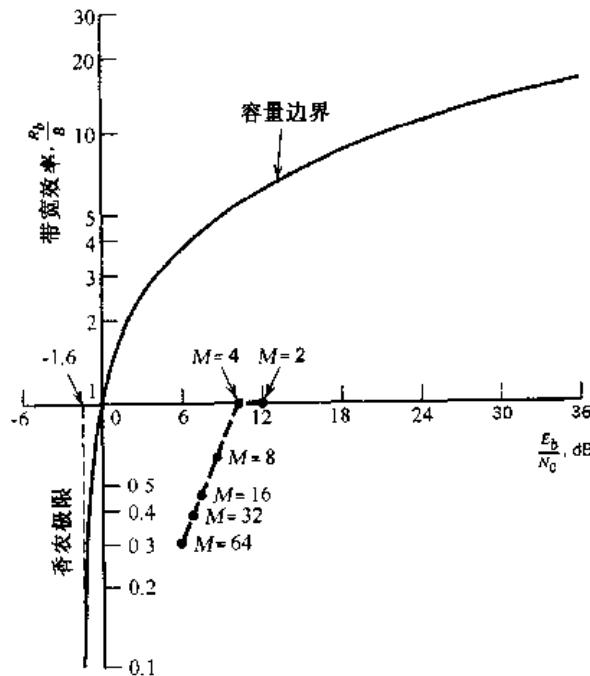
$$I(X; Y) = h(Y) - h(Y|X)$$

上式中第二项 $h(Y|X)$ 为给定信道输入 X 时, 信道输出 Y 的条件相对熵。由式(9.89)和式(9.93)可知, 这一项就是高斯分布的熵。用 σ^2 表示信道噪声的方差, 可将该条件相对熵表示为

$$h(Y|X) = \frac{1}{2} \log_2(2\pi e\sigma^2)$$

相应地, 上述互信息表达式中的第一项 $h(Y)$ 就是信道输出 Y 的相对熵。由于采用的是二进制正反信号, Y 的概率密度函数(当 $X = x$)可表示为具有相同方差 σ^2 且均值分别为 -1 和 +1 的两个高斯分布的混合, 即

$$f_Y(y_i|x) = \frac{1}{2} \left[\frac{\exp(-(y_i + 1)^2/2\sigma^2)}{\sqrt{2\pi}\sigma} + \frac{\exp(-(y_i - 1)^2/2\sigma^2)}{\sqrt{2\pi}\sigma} \right] \quad (9.109)$$

(a) 当 $P_e=10^{-5}$ 时, 随着 M 的增加, M 进制 PSK 系统与理想系统的比较(b) 当 $P_e=10^{-5}$ 时, 随着 M 的增加, M 进制 FSK 系统与理想系统的比较图 9.17 M 进制 PSK 系统、FSK 系统与理想系统的比较

可利用下面的公式来确定 Y 的相对熵：

$$h(Y) = - \int_{-\infty}^{\infty} f_Y(y_i | x) \log_2 [f_Y(y_i | x)] dy_i$$

其中, $f_Y(y_i | x)$ 由式(9.109)定义。从 $h(Y|X)$ 和 $h(Y)$ 的公式中, 可以很清楚地看出, 互信息只是噪声方差 σ^2 的函数。可用 $M(\sigma^2)$ 将这种函数关系表示为

$$I(X; Y) = M(\sigma^2)$$

遗憾的是, 由于难以确定 $h(Y)$, 也就无法推导出 $M(\sigma^2)$ 的近似表达式。但是, 相对熵 $h(Y)$ 可用蒙特卡罗积分(Monte Carlo integration)进行很好的近似。这种积分较容易用计算机程序进行表示, 见习题 9.36。

由于符号 0 和 1 是等概的, 因此信道容量 C 等于 X 与 Y 之间的互信息。所以, 对于 AWGN 信道上的无误数据传输, 码率 r 必须满足以下的条件:

$$r < M(\sigma^2) \quad (9.110)$$

E_b/N_0 的一个实用的度量是

$$\frac{E_b}{N_0} = \frac{P}{N_0 r} = \frac{P}{2\sigma^2 r}$$

其中, P 为平均发射功率, $N_0/2$ 为信道噪声的双边功率谱密度。为了不失一般性, 设 $P = 1$, 则噪声方差可以表示为

$$\sigma^2 = \frac{N_0}{2E_b r} \quad (9.111)$$

将式(9.111)代入式(9.110)并进行整理, 可得

$$\frac{E_b}{N_0} = \frac{1}{2rM^{-1}(r)} \quad (9.112)$$

其中, $M^{-1}(r)$ 为信道输入和输出之间的互信息的倒数, 可表示为码率 r 的函数。

用蒙特卡罗方法可估计相对熵 $h(Y)$ 和 $M^{-1}(r)$, 如图 9.18 所示。^[12] 图 9.18(a) 显示了对于无误通信, 最小 E_b/N_0 随码率 r 的变化曲线。图 9.18(b) 显示了在不同的 r 值下, 可获得的最小误比特率随 E_b/N_0 的变化趋势。

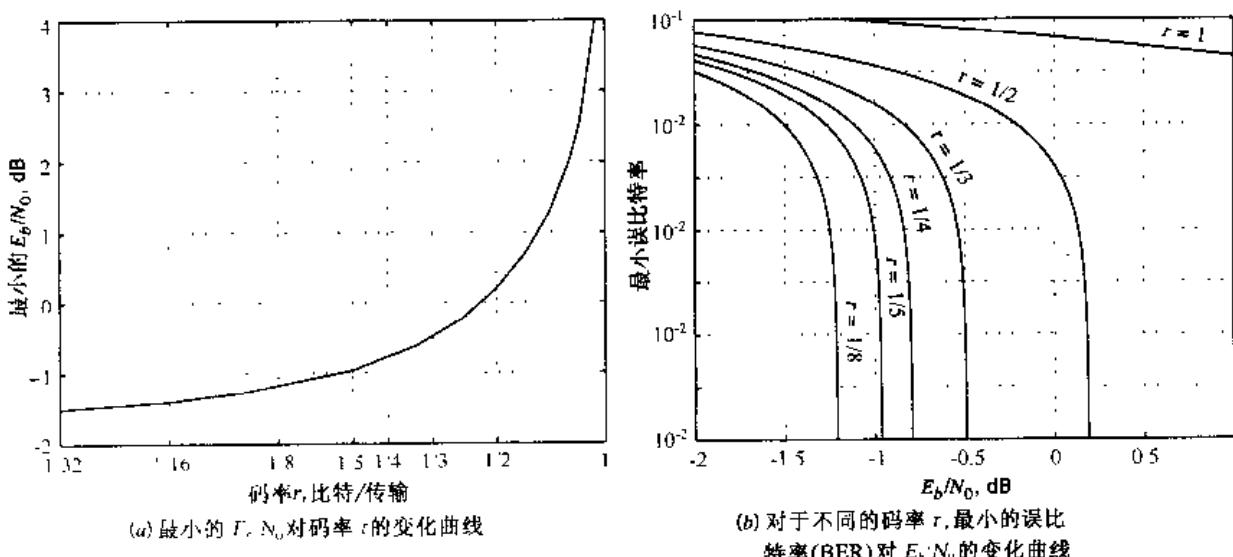


图 9.18 AWGN 信道上的二进制正反信号

由图 9.18 可得如下结论：

- 对于未编码的二进制信号 ($r = 1$)，无噪声通信要求无穷大的 E_b/N_0 ，这与我们所知道的 AWGN 信道上的未编码数据传输是一致的。
- 随着编码效率 r 的降低，最小 E_b/N_0 也随之降低，这与直观事实是相符的。例如，对于 $r = 1/2$, E_b/N_0 的最小值略小于 0.2 dB。
- 当 r 趋近于 0 时，最小 E_b/N_0 趋近于极限值 -1.6 dB，这与之前讨论的香农极限值是一致的，见式(9.100)。

9.12 有色噪声信道^[13]的信息容量

式(9.95)所示的信息容量定理适用于带宽受限的白噪声信道。在本节中，我们将香农信息容量定理推广到更一般的情况：非白色或有色噪声信道。考虑如图 9.19(a)所示的信道模型，其中，信道的转移函数由 $H(f)$ 表示。信道噪声 $n(t)$ 在信道输出端是加性的，其数学模型为零均值且功率谱密度为 $S_N(f)$ 的平稳高斯过程的样本函数。需要解决的问题有两个：

1. 找到由功率谱密度 $S_X(f)$ 描述的输入信号群。该输入信号群能够最大化信道输出 $y(t)$ 与输入 $x(t)$ 之间的互信息，且符合 $x(t)$ 的平均功率为固定值 P 的约束条件。
2. 确定信道的最优信息容量。

因此，要解决的就是最优化问题。解决办法如下：

- 由于信道是线性的，因此可将图 9.19(a)所示的模型替换为图 9.19(b)所示的模型。由信道输出端测得的信号加噪声的频谱特性来看，图 9.19 所示的两种模型是等价的，这里假设图 9.19(b)所示的噪声 $n'(t)$ 的功率谱密度可由图 9.19(a)中噪声 $n(t)$ 的功率谱密度定义为

$$S_{N'}(f) = \frac{S_N(f)}{|H(f)|^2} \quad (9.113)$$

其中 $|H(f)|$ 为信道的幅度响应。

- 为了简化分析，可采用类似于 6.12 节中所描述的分割和解决的原则，将信道分割为大量的相连频段，如图 9.20 所示。各子信道的频率间隔增量 Δf 越小，近似性就越好。

上述两点的结果是通过将图 9.19(a)所示的原始模型用有限 N 个子信道的并行连接替代得到的。其中每个子信道中都存在“带限高斯白噪声”。

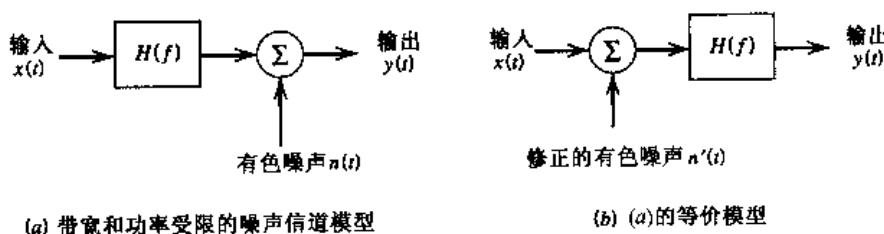
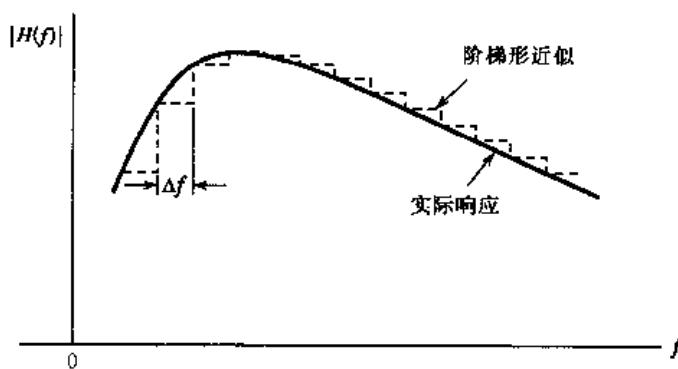


图 9.19 带宽和功率受限的噪声信道等价模型

图 9.20 幅度响应绝对值 $|H(f)|$ 的阶梯形近似; 只显示了响应的正频率部分

在对图 9.19(b)所示模型的近似表示中, 第 k 个子信道可描述为

$$y_k(t) = x_k(t) + n_k(t), \quad k = 1, 2, \dots, N \quad (9.114)$$

信号分量 $x_k(t)$ 的平均功率为

$$P_k = S_x(f_k) \Delta f, \quad k = 1, 2, \dots, N \quad (9.115)$$

其中, $S_x(f_k)$ 为 $f = f_k$ 时输入信号的功率谱密度。噪声分量 $n_k(t)$ 的方差为

$$\sigma_k^2 = \frac{S_N(f_k)}{|H(f_k)|^2} \Delta f, \quad k = 1, 2, \dots, N \quad (9.116)$$

其中, $S_N(f_k)$ 和 $|H(f_k)|$ 分别为频率 f_k 对应的噪声频谱密度和信道幅度响应。第 k 个子信道的信息容量为

$$C_k = \frac{1}{2} \Delta f \log_2 \left(1 + \frac{P_k}{\sigma_k^2} \right), \quad k = 1, 2, \dots, N \quad (9.117)$$

其中, 引入因子 $1/2$ 是因为考虑到 Δf 应用于正、负频率的不同情况。所有的 N 个子信道都是相互独立的。因此, 整个信道的总容量可由下面的求和近似给出:

$$\begin{aligned} C &\approx \sum_{k=1}^N C_k \\ &= \frac{1}{2} \sum_{k=1}^N \Delta f \log_2 \left(1 + \frac{P_k}{\sigma_k^2} \right) \end{aligned} \quad (9.118)$$

下面要考虑的是如何使整体信息容量 C 最大化, 并满足约束条件

$$\sum_{k=1}^N P_k = P = \text{常数} \quad (9.119)$$

解决有约束条件的最优化问题的常规方法是使用拉格朗日因子法(参考第 6 章注释 19)。如式(9.118)和式(9.119), 首先定义一个目标函数(该函数将信息容量 C 和约束条件相联系):

$$J = \frac{1}{2} \sum_{k=1}^N \Delta f \log_2 \left(1 + \frac{P_k}{\sigma_k^2} \right) + \lambda \left(P - \sum_{k=1}^N P_k \right) \quad (9.120)$$

其中, λ 为拉格朗日因子。接着, 将目标函数 J 对 P_k 求微分并使结果等于 0, 得到

$$\frac{\Delta f \log_2 e}{P_k + \sigma_k^2} - \lambda = 0$$

为了得到最优化的结果, 加入如下条件:

$$P_k + \sigma_k^2 = K \Delta f \quad k = 1, 2, \dots, N \quad (9.121)$$

其中, K 对于所有 k 是一个常数。 K 的选择满足平均功率约束。

将式(9.115)和式(9.116)的定义值加入到式(9.121)的最优化条件中, 简化并重新整理, 得到

$$S_X(f_k) = K - \frac{S_N(f_k)}{|H(f_k)|^2}, \quad k = 1, 2, \dots, N \quad (9.122)$$

用 \mathcal{F}_A 表示频率范围, 在这个范围内常数 K 满足如下条件:

$$K \geq \frac{S_N(f)}{|H(f)|^2}$$

当频率间隔增量 Δf 趋近于零, 且子信道的个数 N 也趋于无穷时, 由式(9.122)可知, 实现最优信息容量的输入信号群功率谱密度应当为一个非负的量, 其定义为

$$S_X(f) = \begin{cases} K - \frac{S_N(f)}{|H(f)|^2} & f \in \mathcal{F}_A \\ 0 & \text{其他} \end{cases} \quad (9.123)$$

由于随机过程的平均功率是其功率谱密度曲线下的全部面积, 因此可将信道输入 $x(t)$ 的平均功率表示为

$$P = \int_{f \in \mathcal{F}_A} \left(K - \frac{S_N(f)}{|H(f)|^2} \right) df \quad (9.124)$$

对于预设的 P 和给定的 $S_N(f)$ 和 $H(f)$, 常数 K 为式(9.124)的解。

最后一个问题是找到最优信息容量。将式(9.121)的最优解代入式(9.118), 并应用式(9.115)和式(9.116)的定义值, 可得到

$$C \approx \frac{1}{2} \sum_{k=1}^N \Delta f \log_2 \left(K \frac{|H(f_k)|^2}{S_N(f_k)} \right)$$

当频率间隔增量 Δf 趋于零时, 得到上式的极限形式

$$C = \frac{1}{2} \int_{-\infty}^{\infty} \log_2 \left(K \frac{|H(f)|^2}{S_N(f)} \right) df \quad (9.125)$$

其中, 常数 K 为在预设的输入信号功率 P 下式(9.124)的解。

9.12.1 信息容量定理的注水解释

式(9.123)和式(9.124)对应于图 9.21 所示的曲线。特别地, 由图可知:

- 适当的输入功率谱密度 $S_X(f)$ 对应于函数 $S_N(f)/|H(f)|^2$ 在常数 K 以下的底部区域, 如图中阴影所示。

- 输入功率 P 等于阴影区域的总面积。

图 9.21 所示的频谱区域图称为注水(灌水)解释,这是因为输入功率在函数 $S_N(f)/|H(f)|^2$ 上的分布与容器中水的分布方式相同。

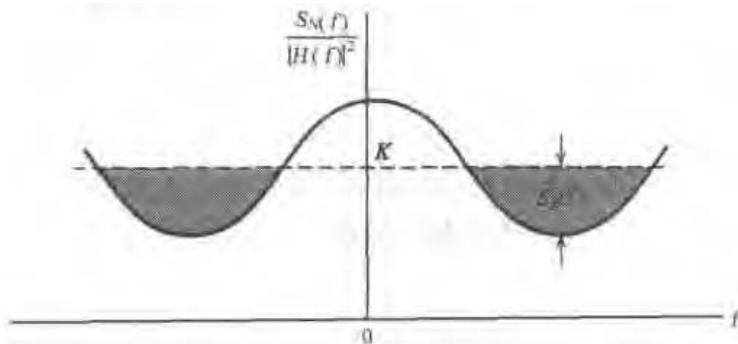


图 9.21 有色噪声信道下,信息容量定理的注水解释

考虑功率谱密度 $N(f) = N_0/2$ 的加性高斯白噪声中的带限信号的理想情况。理想带通滤波器的转移函数 $H(f)$ 为

$$H(f) = \begin{cases} 1, & 0 \leq f_c - \frac{B}{2} \leq |f| \leq f_c + \frac{B}{2} \\ 0, & \text{其他} \end{cases}$$

其中, f_c 为中频, B 为信道带宽。在这种特殊情况下, 式(9.124)和式(9.125)分别简化为

$$P = 2B\left(K - \frac{N_0}{2}\right)$$

和

$$C = B \log_2\left(\frac{2K}{N_0}\right)$$

约掉这两式中的 K , 就得到香农容量定理的标准形式, 如式(9.95)。

例 9.13 近端串话占优势的信道的容量

由 4.8 节可知, DSL 中的主要信道损耗就是近端串话。串话的功率密度可以表示为

$$S_N(f) = |H_{\text{NEXT}}(f)|^2 S_X(f) \quad (9.126)$$

其中, $S_X(f)$ 为发射信号的功率谱密度, $H_{\text{NEXT}}(f)$ 为耦合相邻双绞线的转移函数。这个例子中唯一的约束条件就是对于所有 f , 功率谱密度 $S_X(f)$ 是非负的。将式(9.126)代入式(9.123), 不难发现, 解得的 K 值能够满足上述约束条件:

$$K = \left(1 + \frac{|H_{\text{NEXT}}(f)|^2}{|H(f)|^2}\right) S_X(f)$$

最后, 利用式(9.125)得到的结果, 可得到所求的近端串话占优势的 DSL 信道的容量为

$$C = \frac{1}{2} \int_{\mathcal{F}_A} \log_2 \left(1 + \frac{|H(f)|^2}{|H_{\text{NEXT}}(f)|^2}\right) df$$

其中, \mathcal{F}_A 为使 $S_X(f) > 0$ 的一组正、负频率值。

9.13 率失真定理

根据离散无记忆信源的信源编码定理(见9.3节),平均码长必须至少和理想编码(即信源的理想表示)的信源熵一样大。但在实际情况下存在许多约束,使得编码变得不太理想,从而造成不可避免的失真。例如,对通信信道的约束可能使得可允许的码率及信息信源的平均码长存在上限。再比如,信息信源可能具有与语音信号相同的连续幅度,此时就要求量化由信源产生的每个样点的幅度,使得这些样点能够用有限长度的码字表示(与在PCM中相同)。在这种情况下,问题已涉及到保真度准则下的信源编码,与此相关的信息论的分支称为率失真理论^[14]。率失真理论可应用于下列情况:

- 允许编码字符集不准确表示信源的信源编码。在这种情况下,必须进行有损数据压缩。
- 以高于信道容量的速率进行信息传输。

因此,率失真理论可以认为是香农编码理论的自然延伸。

9.13.1 率失真函数

考虑一个由 M 进制字符集 $X: \{x_i | i = 1, 2, \dots, M\}$ 定义的离散无记忆信源。该信源由相应符号概率为 $\{p_i | i = 1, 2, \dots, M\}$ 的统计独立的符号组成。用 R 表示平均码率,单位为比特/码字。表示码字来自于另一个字符集 $Y: \{y_j | j = 1, 2, \dots, N\}$ 。信源编码定理说明,当 $R > H$ 时,字符集 Y 提供了信源的最佳表示,其中, H 为信源熵。但是,如果必须满足 $R < H$,就不可避免地会存在失真,从而造成信息的丢失。

以 $p(x_i, y_j)$ 表示(信源符号) x_i 和(表示符号) y_j 的联合概率。根据概率理论,有

$$p(x_i, y_j) = p(y_j | x_i)p(x_i) \quad (9.127)$$

其中, $p(y_j | x_i)$ 为转移概率。以符号 y_j 表示信源符号 x_i 的代价用 $d(x_i, y_j)$ 度量, $d(x_i, y_j)$ 称为单个字母失真的度量。在所有可能的信源符号上对 $d(x_i, y_j)$ 求统计平均可得

$$\bar{d} = \sum_{i=1}^M \sum_{j=1}^N p(x_i)p(y_j | x_i) d(x_i, y_j) \quad (9.128)$$

注意,平均失真 \bar{d} 是由信源编/译码器对决定的转移概率 $p(y_j | x_i)$ 的非负连续函数。

当且仅当平均失真 \bar{d} 小于或等于某些可接受的值 D 时,条件概率分布 $p(y_j | x_i)$ 被认为是“ D 允许”的。所有的“ D 允许”的条件概率分布可表示为

$$P_D = \{p(y_j | x_i) : \bar{d} \leq D\} \quad (9.129)$$

每组转移概率的互信息为

$$I(X; Y) = \sum_{i=1}^M \sum_{j=1}^N p(x_i)p(y_j | x_i) \log\left(\frac{p(y_j | x_i)}{p(y_j)}\right) \quad (9.130)$$

率失真函数 $R(D)$ 的定义为平均失真不超过 D 的最小可能码率。以 P_D 表示预设的 D 对应的

条件概率 $p(y_j|x_i)$ 的集合。对于一个固定的 D , 有^[15]

$$R(D) = \min_{p(y_j|x_i) \in P_D} I(X; Y) \quad (9.131)$$

满足约束条件

$$\sum_{j=1}^N p(y_j|x_i) = 1 \quad \text{for } i = 1, 2, \dots, M \quad (9.132)$$

如果将以 2 为底的对数运算应用到式(9.130), 就可以用比特来度量率失真函数 $R(D)$ 。直观上, 我们希望随着率失真函数 $R(D)$ 的增加, 失真 D 会减小。当然也可以反过来认为, 高的失真 D 使得系统可用较低的速率来编码和传输信息。

图 9.22 总结了率失真理论的主要参数。给定信源符号 $\{x_i\}$ 及其概率 $\{p_i\}$, 以及单个字母失真度量的定义 $d(x_i, y_j)$, 率失真函数 $R(D)$ 的计算过程就是寻找符合约束条件的条件概率分布 $p(y_j|x_i)$ 的过程。这是一个变化的问题, 遗憾的是, 其解答往往并不直观。

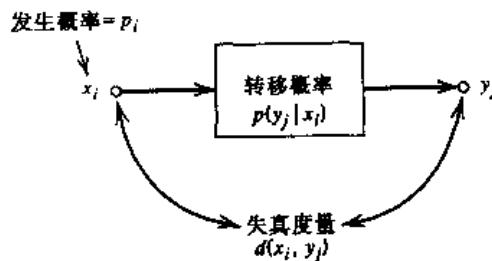


图 9.22 率失真理论的总结

例 9.14 高斯信源

考虑一个零均值且方差为 σ^2 的离散时间无记忆高斯信源。以 x 表示由信源产生的样点值。 y 表示 x 的量化形式, x 可以用有限表示法来表示。平方误差失真为

$$d(x, y) = (x - y)^2$$

上式提供了广泛用于连续字符集的失真度量。具有平均误差失真的高斯信源的率失真函数由下式给出:

$$R(D) = \begin{cases} \frac{1}{2} \log \left(\frac{\sigma^2}{D} \right), & 0 \leq D \leq \sigma^2 \\ 0, & D > \sigma^2 \end{cases} \quad (9.133)$$

在这种情况下, 不难看出: 当 $D \rightarrow 0$ 时, $R(D) \rightarrow \infty$; 当 $D = \sigma^2$ 时, $R(D) = 0$ 。

例 9.15 并行高斯信源集

下面考虑一组 N 个独立高斯随机变量 $\{X_i\}_{i=1}^N$, 其中, X_i 具有零均值和方差 σ_i^2 。应用失真度量

$$d = \sum_{i=1}^N (x_i - \hat{x}_i)^2$$

利用例 9.14 的结果, 可将一组并行高斯信源的率失真函数定义为

$$R(D) = \sum_{i=1}^N \frac{1}{2} \log\left(\frac{\sigma_i^2}{D_i}\right) \quad (9.134)$$

其中, D_i 的定义为

$$D_i = \begin{cases} \lambda & \lambda < \sigma_i \\ \sigma_i^2 & \lambda \geq \sigma_i^2 \end{cases} \quad (9.135)$$

常量 λ 的选择应满足下面的条件:

$$\sum_{i=1}^N D_i = D \quad (9.136)$$

式(9.135)和式(9.136)可解释为一种“反向的注水”,如图 9.23 所示。首先选择一个常量 λ 和方差大于 λ 的随机变量子集。对方差小于 λ 的其余随机变量的子集不用比特进行描述。

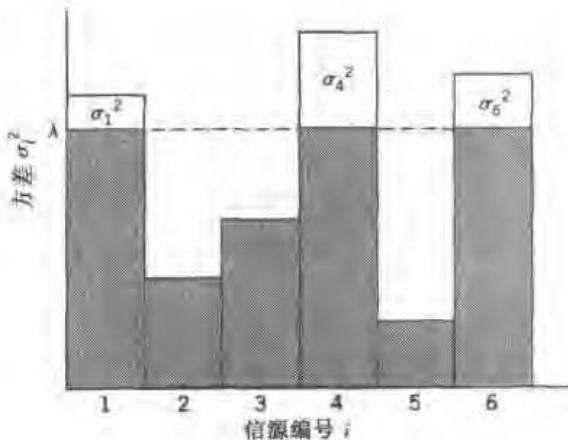


图 9.23 一组并行高斯过程的反向注水图

9.14 数据压缩

与率失真理论密切关联的一个概念就是数据压缩。数据压缩有目的且不可避免地降低了连续或离散信源数据中的信息内容。特别地,可考虑采用数据压缩器或信号压缩器,这是一种提供最短信源输出符号的设备,能够承受可接受的失真。通过以慎重的、有控制的方式使某些细节变得晦涩,数据压缩器保持了信源输出的关键信息内容。因此,数据压缩是有损伤的操作,它降低了信源熵(即信息被丢失),这与信源的类型无关。

对于离散信源,采用数据压缩是为了以小于信源熵的速率对信源输出进行编码。这样就违背了信源编码定理,意味着原始数据的真实再现不再可能。

对于连续信源,熵是无限的,因此必须采用信号压缩来以有限的速率对信源输出进行编码。因此,无法以有限的比特数对一个模拟信号进行数字编码而不产生任何失真。这个结论与第 3 章讨论的 PCM 的思想是一致的。在第 3 章中已经证明,量化作为 PCM 中模 - 数转换过程的基础,总会向发射信号中引入失真(称为量化噪声)。因此,可以认为量化器是信号压缩器。

第 3 章中讨论的均匀和非均匀量化器被认为是标量量化器,因为该量化器每次只处理模

拟信号(即连续信源输出)的一个抽样点。每个抽样点转换为一个量化值,抽样点之间的转换是相互独立的。标量量化器是一种相对简单的信号压缩器,非常适于实际应用。标量量化器能够提供相当好的性能,尤其是在使用非均匀量化器的情况下。

还有另一类量化器,称为矢量量化器。矢量量化器将信源输出的连续抽样点进行分组以构成矢量,每个矢量都被视为一个单独的整体。矢量量化器的关键操作是通过将随机矢量^[16]编码为二进制码字来进行量化,通过与码本进行比较来对矢量进行编码。码本由一组被称为码矢量或图样的预先存储的参考矢量构成。码本中的每个图样都用于表示输入矢量,并遵从保真度最大化准则。矢量量化器中的编码过程可以认为是图样匹配操作。

以 N 表示码本中码矢量的个数, k 表示每个矢量的维数(即每个图样中抽样点的个数), r 表示每个抽样点的编码传输速率,其单位为比特。这三个参量的关系如下:

$$r = \frac{\log_2 N}{k} \quad (9.137)$$

假设码本足够大,矢量量化器的信噪比由下式给出:

$$10 \log_{10}(\text{SNR}) = 6\left(\frac{\log_2 N}{k}\right) + C_k \text{ dB} \quad (9.138)$$

其中, C_k 为一常数(以 dB 表示),取决于维数 k 。根据式(9.138),当码本的大小成倍增加的时候,矢量量化器的信噪比以大约 $6/k$ dB 的速率增加。等价地,可认为在采用标准标量量化器的标准 PCM 系统中,速率(每抽样点的比特数)每增加一个单位,信噪比就增加 6 dB。与标量量化器相比,矢量量化器的优势在于,常量 C_k 具有较高的值,这是因为矢量量化器有效利用了构成矢量的抽样点之间的相关性。对于一个给定的信息信源,随着维数 k 增加,常量 C_k 最后趋于率失真极限。但是,信噪比的提高是以编码复杂度的增加为代价的,对于一个指定的速率 r ,信噪比随维数 k 以指数级增长。这是矢量量化器实际应用中的主要障碍。在某些应用中,通过采用 VLSI 技术将高度复杂的信号处理器集成在一个硅芯片上,就能够解决计算复杂度的问题。例如,在 CDMA 无线通信系统即 IS-95 的 CELP 语音模型中,就采用了这样的技术。从 8.9 节对 CELP 的描述可见,语音 CELP 模型就是矢量量化的一个例子。

9.15 总结与讨论

本章给出了通信系统不同方面的四个基本限制。这些限制分别体现在信源编码定理、信道编码定理、信息容量定理和率失真函数中。

信源编码定理即香农第一定理,提供了研究数据压缩(即由离散无记忆信源产生的数据无损压缩)的数学工具。这个理论表明:可以使每个信源符号的平均二进制码元数(比特数)与信源的熵相等,但是不能小于它。信源的熵是构成信源字符集的符号概率的函数。由于熵是不确定性的度量,当相应概率分布最大不确定性时,熵具有最大值。

信道编码定理即香农第二定理,是信息论中最重要的理论。对于二进制对称信道,信道编码定理表明:对于一个小于或等于信道容量 C 的码率 r ,存在适当的码使得平均差错概率可以达到要求。二进制对称信道是离散无记忆信道的一种最简单形式。发送 0 时接收为 1 和发送 1 时接收为 0 的概率是相等的,所以称之为对称的。差错概率是转移概率的函数。转移概率 p 不仅由

信道输出的加性噪声决定,还由采用的接收机类型决定。 p 的值惟一定义了信道容量 C 。

著名的香农第三定理,即信息容量定理表明,当系统的功率受限时,存在一个任何通信系统都能够可靠(即没有错误地)工作的最大速率。该最大速率称为信息容量,以 bps 度量。当系统工作在高于信息容量的速率下,无论用于发送的信号集或用于处理接收信号的接收机如何选择,都将具有很高的差错概率。

最后,率失真函数提供了信号压缩的数学工具(即解决了保真度准则下的信源编码问题)。率失真函数可以用于离散和连续的无记忆信源。

当信源的输出以无损方式进行压缩时,得到的数据流通常含有冗余比特。通过采用例如霍夫曼编码或 Lempel-Ziv 算法等无损算法,就可以去除这些冗余。可将数据压缩和紧随其后的无失真数据压缩(compaction)视为信源编码分割的两个要素,因为它们涉及到了信源的全部。在一些信源编码的应用中,还存在第三个要素,称为数据加密,一般在无失真数据压缩之后进行。数据加密的目的是使数据(比特)流变得晦涩,以便使得这些数据对于未经认可的接收者来说不具有任何意义。密码系统的一些基础知识,包括加密和解密过程,可从信息论中得到,参见附录 5。密码系统的其他相关问题在附录中也有述及。

最后强调一点。本章介绍的香农信息论,是完全针对无记忆信源和信道的。这个理论可以推广到有记忆信源和信道的情况,此时,每个信号都取决于前面的信号。但这已经超出了本书的范围。^[17]

注释与参考

- [1] 根据 Lucky(1989),香农第一次提到信息论是在 1945 年的题为“*A Mathematical Theory of Cryptography*”的备忘录中。这个概念是相当新颖的,且在香农(奠定了信息论基础的)1948 的论文中未提及。关于信息论的介绍性文章见 Lucky(1989) 和 Wyner(1981) 的第 2 章和 Adámek(1991), Homming(1980) 和 Abramson(1963)。关于该主题的进一步介绍见 Cover 和 Thomas(1991), Blahut(1987), McEliece(1977) 以及 Gallager(1968)。有关信息论发展(包括著名的香农 1948 年论文)的论文集见 Slepian(1974)。香农发表的论文集见 Sloane 和 Wyner(1993)。
- [2] 信息的对数度量首先由 Hartley(1928) 提出,但 Hartley 采用的对数是以 10 为底的。
- [3] 在统计物理中,物理系统的熵定义为(Reif, 1967, pp.147):

$$S = k \log \Omega$$

其中, k 为巴特沃兹常数, Ω 为系统可能的状态个数, \log 表示自然对数运算。这是一个有能量的量纲,因为定义中包含了常数 k 。特别地,该式提供了系统随机程度的量化度量。将统计物理中的熵与信息论中的熵进行比较可得,二者具有相同的形式。关于二者关系的详细讨论见 Pierce(1961, pp.184 – 207) 和 Brillouin(1962)。

- [4] 对信源编码定理的原始证明见 Shannon(1948)。信源编码定理的一般证明可参阅以下著作: Viterbi 和 Omura(1979, pp.13 – 19), McEliece(1977, Chapter 3) 以及 Gallager(1968, pp.38 – 55)。信源编码定理在文献中称为无噪编码定理,无噪是指它使无误编码的条件成为可能。

- [5] 对 Kraft-McMillan 不等式的证明见 Cover 和 Thomas(1991, pp.82 – 84), Blahut(1990, pp.298 – 299)以及 McEliece(1977, pp.239 – 240)。对式(9.23)的证明见 Cover 和 Thomas(1991, pp.87 – 88), Blahut(1990, pp.300 – 301)以及 McEliece(1977, pp.241 – 242)。
- [6] 霍夫曼编码以其发明者 D.A.Huffman(1952)的名字命名。关于霍夫曼编码的通俗论述及其在数据压缩方面的应用见 Adámek(1991)。
- [7] Lempel-Ziv 算法的原始论文是 Ziv 和 Lempel(1977, 1978)。Lempel-Ziv 算法的通俗描述见 Lucky(1989, pp.118 – 122), Blahut(1990, pp.314 – 319), Gitlin 以及 Hayes 和 Weinstein(1992, pp.120 – 122)。Lempel-Ziv 算法在英语文章压缩上的应用见 Lucky(1989, pp.122 – 128)和 Welch(1984), 也可参见 Weiss 和 Shrem(1993)的评论文章。
- [8] 信道编码定理也称为有噪声编码定理。此定理的原始证明由 Shannon(1948)给出。这个定理的证明在 Hamming(1980, 第 9、10 章)也有详细的说明, 其中包括相关结论的综合评论。定理的第二部分在文献中称为编码定理的逆定理。这个定理的证明参见以下的参考文献: Viterbi 和 Omura(1979, pp.28 – 34)以及 Gallager(1968, pp.76 – 82)。
- [9] 式(9.70)中

$$\int_{-\infty}^{\infty} f_Y(x) \log_2 \left(\frac{f_X(x)}{f_Y(x)} \right) dx$$

称为概率密度函数 $f_X(x)$ 和 $f_Y(x)$ 之间的相关熵或 Kullback-Leibler 散度, 见 Kullback(1968)。

- [10] 香农信息容量定理在文献中又称为 Shannon-Hartley 法则, 这是为了纪念 Hartley 对信息传输的早期工作(Hartley, 1928)。Hartley 证明了在给定信道上传输的信息数量与信道带宽和操作时间的乘积成正比例关系。
- [11] 关于球体打包的通俗描述参见 Cover 和 Thomas(1991, pp.242 – 243), 也可参见 Wozencraft 和 Jacobs(1965, pp.323 – 341)。
- [12] 图 9.18 的(a)和(b)部分源自 Frey(1998)的著作中图 6.2 的相应部分。
- [13] 关于有色噪声信道的信息容量的严密论述见 Gallager(1968)。对将图 9.19(a)中的信道模型替换为图 9.19(b)中模型的相关讨论见 Gitlin, Hayes 和 Weinstein(1992)。
- [14] 率失真理论的详细讨论见 Berger(1971)。该主题在 Cover 和 Thomas(1991), McEliece(1977)以及 Gallager(1968)中也有所提到。
- [15] 式(9.131)的推导见 Cover 和 Thomas(1991, pp.345)。式(9.131)定义的率失真函数 $R(D)$ 的算法在 Blahut(1987, pp.220 – 221)以及 Cover 和 Thomas(1991, pp.364 – 367)中有介绍。
- [16] 矢量量化器的早期论文见 Gersho(1979)和 Linde, Buzo 以及 Gray(1980)。关于矢量量化的指导性介绍见 Gray(1984)。式(9.138)定义的矢量量化器的信噪比, 在 Gersho 和 Cuperman(1983)中有讨论。矢量量化的详细讨论见 Gersho 和 Gray(1992)。
- [17] 有记忆的离散信道的详细讨论见 Gallager(1968, pp.97 – 112)和 Ash(1965, pp.211 – 229)。

习题

熵

9.1 用 p 表示事件的概率,对于 $0 \leq p \leq 1$,试给出信息量对事件概率的变化曲线。

9.2 每个信号间隔内,信源发出 4 个可能符号中的一个,其概率分别为:

$$p_0 = 0.4$$

$$p_1 = 0.3$$

$$p_2 = 0.2$$

$$p_3 = 0.1$$

求通过观察信源发出符号的情况而获得的信息量。

9.3 信源分别以 $1/3, 1/6, 1/4, 1/4$ 的概率发出 4 个符号 s_0, s_1, s_2, s_3 中的一个。这些连续符号是统计独立的。计算该信源的熵。

9.4 用 X 表示单次掷一个公平的骰子的结果,问 X 的熵为多少?

9.5 具有零均值和单位方差的高斯过程的样本函数被均匀抽样,并送至具有图 P9.5 所示输入 - 输出幅度特性的均匀量化器。计算量化器输出的熵。

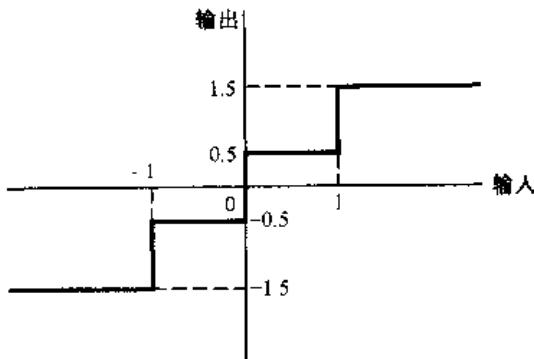


图 P9.5

9.6 考虑一个信源字符集为 $\mathcal{S} = \{s_0, s_1, \dots, s_{K-1}\}$ 且信源统计量为 $\{p_0, p_1, \dots, p_{K-1}\}$ 的离散无记忆信源。该信源的第 n 次扩展为另一个离散无记忆信源,其信源字符集为 $\mathcal{S}^n = \{\sigma_0, \sigma_1, \dots, \sigma_{M-1}\}$,其中, $M = K^n$ 。以 $P(\sigma_i)$ 表示 σ_i 的概率。

(a) 证明

$$\sum_{i=0}^{M-1} P(\sigma_i) = 1$$

这是可以预见的。

(b) 证明

$$\sum_{i=0}^{M-1} P(\sigma_i) \log_2 \left(\frac{1}{P(\sigma_i)} \right) = H(\mathcal{S}), \quad k = 1, 2, \dots, n$$

其中, p_{i_k} 为符号 s_{i_k} 的概率, $H(\mathcal{S})$ 为原始信源的熵。

(c) 证明

$$\begin{aligned} H(\mathcal{S}^n) &= \sum_{i=0}^{M-1} P(\sigma_i) \log_2 \frac{1}{P(\sigma_i)} \\ &= nH(\mathcal{S}) \end{aligned}$$

9.7 考虑信源字符集 $\mathcal{S} = \{s_0, s_1, s_2\}$ 且信源统计量为 $\{0.7, 0.15, 0.15\}$ 的离散无记忆信源。

(a) 计算信源的平均信息量。

(b) 计算信源的二阶扩展的平均信息量。

9.8 存储文字所需要的比特数比存储相同内容的语音要少得多, 请解释原因。

无失真数据压缩

9.9 考虑一个离散无记忆信源, 其字符集由 K 个等概率字母构成。

(a) 试解释为什么使用固定长度的码表示这种信源比其他码更有效。

(b) 码率为 100% 时, K 和码字长度必须满足什么样的条件?

9.10 考虑以下四组码:

符号	码 I	码 II	码 III	码 IV
s_0	0	0	0	00
s_1	10	01	01	01
s_2	110	001	011	10
s_3	1110	0010	110	110
s_4	1111	0011	111	111

(a) 这四组码中的两个为前缀码。指出这两组码并构造它们各自的判决树。

(b) 对码 I, 码 II, 码 III 和码 IV 应用 Kraft-McMillan 不等式。根据(a)中得到的结果讨论你的结论。

9.11 考虑由英文字母构成的序列, 每个字母的发生概率如下:

字母 $a \quad i \quad l \quad m \quad n \quad o \quad p \quad y$

概率 $0.1 \quad 0.1 \quad 0.2 \quad 0.1 \quad 0.1 \quad 0.1 \quad 0.2 \quad 0.1 \quad 0.1$

计算这个字符集的霍夫曼编码。一种情况下, 在编码过程中将组合符号置于尽可能高的位置上, 另一种情况下则将其置于尽可能低的位置上。分别找出这两种码在字母组合上的平均码长及其方差。

9.12 考虑字符集由 7 个符号构成的离散无记忆信源, 每个符号的发生概率如下:

符号 $s_0 \quad s_1 \quad s_2 \quad s_3 \quad s_4 \quad s_5 \quad s_6$

概率 $0.25 \quad 0.25 \quad 0.25 \quad 0.125 \quad 0.125 \quad 0.0625 \quad 0.0625$

计算以尽可能高的位置放置“组合”符号时, 此信源的霍夫曼编码。解释为什么计算得到的信源码具有 100% 的效率。

9.13 考虑具有字符集 $\{s_0, s_1, s_2\}$ 且输出统计量为 $\{0.7, 0.15, 0.15\}$ 的离散无记忆信源。

- (a) 对这个信源应用霍夫曼算法, 证明霍夫曼码的平均码长等于 1.3 比特/符号。
 (b) 将这个信源扩展为二阶。对得到的扩展信源应用霍夫曼算法, 并证明新的码字的平均码长等于 1.1975 比特/符号。
 (c) 将(b)中计算得到的平均码长与原始信源的平均信息量进行比较。

9.14 图 P9.14 为霍夫曼树。由此霍夫曼树表示的符号 A, B, C, D, E, F, G 的码字分别是什么? 每个码长分别为多少?

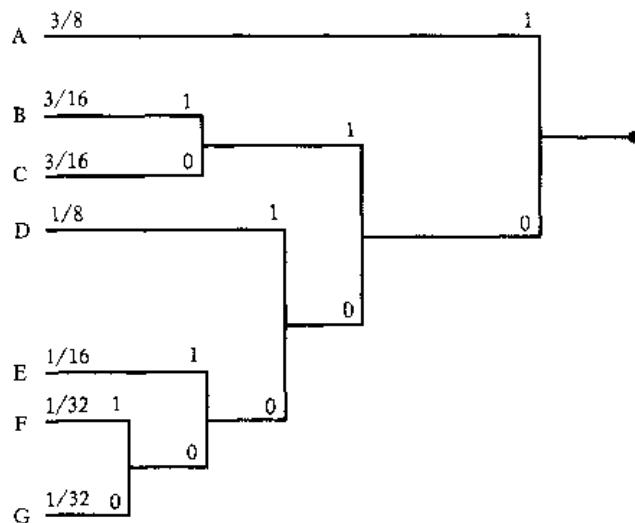


图 P9.14

9.15 计算机执行 4 条指令, 分别由码字(00, 01, 10, 11)表示。假设独立使用每个指令的概率分别为 $\{1/2, 1/8, 1/8, 1/4\}$, 计算采用最佳信源编码时, 所用指令对应的比特数可以减少多少百分比, 并构造霍夫曼编码来实现。

9.16 考虑以下的二进制序列

11101001100010110100...

使用 Lempel-Ziv 算法对这个序列进行编码。假设二进制符号 0 和 1 已经在码本中。

二进制对称信道

9.17 考虑图 9.8 所示的二进制对称信道的转移概率图。二进制符号 0 和 1 以同等的概率发生。找出二进制符号 0 和 1 在信道输出中出现的概率。

9.18 重复习题 9.17 的计算, 假设输入二进制符号 0 和 1 发生的概率分别为 $1/4$ 和 $3/4$ 。

互信息和信道容量

9.19 考虑以转移概率 p 表征的二进制对称信道。画出信道互信息对 p_1 的曲线, p_1 是信道输入符号 1 的先验概率。分别对转移概率 $p = 0, 0.1, 0.2, 0.3, 0.5$ 进行计算。

9.20 图 9.10 显示了具有转移概率 p 的二进制对称信号的信道容量的变化情况。应用习题 9.19 得到的结果解释这种变化。

9.21 考虑图 9.8 所示的二进制对称信道。以 p_0 表示发送二进制符号 $x_0 = 0$ 的概率, $p_1 = 1 - p_0$ 表示发送二进制符号 $x_1 = 1$ 的概率。 p 表示信道的转移概率。

(a) 证明信道输出和输入之间的互信息为

$$I(\mathcal{X}; \mathcal{Y}) = H(z) - H(p)$$

其中

$$\begin{aligned} H(z) &= z \log_2\left(\frac{1}{z}\right) + (1-z) \log_2\left(\frac{1}{1-z}\right) \\ z &= p_0 p + (1-p_0)(1-p) \end{aligned}$$

和

$$H(p) = p \log_2\left(\frac{1}{p}\right) + (1-p) \log_2\left(\frac{1}{1-p}\right)$$

(b) 证明使 $I(\mathcal{X}; \mathcal{Y})$ 最大的 p_0 等于 $1/2$ 。

(c) 证明信道容量等于

$$C = 1 - H(p)$$

9.22 两个二进制对称信道以图 P9.22 所示的方式进行级联。假设两个信道具有与图 9.8 相同的转移概率图，试给出该级联的整体信道容量。

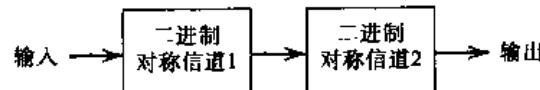


图 P9.22

9.23 二进制删除信道具有两个输入和三个输出，如图 P9.23 所示。输入标以 0 和 1，输出标以 $0, 1, e$ 。输入比特的 α 部分被信道删除。求信道的容量。

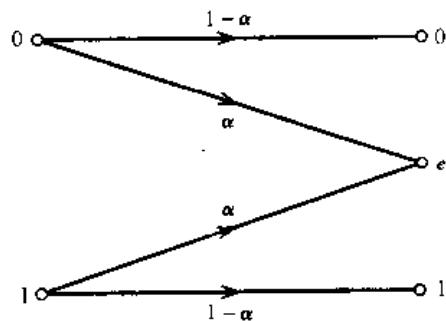


图 P9.23

9.24 考虑一个数字通信系统，其信道编码/译码中采用重复码。每个传输重复 n 次，其中， $n = 2m + 1$ ，且 n 为一奇数。译码器的工作方式为：如果在 n 比特的接收块中，0 的个数超过了 1 的个数，译码器译码为 0；反之译码为 1。当 $n = 2m + 1$ 中有 $m + 1$ 或更多个输出是不正确的，就会出错。假设信道为一个二进制对称信道。

(a) 对于 $n = 3$ ，证明平均差错概率为

$$P_e = 3p^2(1-p) + p^3$$

其中， p 为信道的转移概率。

(b) 对于 $n = 5$ ，证明平均差错概率为

$$P_e = 10p^3(1-p)^2 + 5p^4(1-p) + p^5$$

(c) 对于一般的情况, 证明平均差错概率为

$$P_e = \sum_{i=m+1}^n \binom{n}{i} p^i (1-p)^{n-i}$$

相对熵

9.25 以 X_1, X_2, \dots, X_n 表示高斯矢量 \mathbf{X} 的元素。 X_i 是独立的, 均值为 μ_i , 方差为 σ_i^2 , $i = 1, 2, \dots, n$ 。证明矢量 \mathbf{X} 的相对熵等于

$$h(\mathbf{X}) = \frac{n}{2} \log_2 [2\pi e (\sigma_1^2 \sigma_2^2 \cdots \sigma_n^2)^{1/n}]$$

如果方差相等, $h(\mathbf{X})$ 减至多少?

9.26 连续随机变量 X 的幅度峰值为 M , 即 $-M < X < M$ 。

(a) 证明当 X 均匀分布的时候, X 的相对熵最大, 即

$$f_X(x) = \begin{cases} 1/2M, & -M < x \leq M \\ 0, & \text{其他} \end{cases}$$

(b) 证明 X 的相对熵的最大值等于 $\log_2 2M$ 。

9.27 证明式(9.79)~式(9.81)所示的互信息 $I(X; Y)$ 的性质。

9.28 考虑定义如下的连续随机变量 Y :

$$Y = X + N$$

其中, X 和 N 是统计独立的。证明给定 X 时, Y 的条件相对熵等于

$$h(Y|X) = h(N)$$

其中, $h(N)$ 是 N 的相对熵。

信息容量

9.29 电话网络的一个语音信道的带宽为 3.4 kHz。

(a) 计算电话信道的信息容量, 其中给定信噪比为 30 dB。

(b) 为了能够支持以 9600 bps 的速率在电话信道中进行信息传输, 计算所需要的最小信噪比。

9.30 包含文字和数字的数据从远端通过语音电话信道输入计算机。信道带宽为 3.4 kHz, 输出信噪比为 20 dB。终端总共有 128 个符号。假设这些符号都是等概的, 且连续传输是统计独立的。

(a) 计算信道的信息容量。

(b) 计算信道上无误传输所允许的最大符号速率。

9.31 黑白电视机的画面可认为是由大约 3×10^5 个元素构成, 每个元素等概地占据 10 个不同亮度中的一个。假设:

(1) 传输速率等于 30 图像帧/秒,

(2) 信噪比等于 30 dB。

使用信息容量定理,计算支持最终视频信号传输所需要的最小带宽。

注意:商业电视传输中一般采用的带宽为 4.2 MHz,以便与分配的 6 MHz 带宽相适应。

- 9.32 本题中将继续研究例 9.9。假设图 9.15(b)所示的紧密压缩的星座图按比例增加,以便将每个符号的发射信号能量维持在与图 9.15(a)中 64-QAM 星座图相同的平均值上。试给出该缩放比例下的新星座图。这个新星座图的误比特率与图 9.15(a)中相比如何?证明你的结论。

- 9.33 双绞线信道的平方幅度响应的数学模型为

$$|H(f)|^2 = \exp(-\alpha\sqrt{f})$$

常数 α 的定义为

$$\alpha = \frac{kl}{l_0}$$

其中, k 为取决于电线规格的常量, l_0 为参考长度, l 为我们研究的双绞线的实际长度。造成近端串话的平方幅度响应具有如下形式:

$$|H_{\text{NEXT}}(f)|^2 = \beta f^{3/2} \quad (9.1)$$

其中, β 为取决于电缆型号的常量。

试给出此处描述的近端串话占优势的信道的信息容量表达式。

数据压缩

- 9.34 矢量量化器的信噪比公式(9.138)包含了式(3.33)所示的标准 PCM 在 $k = 1$ 时的特殊情况。证明这个结论的正确性。
- 9.35 所有实际的数据压缩和数据传输方式都受到率失真函数和信道容量定理这两个限制。这两个理论以不同的方式包括了互信息的概念,详细阐述这两个论断引出的论点。

计算机实验

- 9.36 回顾例 9.12,其中研究了 AWGN 信道中的二进制正反信号。由式(9.112)和下述理论开始,构造一个计算给定误比特率下的最小 E_b/N_0 的软件包,其中, E_b 为每比特信号能量, $N_0/2$ 为噪声功率谱密度。试计算图 9.18 中(a)和(b)的结果。

如例 9.12 中所述,信道输入和输出的互信息可由蒙特卡罗积分来进行很好地近似。为了解释这个方法,考虑很难随机抽样的函数 $g(y)$ 。这里,函数 $g(y)$ 表示信道输出的相对熵公式中复杂的被积函数。计算过程如下:

- 找到一个区域 A ,其中包含了我们感兴趣且容易抽样的区域。
- 选择 N 个点,这些点均匀随机分布在区域 A 中。

蒙特卡罗积分定理说明:函数 $g(y)$ 对 y 的积分近似地等于区域 A 与 $g(y)$ 曲线下的点占区域 A 的比例的乘积,如图 P9.36 所示。这种近似方法的准确度随着 N 的增大而提高。

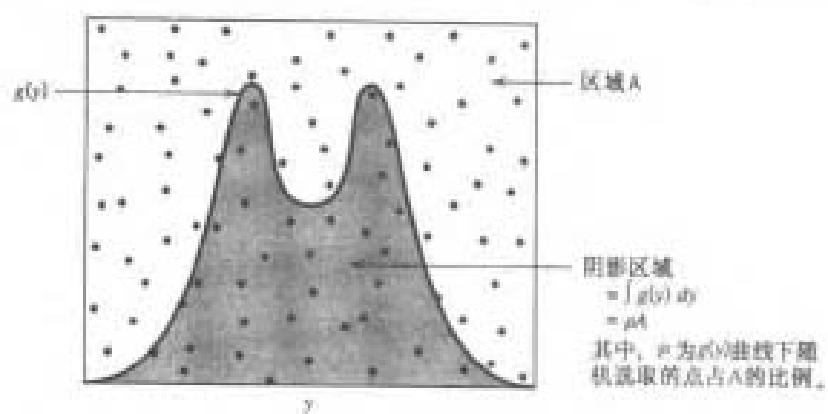


图 19.36

第 10 章 差错控制编码

本章是前一章中香农信息论的自然延续。本章将特别介绍几种差错控制编码技术，这些技术从不同的途径实现了香农信道编码定理。每种差错控制编码技术均包括发射机中的信道编码和接收机中的译码算法。

这里讨论的差错控制编码主要分为以下几种：

- 线性分组码，
- 循环码。
- 卷积码。
- 复合码，例如 Turbo 码、低密度奇偶校验码及其不规则变型。

10.1 引言

数字通信系统的设计者所面临的任务就是提供一种经济而有效的设备，使得系统的一端能以一定的速率发送信息，同时另一端能以一定的质量和可靠性接收信息。设计者可利用的两个关键的系统参数就是信号的发射功率和信道带宽。这两个参数和接收噪声的功率谱密度一起决定了每比特信号能量与噪声功率谱密度的比值 E_b/N_0 。在第 6 章中，我们已经指出：对于某一特定的调制方式，这个比值是误比特率的惟一决定因素。在实际应用中，在确定 E_b/N_0 的取值时，通常需要设置一个门限。因此实际上常常会出现这样的情况，即调制方式无法提供可接受的数据质量（即足够低的误码性能）。对于给定的 E_b/N_0 值，改善数据质量，使之从不可接受变为可接受，惟一可行的办法就是采用差错控制编码。

编码的另一个目的是为了在一定的误比特率下，降低系统所要求的 E_b/N_0 。 E_b/N_0 的降低有利于降低发射功率。在无线通信中，发射功率的降低意味着可以采用尺寸更小的天线，从而降低系统的硬件成本。

为了不破坏数据的完整性，差错控制^[1]借助于前向纠错(FEC)的方法来实现。图 10.1(a)给出了采用这种方法的数字通信系统模型。离散信源产生二进制符号形式的信息。发射机中的信道编码器接收比特信息，然后按指定的规则加上冗余，从而产生更高比特率的编码数据。接收机中的译码器充分利用这些冗余信息来判断发端实际发送的比特信息。信道编码器和信道译码器相结合的目的就是为了尽量降低信道噪声的影响，也就是说，尽量保证信道编码器的输入（来自信源）和信道译码器的输出（传送给用户）之间的差错数最小。

对于某一固定的调制方式来说，编码后的信息中冗余的增加意味着需要提高传输带宽。此外，采用差错控制的同时增加了系统的复杂度，尤其是在接收机的译码过程中。因此，在采用差错控制编码以获取可接受的误码性能时，系统设计要考虑带宽和系统复杂度两方面的折中。

纠错码有很多种类，分别源自不同的数学方法。历史上将其分为分组码和卷积码。区分这两类码的主要依据就是是否含有存储器。

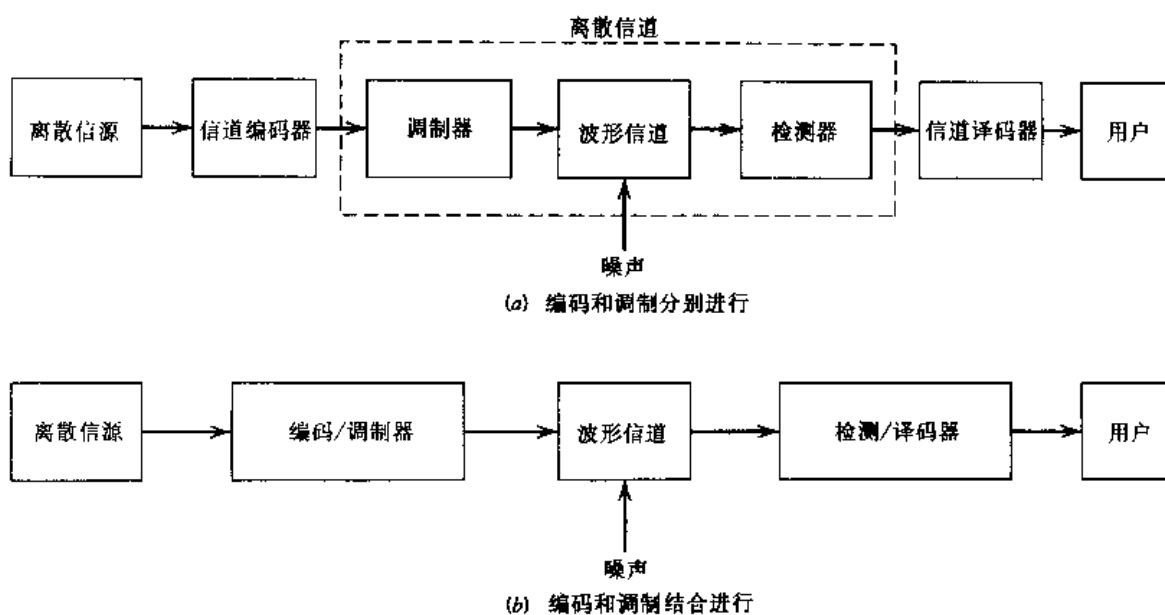


图 10.1 数字通信系统的简化模型

为了产生 (n, k) 的分组码, 信道编码器接收到以连续的 k 比特为一组的信息, 并为每个分组加上与 k 个信息比特数相关的 $n - k$ 个冗余比特, 从而产生全长为 n 比特的码组, 其中 $n > k$ 。此 n 比特码组称为一个码字, n 称为码的分组长度。信道编码器产生速率为 $R_0 = (n/k)R_s$ 的比特流, R_s 为信源的比特速率。无量纲比值 $r = k/n$ 称为码率, $0 < r < 1$ 。编码器输出的比特速率 R_0 称为信道数据速率。由于信源产生的数据速率和信道的数据速率都是以 bps 作为度量的, 因此编码速率 r 是一个无量纲的比值。

卷积码的编码过程可以看成是输入序列与编码器冲激响应的时域离散卷积。冲激响应的持续时间与编码器的存储器容量相当。因此, 卷积码的编码器采用一种持续时间与其存储器容量相等的“滑动窗口”来处理输入的信息序列。也就是说, 与分组码不同, 在卷积码中, 信道编码器依次连续地接收信息比特, 从而产生一个速率更高的连续码序列。

在图 10.1(a)所示的模型中, 发射机中的编码和调制是分开进行的, 接收机中的检测和译码也是分开进行的。然而, 当带宽利用率成为主要考虑因素时, 实现前向纠错编码的最有效办法就是将编码与调制结合起来成为一个单一的功能模块, 如图 10.1(b)所示。这种方法将编码过程改进为在发射信号上叠加某种信号图样的过程。

10.1.1 自动重发请求

FEC 依赖于在发送码字中有控制地加入冗余信息。信号在噪声信道中进行传输时会产生误码现象, 从而需要进行误码检测和纠正。不管接收码字的译码是否成功, 接收机都不进行任何进一步处理。因此, 适用于 FEC 的信道编码技术只要求发射机和接收机之间是单向的链接。

此外, 还有一种解决差错控制的技术, 称为自动重发请求(ARQ)^[2]。ARQ 的基本思想完全不同于 FEC。ARQ 采用冗余的目的仅仅是为了进行误码检测。当在传输的码字中检测到错误时, 接收机就要求重传错误的码字, 这就需要使用一条返回路径(反馈信道)。因此, ARQ 只能用于半双工或全双工的链路。在半双工链路上, 链路可以在任一方向上传送数据, 但不能同时

进行两个方向上的数据传输。另一方面,在全双工链路上,数据可以在两个方向上同时传送。

半双工链路采用的是最简单的 ARQ 方案,称为停止和等待策略。这种方法先将信息比特组编码为码字并在信道上传输,然后发射机停止发送,等待接收机的反馈。反馈信号可能是码字正确接收的确认,也可能是译码过程中发现误码而要求重发码字的请求。如果是后一种情况,发射机在发送下一信息比特组之前,重新发送问题码字。

在停止 - 等待 ARQ 中,停顿问题会导致数据吞吐量的降低。另有一种带后退的连续 ARQ,就能够减少这种停顿。这种连续 ARQ 采用全双工链路,因此,在发射机沿前向信道发送码字的同时,接收机就可以发送反馈信号。特别地,发射机可以持续地发送连续的码字,直到收到由接收机(反馈信道上的)发出的要求重发的请求才停止。此时,发射机停止发送,后退到那个不能被接收机正确译码的码字,然后从那个中断的码字开始重发整个码字序列。

连续 ARQ 的一种改进版本称为选择重发连续 ARQ,这种策略只重发那个错误接收的码字,从而使得数据的吞吐量得到进一步提高。换言之,就是在重发时,不再发送那些错误码字后面的已被成功接收的码字。

上述三种类型的 ARQ 都在全双工或半双工链路的需要与通信资源的有效利用之间找到了平衡点。这三种 ARQ 都依赖于下列两个前提:

- 错误检测,这使得译码器的设计相对简单。
- 无噪声的反馈信道,这不是一个严格的约束条件,因为典型的反馈信道的信息速率是相当低的。

正因如此,ARQ 在计算机通信系统中得到了广泛的应用。

但是,由于 FEC 的应用仅要求单向的链路,这使得 FEC 的应用比 ARQ 更为广泛。同时,随着微处理器和 VLSI 电路的应用,译码器的性价比有了很大改善,从前由于需要将检错和纠错结合在一起而造成的 FEC 译码器的复杂性,已经不再是一个主要问题了。

10.2 离散无记忆信道

回到图 10.1(a)所示的模型,如果给定时间间隔上的译码器输出只依赖于该时间间隔上的发射信号,而与之前的发射信号无关,就称这个波形信道为无记忆的。在这种情况下,可将调制器、波形信道和检测器合成在一起的模型称为离散无记忆信道。这种信道可以用一组转移概率 $p(j|i)$ 来完全描述,其中 i 表示调制器的输入符号, j 表示解调器的输出符号, $p(j|i)$ 表示在发送符号 i 的情况下,接收到 j 的概率。离散无记忆信道已在前面的 9.5 节有所描述。

最简单的离散无记忆信道产生于二进制输入和输出符号。采用二进制编码时,调制器的输入只可能是二进制符号 0 和 1。同样,如果解调器输出使用二进制量化,译码器也只有二进制的输入。也就是说,解调器的输出采用硬判决来判断哪个是实际发送的符号。转移概率如图 10.2 所示的一个二进制对称信道(BSC)就属于这种情况。假设信道的噪声模型是加性高斯白噪声,二进制对称信道可以用转移概率 p 完全描述。由于采用此方案实现简单,大多数编码数字通信系统都采用硬判决译码的二进制编码。硬判决译码器,或者叫代数译码器,采用了特殊的代数结构。这种结构包含在信道编码之中,可使译码实现起来相对简单。

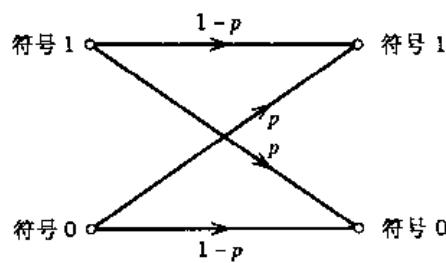
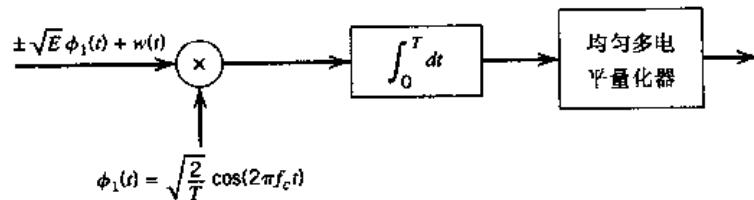
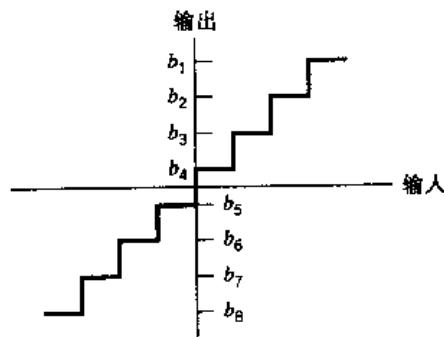


图 10.2 二进制对称信道的转移概率图

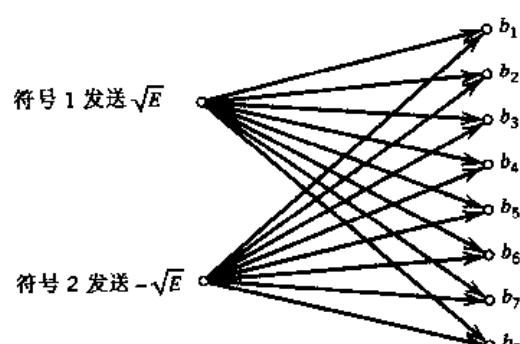
在接收机译码之前采用硬判决, 将对信息造成不可恢复的损失。为了减少这种损失, 可采用软判决。软判决可通过在解调器的输出端加上一个多级的量化器来实现。图 10.3(a)所示为二进制 PSK 信号的情况。量化器的输入输出特性如图 10.3(b)所示。调制器的输入只有二进制的 0 和 1 两种符号, 而解调器的输出则是有 Q 个符号的符号集。假设采用图 10.3(b)所示的量化器, 此时 $Q = 8$ 。这样的信道称为二进制输入 Q 进制输出的离散无记忆信道。相应的转移概率图如图 10.3(c)所示。这种分配形式和译码器的性能首先取决于量化器表示电平的位置, 其次取决于信号电平和噪声的变化。因此, 如果多电平量化器能实现, 则解调器中必须加入自动增益控制。此外, 软判决的采用使译码器的实现变得更加复杂。但由于概率的方法要优于代数的方法, 所以软判决的性能相对于硬判决有重大的改进。相应地, 可将软判决译码器称为概率译码器。



(a) 二进制 PSK 接收机



(b) 多电平量化器的传输特性



(c) 信道转移概率框图

图 10.3 二进制输入 Q 进制输出的离散无记忆信道。(b) 和 (c) 所示为八电平量化

10.2.1 信道编码定理回顾

第 9 章中已经建立了信道容量的概念, 对于离散无记忆信道而言, 信道容量表示每次通过信道传输的最大信息量。信道编码定理是指, 如果一个离散无记忆信道的信道容量为 C , 信源产生的信息速率小于 C , 则存在一种编码技术, 使得信源的输出能以任意低的符号差错概率在

信道中传输。这个定理表明,对于二进制对称信道来说,如果码速率 r 小于信道容量 C ,则能够找到一种码,在信道中实现无误码传输。相反,如果码速率 r 大于信道容量 C ,则不可能找到这样的码。

信道编码定理指出,在离散无记忆信道上实现信息可靠传输(无误码)时,容量 C 对速率的基本限制。只要信噪比足够大,首要问题已不再是信噪比,而是信道输入的编码方法。

但是,信道编码定理最不尽如人意之处就在于它未提供编码方案。此定理说明了有效编码的存在,但并未给出找到它们的方法。这里的有效编码是指这样一系列信道编码,当比特率达到小于信道容量的某一最大值时,它能在噪声信道上提供可靠的信息传输(在任意小的符号差错概率下)。本章所描述的纠错编码技术就提供了设计有效编码的不同方法。

10.2.2 模2运算

本章描述的码字都是二进制码,因此,符号仅由 0 和 1 组成。在这样的码字中,编码和译码函数包含对码字的二进制模2加法和乘法运算。

在本章中,我们用加号(+)表示模2加法。由于本章的全部计算都是基于二进制的,所以这样的术语使用不会引起混淆。因此,应避免使用前面章节用到的特殊符号 \oplus 。根据本章采用的表示法,模2加法的规则如下:

$$\begin{aligned} 0 + 0 &= 0 \\ 1 + 0 &= 1 \\ 0 + 1 &= 1 \\ 1 + 1 &= 0 \end{aligned}$$

因为 $1 + 1 = 0$, 所以有 $1 = -1$ 。因此,在二进制算法中,减法和加法是相同的。模2的乘法规则如下:

$$\begin{aligned} 0 \times 0 &= 0 \\ 1 \times 0 &= 0 \\ 0 \times 1 &= 0 \\ 1 \times 1 &= 1 \end{aligned}$$

除法规则有:

$$1 \div 1 = 1, 0 \div 1 = 0$$

0 不能作为除数。在逻辑运算中,模2加是异或运算,模2乘法则是与运算。

10.3 线性分组码

如果一个码组中的任意两个码字可以进行模2加法,并且产生的第三个码字仍然是该码组中的码字,我们称之为线路码。现在让我们来看一个(n, k)的线性分组码,在 n 比特码字中, k 个比特总是与要发送的信息比特相同,剩余部分的 $n - k$ 个比特是由信息比特根据指定的编码规则计算得到,编码规则决定了这种码的数学结构。因此,这 $n - k$ 个比特是指广义的奇偶校验比特或简称为奇偶比特。我们称发送过程中信息比特不发生变化的分组码为系统

码。对于需要既能检错又能纠错的应用,采用系统分组码可以简化译码器。

取 m_0, m_1, \dots, m_{k-1} 构成一任意 k 比特信息的分组。这样,我们可以获得 2^k 个不同的信息分组。将这个信息比特序列送入一线性分组的编码器,产生一个 n 比特的码字,其各个元素分别用 c_0, c_1, \dots, c_{n-1} 表示。以 $b_0, b_1, \dots, b_{n-k-1}$ 表示码字中的 $n-k$ 个校验比特。由于码字具有系统码的结构,所以一个码字可分为两个部分,一部分是信息比特,另一部分是校验比特。显然,我们可以选择在奇偶校验比特之前发送信息比特,反之亦可。前者如图 10.4 所示,其应用将在后面给出。

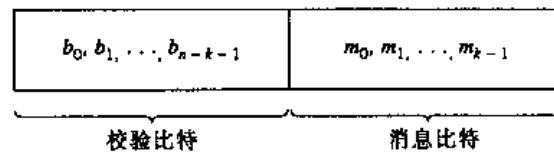


图 10.4 系统码的码字结构

根据图 10.4 的表示法,码字的最左边的 $n-k$ 个比特与相应的奇偶校验码相同,最右边 k 个比特与相应的信息比特相同。因此有

$$c_i = \begin{cases} b_i, & i = 0, 1, \dots, n-k-1 \\ m_{i+k-n}, & i = n-k, n-k+1, \dots, n-1 \end{cases} \quad (10.1)$$

$n-k$ 个校验比特是 k 个信息比特的线性和,可以用一般关系式表示:

$$b_i = p_{0i}m_0 + p_{1i}m_1 + \dots + p_{(k-1)i}m_{k-1} \quad (10.2)$$

系数的定义如下:

$$p_{ij} = \begin{cases} 1 & \text{如果 } b_i \text{ 依赖于 } m_j \\ 0 & \text{其他} \end{cases} \quad (10.3)$$

系数 p_{ij} 的选择要使生成矩阵的各行线性独立,且校验等式惟一。

式(10.1)和式(10.2)给出了 (n, k) 线性分组码的数学结构。这两个等式可用矩阵表示法重新表示为一种紧凑的形式。为此,我们定义 $1 \times k$ 的消息矢量(或信息矢量) \mathbf{m} , $1 \times (n-k)$ 的奇偶校验矢量 \mathbf{b} 和 $1 \times n$ 的码矢量 \mathbf{c} ,其形式分别如下:

$$\mathbf{m} = [m_0, m_1, \dots, m_{k-1}] \quad (10.4)$$

$$\mathbf{b} = [b_0, b_1, \dots, b_{n-k-1}] \quad (10.5)$$

$$\mathbf{c} = [c_0, c_1, \dots, c_{n-1}] \quad (10.6)$$

注意,这三个矢量都是行矢量。本章采用行矢量的表示方法是为了与编码相关文献中采用的表示法一致。这样就可用紧凑的矩阵形式将定义奇偶校验比特的联立等式写为

$$\mathbf{b} = \mathbf{mP} \quad (10.7)$$

其中, \mathbf{P} 为 $k \times (n-k)$ 的系数矩阵,其定义如下:

$$\mathbf{P} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0,n-k-1} \\ p_{10} & p_{11} & \cdots & p_{1,n-k-1} \\ \vdots & \vdots & & \vdots \\ p_{k-1,0} & p_{k-1,1} & \cdots & p_{k-1,n-k-1} \end{bmatrix} \quad (10.8)$$

其中, p_y 取值 1 或 0。

由式(10.4)~式(10.6)可知, \mathbf{c} 可以表示为由矢量 \mathbf{m} 和 \mathbf{b} 组成的分块矢量:

$$\mathbf{c} = [\mathbf{b} : \mathbf{m}] \quad (10.9)$$

将式(10.7)代入式(10.9), 并提出公因子(信息矢量) \mathbf{m} , 得

$$\mathbf{c} = \mathbf{m}[\mathbf{P} : \mathbf{I}_k] \quad (10.10)$$

\mathbf{I}_k 为 $k \times k$ 的单位矩阵:

$$\mathbf{I}_k = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \quad (10.11)$$

定义 $k \times n$ 的生成矩阵为

$$\mathbf{G} = [\mathbf{P} : \mathbf{I}_k] \quad (10.12)$$

式(10.12)给出的生成矩阵 \mathbf{G} 被认为是一种典型形式, 这是因为该矩阵的 k 行之间是线性独立的。也就是说, \mathbf{G} 中的任意一行都不可能表示为其他各行的线性组合。利用生成矩阵 \mathbf{G} 的定义, 可将式(10.10)简化为

$$\mathbf{c} = \mathbf{m}\mathbf{G} \quad (10.13)$$

使信息矢量 \mathbf{m} 在 2^k 个二进制 k 元组($1 \times k$ 矢量)的范围内变化, 并利用式(10.13)即可生成全部码字集合, 简称为码。且其中任意两个码字的和为另一个码字。线性分组码的这种特性称为封闭性。为证明这种特性的正确性, 考虑一对码矢量 \mathbf{c}_i 和 \mathbf{c}_j , 与之对应的是一对信息矢量 \mathbf{m}_i 和 \mathbf{m}_j 。根据式(10.13), \mathbf{c}_i 和 \mathbf{c}_j 的和可表示为

$$\begin{aligned} \mathbf{c}_i + \mathbf{c}_j &= \mathbf{m}_i\mathbf{G} + \mathbf{m}_j\mathbf{G} \\ &= (\mathbf{m}_i + \mathbf{m}_j)\mathbf{G} \end{aligned}$$

将 \mathbf{m}_i 和 \mathbf{m}_j 模 2 加, 得到一个新的信息矢量。相应地, 将 \mathbf{c}_i 和 \mathbf{c}_j 模 2 加也得到一个新的码矢量。

线性分组码的信息比特和奇偶校验比特的关系可以用另一种方法来表示。以 \mathbf{H} 表示一个 $(n-k) \times n$ 的矩阵, 其定义为

$$\mathbf{H} = [\mathbf{I}_{n-k} : \mathbf{P}^T] \quad (10.14)$$

其中, \mathbf{P}^T 为一个 $(n-k) \times k$ 的矩阵, 表示系数矩阵 \mathbf{P} 的转置, \mathbf{I}_{n-k} 为 $(n-k) \times (n-k)$ 的单位矩阵。因此分块阵的乘法运算如下:

$$\begin{aligned} \mathbf{H}\mathbf{G}^T &= [\mathbf{I}_{n-k} : \mathbf{P}^T] \begin{bmatrix} \mathbf{P}^T \\ \vdots \\ \mathbf{I}_k \end{bmatrix} \\ &= \mathbf{P}^T + \mathbf{P}^T \end{aligned}$$

其中用到了方阵的一个性质, 即方阵乘以维数相同的单位矩阵后, 其值不变。在模 2 运算中, 有 $\mathbf{P}^T + \mathbf{P}^T = \mathbf{0}$, $\mathbf{0}$ 表示 $(n-k) \times k$ 的空矩阵(所有元素都为零的矩阵)。所以

$$\mathbf{H}\mathbf{G}^T = \mathbf{0} \quad (10.15)$$

同样地,有 $\mathbf{G}\mathbf{H}^T = \mathbf{0}$,此处 $\mathbf{0}$ 是一个新的空矩阵。将式(10.13)右乘 \mathbf{H} 的转置矩阵 \mathbf{H}^T ,根据式(10.15),可得

$$\begin{aligned}\mathbf{c}\mathbf{H}^T &= \mathbf{m}\mathbf{G}\mathbf{H}^T \\ &= \mathbf{0}\end{aligned}\quad (10.16)$$

矩阵 \mathbf{H} 称为线性分组码的奇偶校验矩阵,式(10.16)所列的等式称为奇偶校验方程。

生成方程(10.13)和奇偶校验方程(10.16)是描述和运算线性分组码的基本方程。这两个方程可以用方框图的形式来描述,分别如图 10.5(a)和 10.5(b)所示。

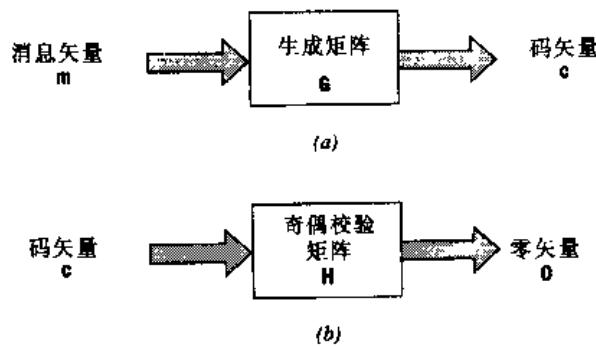


图 10.5 式(10.13)的生成方程和式(10.16)的奇偶校验方程的框图

例 10.1 重复码

重复码是最简单的线性分组码。特别地,单个信息比特被编码为 n 个相同比特的分组,生成了一个 $(n, 1)$ 的分组码。这种码允许规定不同数目的冗余,并且只含有两个码字:全 0 码和全 1 码。

例如,假设重复码中 $k=1, n=5$ 。在这种情况下,可得到 4 个与信息比特相同的奇偶校验比特。因此,单位矩阵 $\mathbf{I}_k=1$,系数矩阵 \mathbf{P} 由一个 1×4 的矢量构成,矢量的元素全为 1。相应地,生成矩阵为一个全 1 的行矢量:

$$\mathbf{G} = [1 \ 1 \ 1 \ 1 \ 1]$$

系数矩阵 \mathbf{P} 的转置矩阵 \mathbf{P}^T 则是一个 4×1 的全 1 矢量。单位矩阵 \mathbf{I}_{n-k} 为一个 4×4 的矩阵。因此,奇偶校验矩阵为:

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

由于信息矢量是由单个的二进制符号 0 或 1 组成,由式(10.13)可知,(5, 1)重复码中只有两个码字:00000 和 11111,这与先前的设想是一致的。此外, $\mathbf{HG}^T = \mathbf{0}$ 的模 2 运算的结果等于 0,这也与式(10.15)一致。

10.3.1 校正子: 定义和性质

生成矩阵 \mathbf{G} 应用于发射机的编码,而奇偶校验矩阵 \mathbf{H} 则用于接收机的译码。对于后者,我们引入一个 $1 \times n$ 的接收矢量 \mathbf{r} ,该矢量是由发送的码矢量 \mathbf{c} 经过噪声信道得到的。将矢量

r 表示为原始的码矢量 c 与另一个矢量 e 的和, 如下所示:

$$r = c + e \quad (10.17)$$

矢量 e 称为误码矢量或错误图样。如果 r 中第 i 个元素与 c 中的第 i 个元素相同, 则 e 中相应的元素为 0; 另一方面, 如果 r 中第 i 个元素与 c 中的第 i 个元素不同, 则 e 中相应的元素为 1, 这种情况就可以认为是在第 i 位产生了误码。也就是说, 对于 $i = 1, 2, \dots, n$, 有

$$e_i = \begin{cases} 1 & \text{在第 } i \text{ 位产生了误码} \\ 0 & \text{其他} \end{cases} \quad (10.18)$$

接收机的任务就是从接收的矢量 r 中译码出码矢量 c 。通常, 译码算法始于对一个称为误码校正子的 $1 \times (n - k)$ 矢量的计算, 误码校正子也可简称为校正子^[3]。校正子的重要之处在于它仅取决于错误图样。

假设有 $1 \times n$ 的接收矢量 r , 则与之对应的校正子可定义为

$$s = rH^T \quad (10.19)$$

校正子有如下的重要性质:

性质 1

校正子仅与错误图样有关, 而与发送的码字无关。

为证明这个性质, 先将式(10.17)代入式(10.19), 然后再代入式(10.16)得到

$$\begin{aligned} s &= (c + e)H^T \\ &= cH^T + eH^T \\ &= eH^T \end{aligned} \quad (10.20)$$

因此, 只要知道奇偶校验矩阵 H , 就可以计算出校正子 s 。 s 仅与错误图样 e 有关。

性质 2

不同码字的所有错误图样都有相同的校正子。

如果有 k 个信息比特, 就存在 2^k 个不同的码矢量 c_i , $i = 0, 1, \dots, 2^k - 1$ 。相应地, 对任意一种错误图样 e , 可定义 2^k 个不同的矢量 e_i :

$$e_i = e + c_i, \quad i = 0, 1, \dots, 2^k - 1 \quad (10.21)$$

其中, 这一系列矢量 $\{e_i, i = 0, 1, \dots, 2^k - 1\}$ 称为接收码字的陪集。换言之, 由于码矢量的不同, 一个陪集刚好包含 2^k 个不同的元素, 这些元素之间最多有一个码矢量的差异。这样, 一个 (n, k) 的线性分组码就具有 2^{n-k} 个可能的陪集。在式(10.21)的两边乘以矩阵 H^T , 可得

$$\begin{aligned} e_i H^T &= e H^T + c_i H^T \\ &= e H^T \end{aligned} \quad (10.22)$$

这个结果是与下标 i 无关的。因此可以说, 码字的每个陪集有惟一的校正子。

将式(10.20)展开, 可以帮助我们全面理解性质 1 和性质 2, 特别是当矩阵 H 采用式(10.14)的系统形式代入时。其中, 矩阵 P 如式(10.8)所定义, 从式(10.20)中不难发现, 校正子的 $n - k$ 个元素 s 是错误图样 e 的 n 个元素的线性组合, 即

$$\begin{aligned}
 s_0 &= e_0 + e_{n-k} p_{00} + e_{n-k+1} p_{10} + \cdots + e_{n-1} p_{k-1,0} \\
 s_1 &= e_1 + e_{n-k} p_{01} + e_{n-k+1} p_{11} + \cdots + e_{n-1} p_{k-1,1} \\
 &\vdots \\
 s_{n-k-1} &= e_{n-k-1} + e_{n-k} p_{0,n-k-1} + \cdots + e_{n-1} p_{k-1,n-k-1}
 \end{aligned} \tag{10.23}$$

显然,这 $n - k$ 个方程组说明校正子中包含与错误图样有关的信息,因此可用于误码的检测。但需要注意的是,这些方程组还是不够的,因为未知数的个数多于方程的个数。因此,错误图样没有惟一解。当然,满足方程(10.23)的错误图样有 2^n 个,但它们都对应于与性质 2 和等式(10.22)相符的同一个校正子。特别地,虽然有 2^{n-k} 个可能出现的校正子矢量,但包含在校正子 s 中的关于错误图样 e 的信息,还不足以使译码器准确地计算出发射端发出的码矢量。但是,校正子 s 使真实错误图样的搜索范围由原来的 2^n 种可能减少为 2^{n-k} 种可能。这样,译码器的任务就是从对应于 s 的陪集中得出最佳选择。

10.3.2 最小距离的考虑

假设有两个码矢量 \mathbf{c}_1 和 \mathbf{c}_2 ,其元素个数相同。将两个码矢量中相应位置上具有不同元素的位置个数定义为汉明距离,记为 $d(\mathbf{c}_1, \mathbf{c}_2)$ 。

将码矢量中非零元素的个数定义为码矢量 \mathbf{c} 的汉明重量 $w(\mathbf{c})$ 。同样地,可以认为码矢量的汉明重量是该码矢量与全零码之间的距离。

线性分组码的最小距离 d_{\min} 是指码组中任意两个码矢量之间的最小汉明距离。换言之,最小距离等于任意两个码矢量之差的最小汉明重量。由线性分组码的封闭性可知,两个码矢量的和(或差)得到另一个码矢量,因此,可以认为线性分组码的最小距离就是码组中非零码矢量的最小汉明重量。

从根本上讲,最小距离 d_{\min} 与码组的奇偶校验矩阵 \mathbf{H} 的结构有关。由式(10.16)可知,线性分组码是由满足 $\mathbf{c}\mathbf{H}^T = 0$ 的所有码矢量组成,其中, \mathbf{H}^T 为奇偶校验矩阵 \mathbf{H} 的转置矩阵。将此矩阵 \mathbf{H} 表示成列矢量的形式,有

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n] \tag{10.24}$$

这样,满足 $\mathbf{c}\mathbf{H}^T = 0$ 的码矢量 \mathbf{c} 在一些位置上必须为 1,在这些位置上, \mathbf{H}^T 中的相应行求和为 0。然而,根据定义,码矢量中 1 的个数为码矢量的汉明重量,且线性分组码中非零码矢量的最小汉明重量就是该码的最小距离。因此,线性分组码的最小距离可以定义为矩阵 \mathbf{H}^T 中行求和为 0 的矢量的最少行数。

在线性分组码中,最小距离是一个重要的参数,因为它决定了码的纠错能力。假设有一 (n, k) 的线性分组码,要求能检测并纠正所有的错误图样(通过二进制对称信道),其中该码的汉明重量小于等于 t 。也就是说,如果发送码组中的一个码矢量 \mathbf{c}_i ,接收矢量为 $\mathbf{r} = \mathbf{e} + \mathbf{c}_i$,则只要求错误图样的汉明重量 $w(\mathbf{e}) \leq t$,译码器的输出 $\hat{\mathbf{c}} = \mathbf{c}_i$ 。假设码组中的 2^k 个码矢量以相等的概率发送,译码器可采用的最好策略就是选取最接近接收矢量 \mathbf{r} 的码矢量,也就是选择汉明距离 $d(\mathbf{c}, \mathbf{r})$ 最小的码矢量。如果线性分组码的最小距离大于等于 $2t + 1$,采用这种方法,译码器就可以检测并纠正所有汉明重量 $w(\mathbf{e}) \leq t$ 的错误图样。下面通过这个问题的几何解释来证明上述条件的正确性。实际上, $1 \times n$ 的码矢量和 $1 \times n$ 的接收矢量可以用 n 维空间的点来表示。构造两个球,半径均为 t ,码矢量 \mathbf{c}_i 和 \mathbf{c}_j 分别表示两个球心。如图 10.6(a)所示,

这两个球是不相交的。要满足这个条件,就要求 $d(\mathbf{c}_i, \mathbf{c}_j) \geq 2t + 1$ 。如果发送的码矢量为 \mathbf{c}_i , 并且汉明距离 $d(\mathbf{c}_i, \mathbf{r}) \leq t$, 显然,译码器应该选择 \mathbf{c}_i 作为最接近接收矢量 \mathbf{r} 的码矢量。换言之,如果汉明距离 $d(\mathbf{c}_i, \mathbf{c}_j) \leq 2t$,围绕着 \mathbf{c}_i 和 \mathbf{c}_j 的两个球的交叉部分则如图 10.6(b)所示。这时,如果发送 \mathbf{c}_i ,就会存在满足上述条件的接收矢量 \mathbf{r} 。虽然汉明距离 $d(\mathbf{c}_i, \mathbf{r}) \leq t$,但 \mathbf{r} 与 \mathbf{c}_i 和 \mathbf{c}_j 的距离相等。显然,此时存在着译码器错选矢量 \mathbf{c}_j 的可能性。因此,我们可以得出结论:线性分组码(n, k)能够纠正汉明距离为 t 的错误图样的充要条件是

$$d(\mathbf{c}_i, \mathbf{c}_j) \geq 2t + 1 \quad \text{对于所有 } \mathbf{c}_i \text{ 和 } \mathbf{c}_j$$

由定义可知,码组中任意两个码矢量的最小距离就是该码的最小距离 d_{\min} 。因此,我们可以说,一个最小距离为 d_{\min} 的线性分组码(n, k)最多可纠正 t 个错误的充要条件是

$$t \leq \lfloor \frac{1}{2}(d_{\min} - 1) \rfloor \quad (10.25)$$

上式中 $\lfloor \cdot \rfloor$ 表示小于或等于内部量的最大整数。式(10.25)定量地说明了线性分组码的纠错能力。

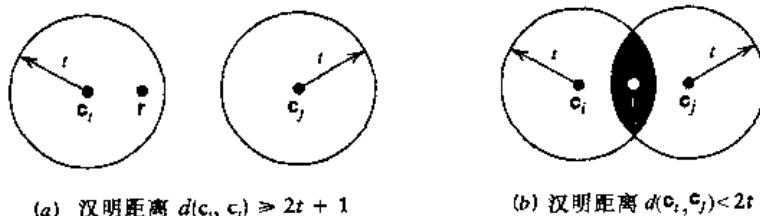


图 10.6 汉明距离;接收矢量用 \mathbf{r} 表示

10.3.3 校正子译码

下面讨论基于校正子的线性分组码的译码方案。用 $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_{2^k}$ 表示线性分组码(n, k)的 2^k 个码矢量, \mathbf{r} 表示接收矢量,且有 2^n 个可能的取值。接收机的任务就是将这 2^n 个可能的接收矢量分别放入 2^k 个互不相交的子集 $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_{2^k}$ 中, 第 i 个子集 \mathcal{D}_i 与码矢量 \mathbf{c}_i 相对应, 其中, $1 \leq i \leq 2^k$ 。如果接收矢量 \mathbf{r} 落在第 i 个子集中, 则将被译码为 \mathbf{c}_i 。为了能正确地译码, 接收矢量 \mathbf{r} 必须落在实际发送码矢量 \mathbf{c}_i 所属的那个子集中。

这里所述的 2^k 个子集组成了一个线性分组码的标准矩阵。我们可以利用码组的线性结构,按如下步骤来构造标准矩阵:

1. 2^k 个码矢量排成一行,全零码矢量 \mathbf{c}_1 放在最左边。
2. 选取一种错误图样 \mathbf{e}_2 置于 \mathbf{c}_1 之下, 将 \mathbf{e}_2 加到第一行剩余的各码矢量上, 可得到第二行。其中最关键的一点是,一行中第一个元素的错误图样在此前必须从未在标准矩阵中出现过。
3. 重复第二步,直到所有可能的错误图样都得以解决。

图 10.7 给出了按上述步骤构造的标准矩阵的结构。矩阵的 2^k 列表示互不相交的子集 $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_{2^k}$, 2^{n-k} 行表示码组的陪集, 陪集第一个元素 $\mathbf{e}^2, \dots, \mathbf{e}_{2^{n-k}}$ 称为陪集首。

对于一个给定的信道,当选择最有可能出现(发生概率最大)的错误图样作为陪集首时,译码的错误概率最小。在二进制对称信道中,错误图样的汉明重量越小,则其发生的可能性就越

大。因此，标准矩阵应该由陪集中汉明重量最小的陪集首来构造。

下面将线性分组码的译码过程描述为：

1. 对于接收矢量 r , 计算校正子 $s = rH^T$ 。
2. 在以校正子 s 作为表征的陪集中, 鉴别出陪集首(具有最大发生概率的错误图样) e_0 。
3. 计算码矢量

$$c = r + e_0 \quad (10.26)$$

作为接收矢量 r 的译码输出。

上述过程称为校正子译码。

$$\begin{array}{ccccccc} c_1 = 0 & c_2 & c_3 & \cdots & c_i & \cdots & c_{2^k} \\ 0_2 & 0_2 + 0_2 & 0_3 + 0_2 & \cdots & 0_i + 0_2 & \cdots & 0_{2^k} + 0_2 \\ 0_3 & 0_2 + 0_3 & 0_3 + 0_3 & \cdots & 0_i + 0_3 & \cdots & 0_{2^k} + 0_3 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 0_j & 0_2 + 0_j & 0_3 + 0_j & \cdots & 0_i + 0_j & \cdots & 0_{2^k} + 0_j \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ 0_{2^{n-k}} & 0_2 + 0_{2^{n-k}} & 0_3 + 0_{2^{n-k}} & \cdots & 0_i + 0_{2^{n-k}} & \cdots & 0_{2^k} + 0_{2^{n-k}} \end{array}$$

图 10.7 (n, k) 分组码的标准阵列

例 10.2 汉明码^[4]

假设一组(n, k)的线性分组码具有如下的参数：

$$\begin{array}{ll} \text{分组长度:} & n = 2^m - 1 \\ \text{消息比特数:} & k = 2^m - m - 1 \\ \text{奇偶校验比特数:} & n - k = m \end{array}$$

其中, $m \geq 3$ 。这种码称为汉明码。

例如,(7, 4)汉明码对应的 $n = 7, k = 4, m = 3$ 。则码的生成矩阵必须具有与式(10.12)一致的结构。下面的矩阵就是一个正确的(7, 4)汉明码的生成矩阵：

$$G = \left[\begin{array}{cccc|cccc} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right] \quad \begin{matrix} P \\ I_k \end{matrix}$$

相应的奇偶校验矩阵如下：

$$H = \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{array} \right] \quad \begin{matrix} I_{n-k} \\ P^T \end{matrix}$$

表 10.1 为(7, 4)汉明码的码字表。当 $k = 4$ 时, 共有 $2^k = 16$ 种不同的信息码字。对于每个给定的信息码字, 根据式(10.13)可以得到一个相应的码字。应用这个等式, 可得到列于表 10.1 中的 16 种码字。

表 10.1 (7,4)汉明码的码字

消息码字	码字	码重	消息码字	码字	码重
0000	0000000	0	1000	1101000	3
0001	1010001	3	1001	0111001	4
0010	1110010	4	1010	0011010	3
0011	0100011	3	1011	1001011	4
0100	0110100	3	1100	1011100	4
0101	1100101	4	1101	0001101	3
0110	1000110	3	1110	0101110	4
0111	0010111	4	1111	1111111	7

在表 10.1 中, 还列出了(7,4)汉明码的每个码字的汉明重量。由非零码的最小汉明重量为 3 可得, 码的最小距离也是 3。事实上, 汉明码具有如下性质: 最小距离 $d_{\min} = 3$, 且独立于奇偶校验比特数 m 。

为了证明最短距离 d_{\min} 与奇偶校验矩阵 H 的关系, 假设有一个码为 0110100。按照式(10.16)定义的矩阵乘法, 用码字的非零元素“筛选”出矩阵 H 的第 2, 第 3 和第 5 列得

$$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

对剩余的 14 个非零码字, 可以进行类似的计算。不难发现, 在 H 中求和为 0 的最小列数为 3, 从而证明了前面 $d_{\min} = 3$ 的论述。

汉明码具有一个重要性质, 即假定 $t = 1$, 则汉明码满足式(10.25)取等号时对应的条件。也就是说, 汉明码是纠正单个错误的二进制完备码。

若发生的是单个错误, 则可以列出 7 个陪集首, 如表 10.2 的右边一列所示。列于左侧的相应的 2^3 个校正子是根据式(10.20)计算得出的。零校正子意味着传输中无错误。

表 10.2 表 10.1 所示(7,4)汉明码的译码表

校 码	错 误 图 样
000	0000000
100	1000000
010	0100000
001	0010000
110	0001000
011	0000100
111	0000010
101	0000001

例如, 假设发送的码矢量是 [1110010], 接收矢量在第 3 位发生误码, 为 [1100010]。根据式(10.19), 校正子为

$$s = [1100010] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = [0 \ 0 \ 1]$$

再由表 10.2 可知,与之对应的陪集首(发生概率最高的错误图样)为 [0010000],说明接收矢量中第 3 位发生了错误。于是,根据式(10.26),在接收矢量上加上错误图样,就能够得到实际发送的正确码字。

10.3.4 对偶码

设有一线性分组码,可以按下述方式定义其对偶码。将式(10.15)的两边同时转置,得

$$\mathbf{G}\mathbf{H}^T = \mathbf{0}$$

\mathbf{H}^T 是线性分组码的奇偶校验矩阵的转置矩阵, $\mathbf{0}$ 是一个新的零矩阵。这个等式说明,生成矩阵为 \mathbf{G} 、奇偶校验矩阵为 \mathbf{H} 的 (n, k) 线性分组码均存在一个参数为 $(n, n-k)$ 、生成矩阵为 \mathbf{H} 、奇偶校验矩阵为 \mathbf{G} 的对偶码。

10.4 循环码

循环码是线性分组码的一个子类。现有的重要线性分组码大都是循环码或与循环码密切相关。与其他大多数类型的码相比,循环码的优势在于它易于编码。此外,循环码具有明确定义的数学结构,因此比较容易得到对应的有效译码方案。

二进制码只要满足以下的两个基本特性,就可以称为循环码:

1. 线性特性,即码组中任意两个码的和仍然在这个码组中。
2. 循环性,即码组中任一码字经循环移位后仍然在这个码组中。

特性 1 重申了循环码是线性分组码(也可看成是奇偶校验码)。为了从数学的角度重新说明特性 2,取 n 维数组 $(c_0, c_1, \dots, c_{n-1})$ 来表示线性分组码 (n, k) 的一个码字。如果 n 维数组

$$\begin{aligned} &(c_{n-1}, c_0, \dots, c_{n-2}), \\ &(c_{n-2}, c_{n-1}, \dots, c_{n-3}), \\ &\vdots \\ &(c_1, c_2, \dots, c_{n-1}, c_0) \end{aligned}$$

都是码组中的码字,此码组就是循环码。

为了研究循环码的代数特性,用码字的各元素 c_0, c_1, \dots, c_{n-1} 来定义码多项式:

$$c(X) = c_0 + c_1X + c_2X^2 + \dots + c_{n-1}X^{n-1} \quad (10.27)$$

其中, X 是一个未知数。自然地,对二进制码来说,系数为 1 或 0。多项式 $c(X)$ 中 X 的每个幂

次表示1比特的移位。因此,多项式 $c(X)$ 乘以 X 可以看成是右移一位。关键问题是怎样产生这样的循环移位。下面就回答这个问题。

将多项式 $c(X)$ 乘上 X^i ,得到

$$\begin{aligned} X^i c(X) &= X^i(c_0 + c_1X + \cdots + c_{n-i-1}X^{n-i-1} + c_{n-i}X^{n-i} \\ &\quad + \cdots + c_{n-1}X^{n-1}) \\ &= c_0X^i + c_1X^{i+1} + \cdots + c_{n-i-1}X^{n-1} + c_{n-i}X^n \\ &\quad + \cdots + c_{n-1}X^{n+i-1} \\ &= c_{n-i}X^n + \cdots + c_{n-1}X^{n+i-1} + c_0X^i + c_1X^{i+1} \\ &\quad + \cdots + c_{n-i-1}X^{n-1} \end{aligned} \quad (10.28)$$

最后一行仅调整了上一行中各项的顺序。在模2加中, $c_{n-i} + c_{n-i-1} = 0$,因此可对式的前 i 项进行如下处理:

$$\begin{aligned} X^i c(X) &= c_{n-i} + \cdots + c_{n-1}X^{i-1} + c_0X^i + c_1X^{i+1} + \cdots + c_{n-i-1}X^{n-1} \\ &\quad + c_{n-i}(X^n + 1) + \cdots + c_{n-1}X^{i-1}(X^n + 1) \end{aligned} \quad (10.29)$$

接下来,引入下面的定义:

$$\begin{aligned} c^{(i)}(X) &= c_{n-i} + \cdots + c_{n-1}X^{i-1} + c_0X^i + c_1X^{i+1} \\ &\quad + \cdots + c_{n-i-1}X^{n-1} \end{aligned} \quad (10.30)$$

$$q(X) = c_{n-i} + c_{n-i+1}X + \cdots + c_{n-1}X^{i-1} \quad (10.31)$$

因此,式(10.29)可写为紧凑形式:

$$X^i c(X) = q(X)(X^n + 1) + c^{(i)}(X) \quad (10.32)$$

多项式 $c^{(i)}(X)$ 是码字 $(c_0, c_1, \dots, c_{n-i-1}, c_{n-i}, \dots, c_{n-1})$ 经过 i 次循环移位得到的码字 $(c_{n-i}, \dots, c_{n-1}, c_0, c_1, \dots, c_{n-i-1})$ 的码多项式。由式(10.32)不难看出, $c^{(i)}(X)$ 是 $X^i c(X)$ 除以 $(X^n + 1)$ 的余数。这样就可以正式地用多项式表示法来说明循环特性:如果 $c(X)$ 是一个码多项式,那么多项式

$$c^{(i)}(X) = X^i c(X) \bmod(X^n + 1) \quad (10.33)$$

是循环移位 i 次后的码多项式。 \bmod 是取模的缩写形式。式(10.33)中所描述的多项式相乘的特殊形式可看成是模 $X^n + 1$ 的乘法运算。实际上,在这个乘法运算中,有 $X^n = 1$,这也使得多项式中所有 $i < n$ 的 $X^i c(X)$ 变为 $n - 1$ 次。(注意:在模2的代数运算中, $X^n + 1$ 和 $X^n - 1$ 是相等的。)

10.4.1 生成多项式

在循环码的生成中,多项式 $X^n + 1$ 及其因子起着主要的作用。令 $g(X)$ 为 $X^n + 1$ 的 $n - k$ 阶多项式因子。同样, $g(X)$ 是码组中最低阶的多项式。通常,对 $g(X)$ 可以做如下展开:

$$g(X) = 1 + \sum_{i=1}^{n-k-1} g_i X^i + X^{n-k} \quad (10.34)$$

系数 g_i 等于1或0。按照这个展开式,多项式 $g(X)$ 中系数为1的有两项,且被 $n - k - 1$ 项分割开。 $g(X)$ 称为循环码的生成多项式。一个循环码是由生成多项式 $g(X)$ 唯一确定的,这是

因为,码组中的每个码多项式都可以表示为如下的多项式乘积的形式:

$$c(X) = a(X)g(X) \quad (10.35)$$

其中, $a(X)$ 是 X 的 $k - 1$ 阶多项式。由此构成的 $c(X)$ 满足式(10.33)的条件,因为 $g(X)$ 是 $X^n + 1$ 的因子。

假设给定了一个生成多项式 $g(X)$,并要求将信息序列 $(m_0, m_1, \dots, m_{k-1})$ 编码成一个 (n, k) 系统循环码。也就是说,信息比特以不变的形式发送,其结构如下(见图 10.4):

$$\begin{array}{c} (b_0, b_1, \dots, b_{n-k-1}, \underbrace{m_0, m_1, \dots, m_{k-1}}) \\ \hline n-k \text{ 个奇偶比特} \qquad \qquad k \text{ 个消息比特} \end{array}$$

令信息多项式为

$$m(X) = m_0 + m_1X + \dots + m_{k-1}X^{k-1} \quad (10.36)$$

令

$$b(X) = b_0 + b_1X + \dots + b_{n-k-1}X^{n-k-1} \quad (10.37)$$

根据式(10.1),将码多项式化为如下形式:

$$c(X) = b(X) + X^{n-k}m(X) \quad (10.38)$$

因此,由式(10.35)和式(10.38)可得

$$a(X)g(X) = b(X) + X^{n-k}m(X)$$

同样,根据模 2 加法的性质,有

$$\frac{X^{n-k}m(X)}{g(X)} = a(X) + \frac{b(X)}{g(X)} \quad (10.39)$$

式(10.39)说明,多项式 $b(X)$ 是 $X^{n-k}m(X)$ 除以 $g(X)$ 的余式。

下面将系统结构的 (n, k) 循环码的编码步骤总结为:

1. 信息多项式 $m(X)$ 乘以 X^{n-k} 。
2. 用多项式 $X^{n-k}m(X)$ 除以生成多项式 $g(X)$,得到余式 $b(X)$ 。
3. $b(X)$ 与 $X^{n-k}m(X)$ 相加,得到码多项式 $c(X)$ 。

10.4.2 奇偶校验多项式

一个 (n, k) 循环码是由它的 $(n - k)$ 阶生成多项式 $g(X)$ 惟一确定的。这个码也可以由另一个 k 阶多项式惟一确定,该多项式称为奇偶校验多项式。其定义如下:

$$h(X) = 1 + \sum_{i=1}^{k-1} h_i X^i + X^k \quad (10.40)$$

其中,系数 h_i 为 0 或 1。校验多项式具有与生成多项式相似的形式,其中也有两项的系数等于 1,且被 $k - 1$ 项分开。

用生成多项式 $g(X)$ 来描述码组,与用生成矩阵 \mathbf{G} 来描述码组是等价的。相应地,奇偶校验多项式 $h(X)$ 与奇偶校验矩阵 \mathbf{H} 的表示法也是等价的。不难发现,由式(10.15)表示的线性分组码的矩阵关系 $\mathbf{HG}^T = 0$ 与下面的关系是一致的:

$$g(X)b(X) \bmod(X^n + 1) = 0 \quad (10.41)$$

因此,可以说生成多项式 $g(X)$ 和校验多项式 $b(X)$ 都是多项式 $X^n + 1$ 的因子,即

$$g(X)b(X) = X^n + 1 \quad (10.42)$$

这个性质提供了选择循环码的生成多项式或校验多项式的基本方法。特别地,可以认为如果 $g(X)$ 是 $n - k$ 阶多项式,并且是 $X^n + 1$ 的因子,那么 $g(X)$ 就是一个 (n, k) 循环码的生成多项式。同理,也可以认为,如果 $b(X)$ 是一个 k 阶多项式,并且是 $X^n + 1$ 的因子,那么 $b(X)$ 就是一个 (n, k) 循环码的奇偶校验多项式。

任意一个 $X^n + 1$ 的 $n - k$ (校验比特的数目) 阶因子均可以作为生成多项式。当 n 值较大时,多项式 $X^n + 1$ 可以有许多的 $n - k$ 阶因子。其中,有些多项式因子可生成好的循环码,但有些却只能生成坏的循环码。怎样选择那些能生成好循环码的生成多项式,是非常困难的一件事。事实上,编码理论学家们已经在寻找好的循环码上进行了许多努力。

10.4.3 生成矩阵和奇偶校验矩阵

假设一个 (n, k) 循环码的生成多项式是 $g(X)$,可通过记录生成码组的 k 个多项式 $g(X), Xg(X), \dots, X^{k-1}g(X)$ 来构造生成矩阵 \mathbf{G} 。因此,与多项式相对应的 n 维数组可以作为 $k \times n$ 阶生成矩阵 \mathbf{G} 的行。

然而,由循环码的奇偶校验多项式 $b(X)$ 构造奇偶校验矩阵 \mathbf{H} 时,有些方面需要特别注意。将式(10.42)乘上 $a(x)$,然后利用式(10.35)得到

$$c(X)b(X) = a(X) + X^n a(X) \quad (10.43)$$

多项式 $c(X)$ 和式 $b(X)$ 分别由式(10.27)和式(10.40)定义,也就是说,式(10.43)中左边的乘积包含最高幂次为 $n + k - 1$ 的项。另一方面,由于 $a(X)$ 的幂次为 $k - 1$ 或更小,也就是说,在式(10.43)右边的多项式中,不会出现幂次为 $X^k, X^{k+1}, \dots, X^{n-1}$ 的项。令 $c(X)b(X)$ 的乘积多项式的展开式中 $X^k, X^{k+1}, \dots, X^{n-1}$ 的系数等于 0,可得到一个 $n - k$ 个方程组成的方程组

$$\sum_{i=j}^{j+k} c_i b_{k+i-j} = 0 \quad 0 \leq j \leq n - k - 1 \quad (10.44)$$

对比式(10.44)与式(10.16)之间的对应关系,可以得到如下重要发现:式(10.44)所描述的多项式乘积中包含的校验多项式 $b(X)$ 的系数,与式(10.16)中描述的形成矢量内积中包含的奇偶校验矩阵 \mathbf{H} 的系数排列顺序正好相反。根据这个发现,可定义奇偶校验多项式的倒数为

$$\begin{aligned} X^k b(X^{-1}) &= X^k \left(1 + \sum_{i=1}^{k-1} b_i X^{-i} + X^{-k} \right) \\ &= 1 + \sum_{i=1}^{k-1} b_{k-i} X^i + X^k \end{aligned} \quad (10.45)$$

它也是 $X^n + 1$ 的因子。因此,对应于 $n - k$ 个多项式 $X^k b(X^{-1}), X^{k+1} b(X^{-1}), \dots, X^{n-1} b(X^{-1})$ 的 n 维数组可以作为 $(n - k) \times n$ 的奇偶校验矩阵 \mathbf{H} 的行。

通常,按照上述方法构造的生成矩阵 \mathbf{G} 和奇偶校验矩阵 \mathbf{H} 都不是系统形式的,但通过各行的初等变换就可以将它们转变成系统形式,如例 10.3 所示。

10.4.4 循环码编码器

前面已经指出,系统形式的(n, k)循环码的编码过程分为三步:(1)信息多项式 $m(X)$ 乘以 X^{n-k} ;(2) $X^{n-k}m(X)$ 除以生成多项式 $g(X)$,得到余式 $b(X)$;(3) $b(X)$ 加上 $X^{n-k}m(X)$ 。这样就可得到我们所期望的码多项式。这三步可以通过图 10.8 所示的编码器来实现,该编码器由 $n - k$ 级线性反馈移位寄存器构成。

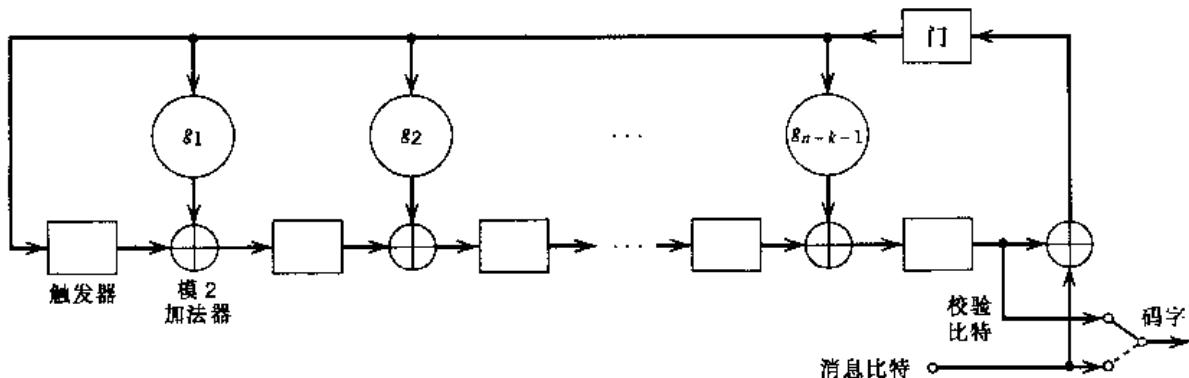


图 10.8 (n, k)循环码的编码器

图 10.8 中的方框表示触发器,或者是单位延时元件。触发器是能驻留在某一种状态的器件。其状态有两种可能性,分别用 0 或 1 表示。一个外部的时钟(未在图 10.8 中示出)控制所有触发器的运行。时钟每跳动一次,触发器(初始化为 0)的内容就按箭头的方向移出。除触发器以外,图 10.8 的编码器还包含有第二种逻辑元件,称为加法器。加法器能够计算各自输入的模 2 和。最后,乘法器将其输入乘上相关系数。特别地,如果系数 $g_i = 1$,乘法器就是直接的“连接”。另一方面,如果系数 $g_i = 0$,乘法器就相当于是断开的。

图 10.8 所示编码器的工作步骤如下:

1. 门打开, k 个信息比特被送入信道。一旦 k 个信息比特进入移位寄存器,寄存器中得到的 $n - k$ 个比特就形成奇偶校验比特。奇偶校验比特与余式 $b(X)$ 的系数相等。
2. 门关闭,从而切断反馈连接。
3. 移位寄存器中的内容被读入到信道。

10.4.5 校正子的计算

假设码字 $(c_0, c_1, \dots, c_{n-1})$ 经过噪声信道在接收端得到的接收码字为 $(r_0, r_1, \dots, r_{n-1})$ 。由 10.3 节可知,在线性分组码的译码过程中,第一步就是计算接收码字的校正子。如果校正子为 0,则接收码字中没有因传输造成的误码。如果校正子不等于 0,则接收的码字中存在传输误码,有待纠错。

在系统循环码中,校正子可以比较容易地计算出来。用一个 $n - 1$ 阶或更低阶的多项式表示接收码字,即

$$r(X) = r_0 + r_1X + \dots + r_{n-1}X^{n-1} \quad (10.46)$$

用 $q(X)$ 表示 $r(X)$ 除以 $g(X)$ 的商式, $s(X)$ 表示余式。因此, $r(X)$ 可表示为

$$r(X) = q(X)g(X) + s(X) \quad (10.47)$$

余式 $s(X)$ 是 $n - k - 1$ 阶或更低阶的多项式。由于余式的系数组成了 $(n - k) \times 1$ 的校正子，所以称为校正子多项式。

图 10.9 为一个校正子计算器，除了接收比特是由左边的反馈移位寄存器的第 $n - k$ 级输入的以外，它与图 10.8 所示的编码器基本相同。当所有接收比特都被移入移位寄存器后，校正子 s 就相应地确定了。

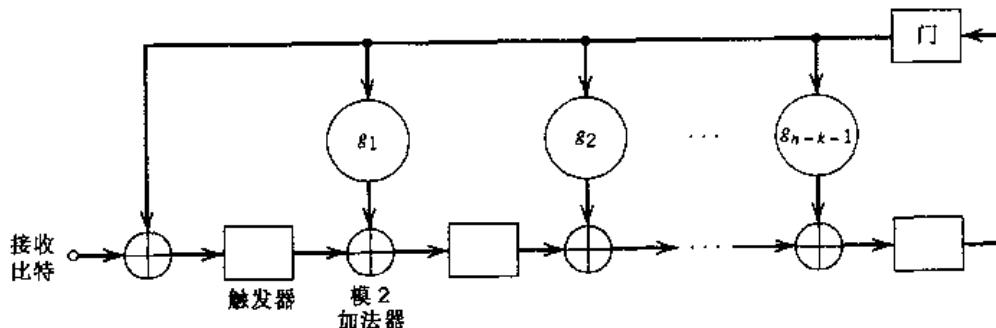


图 10.9 (n, k) 循环码的校正子计算器

由式(10.47)的定义可知，校正子多项式 $s(X)$ 具有如下的有用特性：

1. 接收码多项式的校正子也是相应的误码多项式的校正子。

设有一循环码，其码多项式为 $c(X)$ ，将其发送至一噪声信道，则接收码多项式的定义为

$$r(X) = c(X) + e(X) \quad (10.48)$$

其中， $e(X)$ 是误码多项式。同理可得

$$e(X) = r(X) + u(X) \quad (10.49)$$

将式(10.35)和式(10.47)代入式(10.49)，可得

$$e(X) = u(X)g(X) + s(X) \quad (10.50)$$

其中， $u(X) = a(X) + q(X)$ 。式(10.50)说明， $s(X)$ 也是误码多项式 $e(X)$ 的校正子。这个特性表明，当校正子多项式 $s(X)$ 不为零时，将检测到接收码字中存在着传输误码。

2. 如果接收码多项式 $r(X)$ 的校正子是 $s(X)$ ，那么， $r(X)$ 循环移位后的 $Xr(X)$ 的校正子是 $Xs(X)$ 。

对式(10.47)的两边同时进行循环移位，可得

$$Xr(X) = Xq(X)g(X) + Xs(X) \quad (10.51)$$

因此，不难看出 $Xs(X)$ 是 $Xr(X)$ 除以 $g(X)$ 的余式。于是， $Xs(X)$ 是 $Xr(X)$ 的校正子。可将这个结论归纳为：如果 $s(X)$ 是 $r(X)$ 的校正子，则 $Xs(X)$ 是 $Xr(X)$ 的校正子。

3. 假如误码被限制为 $n - k$ 个接收码多项式 $r(X)$ 的奇偶校验比特，那么校正子多项式 $s(X)$ 与误码多项式 $e(X)$ 是相同的。

这里所做的假设等价于：误码多项式 $e(X)$ 的阶数小于或等于 $(n - k - 1)$ 。由于生成多项式 $g(X)$ 是 $n - k$ 阶的，根据定义，只有在 $u(X)$ 等于 0 时，才能满足式(10.50)的要求。

求。换言之,误码多项式 $e(X)$ 和校正子多项式 $s(X)$ 是同一个(或相等的)多项式。因此,性质 3 说明,在上述条件下,纠错可以很容易地通过将校正子多项式 $s(X)$ 加上接收码多项式 $r(X)$ 来实现。

例 10.3 汉明码的回顾

为了举例说明与循环码的多项式表示法有关的一个观点,考虑一组(7,4)循环码的产生过程。由于分组的长度为 $n = 7$,先将 $X^7 + 1$ 分解成三个不可约的多项式:

$$X^7 + 1 = (1 + X)(1 + X^2 + X^3)(1 + X + X^3)$$

不可约多项式是指已经无法用二进制系数的多项式来进行因式分解的多项式。如果一个 m 阶的不可约多项式能整除 $X^n + 1$,且最小正整数 $n = 2^m - 1$,则称这个不可约多项式为本原多项式。例如 $(1 + X^2 + X^3)$ 和 $(1 + X + X^3)$ 就是两个本原多项式。取

$$g(X) = 1 + X + X^3$$

作为生成多项式,其阶数等于奇偶校验比特数。也就是说,奇偶校验多项式

$$\begin{aligned} h(X) &= (1 + X)(1 + X^2 + X^3) \\ &= 1 + X + X^2 + X^4 \end{aligned}$$

的阶数等于信息比特数 $k = 4$ 。

下面用该多项式作为生成多项式对信息序列 1001 进行编码,以此为例来说明构造一个码字的过程。其中相应的信息多项式为

$$m(X) = 1 + X^3$$

因此,再乘上 $X^{n-k} = X^3$ 得

$$X^{n-k}m(X) = X^3 + X^6$$

第二步是用 $X^{n-k}m(X)$ 除以 $g(X)$,具体过程如下:

$$\begin{array}{r} X^3 + X \\ \hline X^3 + X + 1 | X^6 + X^3 \\ \hline X^6 + X^4 + X^3 \\ \hline X^4 \\ \hline X^4 + X^2 + X \\ \hline X^2 + X \end{array}$$

注意,在这个长除式中,我们认为减法和加法是相同的,因为这里进行的是模 2 的运算。于是有

$$\frac{X^3 + X^6}{1 + X + X^3} = X + X^3 + \frac{X + X^2}{1 + X + X^3}$$

也就是说,商式 $a(X)$ 和余式 $b(X)$ 分别为

$$\begin{aligned} a(X) &= X + X^3 \\ b(X) &= X + X^2 \end{aligned}$$

因此,根据式(10.38),可以得到所需的多项式

$$\begin{aligned} c(X) &= b(X) + X^{n-k}m(X) \\ &= X + X^2 + X^3 + X^6 \end{aligned}$$

得到的码字为 0111001。最右边的四个比特 1001 就是信息比特。左边的三个比特 011 是奇偶校验比特。这样产生的码字和图 10.1 中给出的(7, 4)汉明码中相应的码字相同。

我们可以将这一结果归纳为：任意由本原多项式产生的循环码是最小距离为 3 的汉明码。下面说明生成多项式 $g(X)$ 和奇偶校验多项式 $b(X)$ 分别惟一确定生成矩阵 \mathbf{G} 和奇偶校验矩阵 \mathbf{H} 的过程。

为了构造 4×7 的生成矩阵 \mathbf{G} ，从 $g(X)$ 和三个 $g(X)$ 循环移位对应的四个多项式入手，即

$$\begin{aligned} g(X) &= 1 + X + X^3 \\ Xg(X) &= X + X^2 + X^4 \\ X^2g(X) &= X^2 + X^3 + X^5 \\ X^3g(X) &= X^3 + X^4 + X^6 \end{aligned}$$

多项式 $g(X), Xg(X), X^2g(X)$ 和 $X^3g(X)$ 分别表示(7, 4)汉明码的码多项式。如果将这些多项式的系数作为 4×7 的矩阵各行的元素，那么，就可以得到如下的生成矩阵：

$$\mathbf{G}' = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

显然，按照这种方法构造的生成矩阵 \mathbf{G}' 不是系统形式的。可以通过将第 1 行加到第 3 行上，将前两行的和加到第 4 行上将该矩阵转化成系统形式。经过这些处理我们就能够得到想要的生成矩阵

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

这正好与例 10.2 的结果相同。

接下来介绍怎样由奇偶校验多项式 $b(X)$ 构造 3×7 的奇偶校验矩阵 \mathbf{H} 。为此，先取 $b(X)$ 的倒数 $X^4 b(X^{-1})$ ，然后构造 $X^4 b(X^{-1})$ 和它的两个移位对应的三个多项式，即

$$\begin{aligned} X^4 b(X^{-1}) &= 1 + X^2 + X^3 + X^4 \\ X^5 b(X^{-1}) &= X + X^2 + X^4 + X^5 \\ X^6 b(X^{-1}) &= X^2 + X^4 + X^5 + X^6 \end{aligned}$$

用这三个多项式的系数作为 3×7 奇偶校验矩阵各行的元素，有

$$\mathbf{H}' = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

矩阵 \mathbf{H}' 也不是系统形式的，为将其转化为系统形式，将第 3 行加到第 1 行上，得到

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

这也和例 10.2 中结果相同。

图 10.10 为由多项式 $g(X) = 1 + X + X^3$ 生成的(7,4)循环汉明码的编码器。为了举例说明编码器的工作过程,假设输入一个信息序列 1001。如表 10.3 所示,随着不同信息比特的输入,寄存器的内容也相应地发生变化。经过 4 次移位之后,寄存器中的内容变为 011,即奇偶校验比特。加上这些奇偶校验比特,可得码字 0111001。这个结果与前面例题中得出的结论相同。

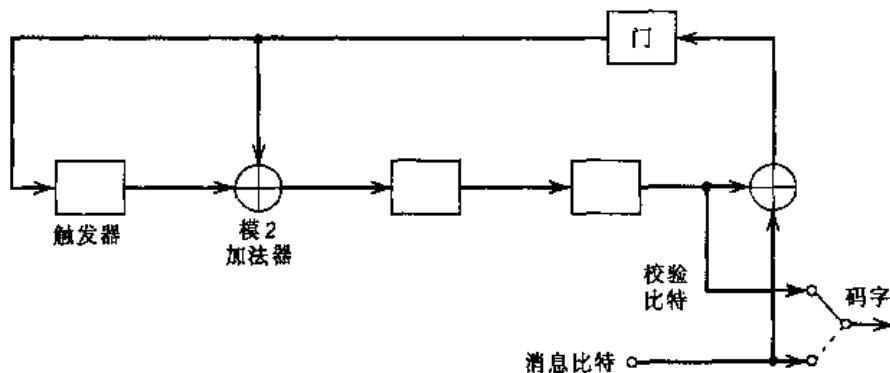


图 10.10 由多项式 $g(X) = 1 + X + X^3$ 生成的(7,4)循环汉明码的编码器

表 10.3 输入信息序列 1001 时图 10.10 所示编码器中移位寄存器的内容

移位	输入	寄存器内容
		000(初始状态)
1	1	110
2	0	011
3	0	111
4	1	011

图 10.11 给出了相应的(7,4)汉明码校正子计算器。如果发送的码字为 0111001,而接收的码字为 0110001,也就是说中间比特发生了误码。移位寄存器初始化为 0,当接收比特被送入移位寄存器时,寄存器的内容就发生了改变,如表 10.4 所示。当第 7 次移位结束时,校正子就可以从寄存器的内容中被识别出来,即为 110。由于校正子不是全 0,所以接收码字中存在着误码。此外,由表 10.2 可知,对应于这个校正子的错误图样是 0001000。这说明误码的确是发生在接收码字的中间比特。

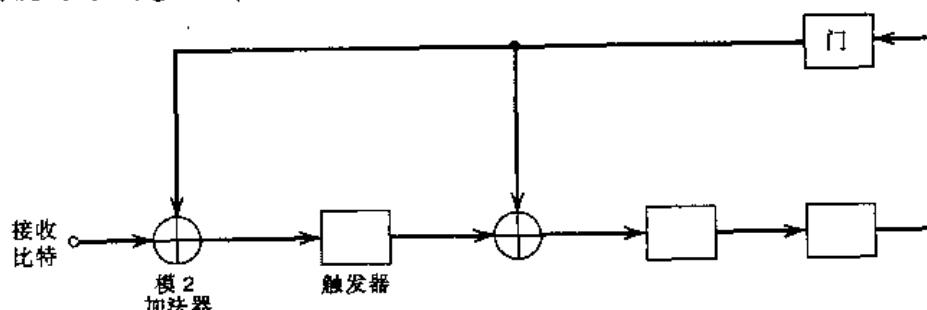


图 10.11 由多项式 $g(X) = 1 + X + X^3$ 生成的(7,4)循环汉明码的校正子计算器

表 10.4 接收比特为 0110001 时, 图 10.11 所示校正子计算器中寄存器的内容

移位	输入比特	寄存器内容
		000(初始状态)
1	1	100
2	0	010
3	0	001
4	0	110
5	1	111
6	1	001
7	0	110

例 10.4 最大长度码

对于任意的正整数 $m \geq 3$, 存在着一个最大长度码, 其参数如下:

$$\begin{array}{ll} \text{分组长度:} & n = 2^m - 1 \\ \text{消息比特数:} & k = m \\ \text{最小距离:} & d_{\min} = 2^{m-1} \end{array}$$

最大长度码的生成多项式为

$$g(X) = \frac{1 + X^n}{b(X)} \quad (10.52)$$

其中, $b(X)$ 是一个 m 阶的本原多项式。前面我们已提到, 由本原多项式生成的任何循环码都是最小距离为 3 的汉明码(见例 10.3)。因此, 最大长度码就是汉明码的对偶码。

多项式 $b(X)$ 决定了编码器中的反馈连接方式。假如编码器的初始状态是 00…01, 那么, 生成多项式 $g(X)$ 就定义了一个周期的最大长度码。为了说明这一点, 以一个(7,3)的最大长度码为例。如例 10.3 所述, 这个码就是(7,4)汉明码的对偶码。选取

$$b(X) = 1 + X + X^3$$

则(7,3)最大长度码的生成多项式是

$$g(X) = 1 + X + X^2 + X^4$$

图 10.12 为(7,3)最大长度码的编码器, 其反馈连接与第 8 章图 8.2 中完全相同。码周期为 $n = 7$ 。因此, 假设编码器的初始状态为 001(如图 10.12 所示), 则输出序列为

$$\underbrace{\begin{matrix} 1 & 0 & 0 \end{matrix}}_{\text{初始状态}} \quad \underbrace{\begin{matrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \end{matrix}}_{g(X) = 1 + X + X^2 + X^4}$$

通过图 10.12 所示编码器的循环, 可以容易地验证上述结果。

需要注意的是, 如果我们选择其他的本原多项式, 如

$$b(X) = 1 + X^2 + X^3$$

来产生(7,3)最大长度码, 则只能得到上述码的“图像”, 输出序列将被及时地“翻转”。

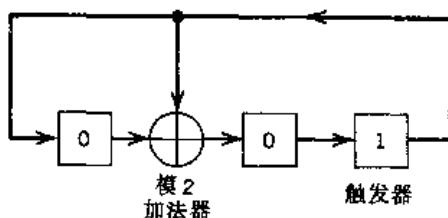


图 10.12 (7,3)最大长度码编码器(图中为编码器的初始状态)

10.4.6 其他循环码

下面以对其他三类重要的循环码特征的介绍来结束循环码的讨论。

循环冗余校验码

循环码特别适合于进行误码检测,主要有两个原因。第一,循环码可以被设计来检测可能误码的多种组合;第二,循环码中用于实现编码和误码检测的电路是实用的。正因如此,实际应用中许多能进行误码检测的码都属于循环码。用于误码检测的循环码称为循环冗余检测(CRC)码。

如果 B 比特连续码序列的首位和末位,或中间任意位在接收时发生误码,则定义该连续码序列为 n 比特接收码字中的长度为 B 的突发错误。二进制的 (n, k) CRC 码可以检测下面的错误图样:

1. 长度小于等于 $n - k$ 的突发错误。
2. 部分长度为 $n - k + 1$ 的突发错误,可检测的部分占总数的 $1 - 2^{-(n-k+1)}$ 。
3. 部分长度大于 $n - k + 1$ 的突发错误,可检测的部分占总数的 $1 - 2^{-(n-k+1)}$ 。
4. 所有码距 $\leq d_{\min} - 1$ 的误码的组合。
5. 码组的生成多项式 $g(X)$ 中包含偶数个非零系数时,所有的奇数个错误。

表 10.5 列出了成为国际标准的三个 CRC 码的生成多项式。所有这些码都包含基本因子 $1 + X$ 。CRC-12 码用于 6 比特字符,其他两个码用于 8 比特字符。CRC 码为 10.1 节讨论的 ARQ 协议和第 4 章讨论的 DSL 中的错误检测提供了一种有效的方法。

表 10.5 CRC 码

码字	生成多项式 $g(X)$	$n - k$
CRC-12 码	$1 + X + X^2 + X^3 + X^{11} + X^{12}$	12
CRC-16 码(美国)	$1 + X^2 + X^{15} + X^{16}$	16
CRC-ITU 码	$1 + X^5 + X^{12} + X^{16}$	16

BCH 码^[5]

BCH(Bose-Chaudhuri-Hocquenghem)码是最重要且最有效的线性分组码之一,也是参数可大范围变化的循环码。本原 BCH 码是最常见的二进制 BCH 码。这种码具有如下特征:对于任意的正整数 m (大于或等于 3)和小于 $(2^m - 1)/2$ 的 t ,有

分组长度:	$n = 2^m - 1$
消息比特数:	$k \geq n - mt$
最小距离:	$d_{\min} \geq 2t + 1$

每个 BCH 码都是纠正 t 个错误的码 (t -error correcting code), 因为它的每个码字能检测并纠正多达 t 个随机的误码。纠正单个误码的汉明码也可以称为 BCH 码。BCH 码能灵活地选择码参数, 如分组长度和码率。而且当分组长度为几百或更少时, BCH 码被公认为同样分组长度和码率的编码中最好的码之一。

详细分析 BCH 码的构造已超出了我们当前讨论的范围。为了对 BCH 码的性能有更多的感性认识, 我们在表 10.6 中给出了长度达 $2^5 - 1$ 的二进制分组 BCH 码的码参数的生成多项式。例如, 我们现在想构造(15, 7)BCH 码的生成多项式, 由表 10.6 可得, 生成多项式的系数为 (111 010 001), 即

$$g(X) = X^8 + X^7 + X^6 + X^4 + 1$$

表 10.6 码长达 $2^5 - 1$ 的二进制 BCH 码

n	k	t	生成多项式					
			1	011	10	011	111	010
7	4	1						001
15	11	1					10	011
15	7	2				111	010	001
15	5	3			10	100	110	111
31	26	1					100	101
31	21	2			11	101	101	001
31	16	3		1	000	111	110	101
31	11	5	101	100	010	011	011	010
31	6	7	11	001	011	011	110	101
						000	100	111

注: n = 分组长度

k = 消息比特数

t = 可检测的最大差错数

生成多项式 $g(X)$ 的高阶系数在左边。

里德 - 索罗蒙码^[6]

里德 - 索罗蒙 (Reed-Solomon) 码是非二进制 BCH 码的一个重要子类, 常简称为 RS 码。RS 码的编码器不同于二进制的编码器, 这是因为其中采用的是多比特, 而不是单比特。一个 (n, k) RS 码用来将一个符号组编码为由 $n = 2^m - 1$ 个符号组成的分组, 其中, 每个符号由 m 比特组成。也就是说, 共有 $m(2^m - 1)$ 个比特, 其中 $m \geq 1$ 。这样, 编码算法就是, 通过加上 $n - k$ 个冗余的符号, 将一个 k 个符号的分组扩充为 n 个符号的分组。当 m 等于 2 的整数次幂时, 称这 m 比特为一个字节。通常 m 等于 8, 8 比特的 RS 码非常有用。

可纠正 t 个错误的 RS 码具有如下参数:

分组长度:	$n = 2^m - 1$
消息长度:	k
部分校验长度:	$n - k = 2t$
最小距离:	$d_{\min} = 2t + 1$

RS 码的分组长度比码符号的数目小 1, 最小距离比奇偶校验符号的数目多 1。RS 码提高了冗余的利用率, 且分组长度和符号的大小可以随时调整, 以适应不同的信息长度。同时, RS 码还

提供了大范围的码率,这样就可以选择不同的码率来优化性能。最后要指出的一点是,采用RS码可以使有效的译码技术得到运用,这也是它得到广泛应用的另一个原因(例如在CD数字音响系统中)。

10.5 卷积码^[1]

在分组码中,编码器接收一个 k 比特的信息分组并生成一个 n 比特的码字。这样,码字就是逐组产生的。显然,在生成相关的码字之前,编码器必须缓存完整的信息分组。但是,存在这样一种应用,其中信息比特是串行输入而不是一组同时输入的。在这种情况下,采用缓存器是不现实的。这种情况下,采用卷积码是一种更好的方法。卷积码通过模2的卷积来产生冗余,由此而得名。

如果以比特来表示每个符号,则码率为 $1/n$ 的二进制卷积码的编码器可以看成是有限状态机。该编码器由一个 M 级的移位寄存器组成,每级移位寄存器分别与指定的模2加法器相连,在加法器的输出端串接一个复接器。一个长度为 L 比特的信息序列编码后,产生长度为 $n(L+M)$ 比特的输出序列。因此码率为

$$r = \frac{L}{n(L+M)} \quad \text{比特/符号} \quad (10.53)$$

通常,取 $L \gg M$,所以码率又可以简化为

$$r \approx \frac{1}{n} \quad \text{比特/符号} \quad (10.54)$$

卷积码的约束长度用信息比特来表示,可定义为单个比特可影响编码器输出的移位个数。在有 M 级移位寄存器的编码器中,编码器的存储器容量为 M 个信息比特。从一个信息比特进入移位寄存器直到它从中输出,共需要 $K = M + 1$ 次移位。因此,编码器的约束长度为 K 。

图10.13(a)所示为 $n=2, K=3$ 的卷积码编码器。编码器的码率为 $1/2$ 。图10.13(a)所示的编码器每次处理输入信息序列中的一个比特。

我们可以用连接 n 个模2加法器、一个输入复接器和一个输出复接器的相互独立的 k 个移位寄存器来生成码率为 k/n 的二进制卷积码。图10.13(b)为这种编码器的一个例子,其中 $k=2, n=3$,在两个移位寄存器中都有 $K=2$,码率等于 $2/3$ 。在这个例子中编码器一次处理2比特的信息序列。

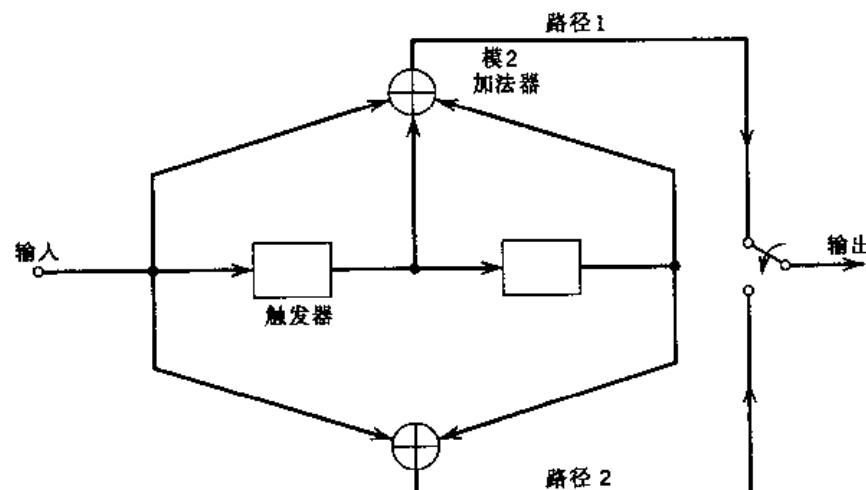
图10.13中编码器产生的卷积码是非系统码。与分组码不同,在卷积码中非系统码的应用一般多于系统码。

卷积码编码器中每一条连接输出到输入的路径特性,都可以根据每个触发器被初始化为零状态后该路径的冲激响应来表述。冲激响应就是向这条路径输入符号1时得到的响应。同样,如果用生成多项式来进行表述,则该路径特性可定义为冲激响应的单位时延变换。具体地说,就是取生成序列 $(g_0^{(i)}, g_1^{(i)}, g_2^{(i)}, \dots, g_M^{(i)})$ 表示第 i 条路径的冲激响应。其中,系数 $g_0^{(i)}, g_1^{(i)}, g_2^{(i)}, \dots, g_M^{(i)}$ 等于0或1。对应的第 i 条路径的生成多项式定义为

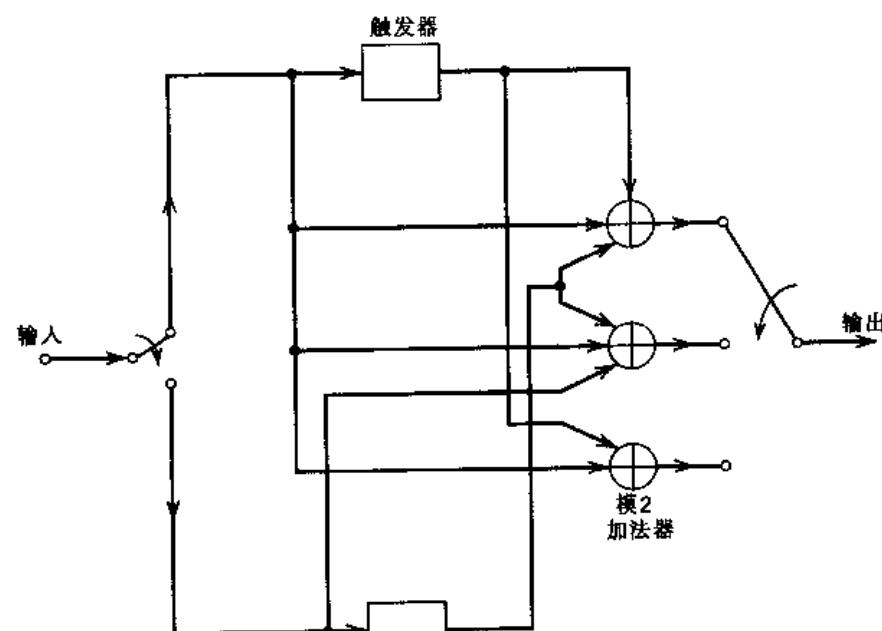
$$g^{(i)}(D) = g_0^{(i)} + g_1^{(i)}D + g_2^{(i)}D^2 + \dots + g_M^{(i)}D^M \quad (10.55)$$

其中, D 表示单位时延变量。完整的卷积码编码器可以用一组生成多项式 $\{g^{(1)}(D), g^{(2)}(D), \dots\}$

$\dots, g^{(n)}(D)$ 来描述。按照惯例, 卷积码和分组码采用不同的变量来表示, 一般 D 用于表示卷积码, X 用于表示分组码。



(a) 约束长度为 3, 码率为 1/2 的卷积码编码器



(b) 约束长度为 2, 码率为 2/3 的卷积码编码器

图 10.13 卷积码编码器

例 10.5

设有一卷积码编码器如图 10.13(a)所示。为方便起见, 将它的两条路径分别标为 1 和 2。路径 1(上面的路径)的冲激响应为 1, 1, 1。因此, 相应的生成多项式为

$$g^{(1)}(D) = 1 + D + D^2$$

路径 2(下面的路径)的冲激响应为 1, 0, 1。因此, 相应的生成多项式为

$$g^{(2)}(D) = 1 + D^2$$

也可将信息序列 10011 用多项式表示为

$$m(D) = 1 + D^3 + D^4$$

由傅里叶变换可知,时域的卷积可以化为 D 域的乘积。因此,路径 1 的输出多项式为

$$\begin{aligned} c^{(1)}(D) &= g^{(1)}(D)m(D) \\ &= (1 + D + D^2)(1 + D^3 + D^4) \\ &= 1 + D + D^2 + D^3 + D^6 \end{aligned}$$

由此可立即推出路径 1 的输出序列为 1111001。类似地,路径 2 的输出序列为

$$\begin{aligned} c^{(2)}(D) &= g^{(2)}(D)m(D) \\ &= (1 + D^2)(1 + D^3 + D^4) \\ &= 1 + D^2 + D^3 + D^4 + D^5 + D^6 \end{aligned}$$

所以,路径 2 的输出序列为 1011111。最后,将路径 1 和路径 2 的输出序列相结合,就可以得到编码的输出序列为

$$c = (11, 10, 11, 11, 01, 01, 11)$$

于是,长度 $L = 5$ 比特的信息序列编码后生成了长度为 $n(L + K - 1) = 14$ 比特的输出序列,我们还注意到由于移位寄存器被恢复到零初始状态,输出序列的最后加上了 $K - 1 = 2$ 个 0 组成的结束序列。该结束序列称为信息尾。

10.5.1 码树、网格图和状态图

通常,卷积码的结构特性可以用下面三种等效的图示法中的任意一种来表示:码树、网格图和状态图。下面以图 10.13(a)所示的卷积码编码器为例来说明这三种图示法的实质。

首先从图 10.14 所示的码树开始。树的每个分支表示一个输入符号,相应的二进制输出符号标注在分支上。通常采用如下方法来区分输入的二进制符号 0 和 1。用分叉点以上的分支表示输入为 0,用分叉点以下的分支表示输入为 1。码树中每一个由左至右的特定路径都对应于一个输入(信息)序列,路径中各分支上的相应编码符号组成了输出序列。例如,将信息序列(10011)加到图 10.13(a)所示的编码器输入端。按上述步骤,可得到相应的编码序列为(11, 10, 11, 11, 01),这与例 10.5 中得到的 $\{c_i\}$ 的前 5 对比特相同。

从图 10.14 所示的码树中还可以发现,在前 3 个分支以后,码树开始重复。事实上,在第 3 个分支以后,标 a 的两个节点是完全相同的。其他同标号的节点对也是如此。码树的这个特性可通过图 10.13(a)所示的编码器来进行检验。由于编码器只能存储 $M = K - 1 = 2$ 个信息比特,因此,当第 3 个比特进入编码器时,第 1 个比特就会移出寄存器。所以,第 3 个分支以后,信息序列 $100m_3m_4\dots$ 和 $000m_3m_4\dots$ 生成的符号码字相同,因此可将标号为 a 的两个节点合并到一起。同理,其他节点也如此。因此,可将图 10.14 所示的码树折叠得到一个如图 10.15 所示的新形式,称为网格图^[8]。之所以这样命名,是因为网格图是一种树形结构,其中重新合并了各分支。图 10.15 中用来区分输入符号是 0 还是 1 的常用方法是:用实线表示输入为 0 的分支,用虚线表示输入为 1 的分支。和前面一样,每个输入(信息)序列对应于一条通过网格图的特定路径。例如在图 10.15 中,输入信息序列 10011 对应的输出编码序列是(11, 10, 11, 11, 01),这与前面的结果一致。

网格图比码树更具启发性,因为从中能明显地看出对应的卷积码编码器是有限状态机。我们可用存在于编码器的移位寄存器中的 $K - 1$ 个信息比特来定义码率为 $1/n$ 的卷积码状态。在 j 时刻,含有相邻 K 比特的信息序列部分记为 $m_{j-K+1}, \dots, m_{j-1}, m_j$,其中, m_j 是当前比特。因此,编码器在 j 时刻的 $K - 1$ 个状态比特可以简单地写成 $m_{j-1}, \dots, m_{j-K+2}, m_{j-K+1}$ 。

在图 10.13(a)所示简单卷积码编码器的情况下,有 $K - 1 = 2$ 。因此,编码器的状态可以是 4 个可能状态中的任意一个,如表 10.7 所示。网格图包含 $L + K$ 级,其中, L 是输入的信息序列的长度, K 是码的约束长度。网格的各级在图 10.15 中分别标为 $j = 0, 1, \dots, L + K - 1$, 其中 $K = 3$ 。级数 j 也可以看成是深度 j 。这两个概念可以互换。前 $K - 1$ 级对应于编码器脱离初始状态的过程,最后 $K - 1$ 级对应于编码器回到初始状态的过程。显然,网格图中前 $K - 1$ 级和后 $K - 1$ 级所对应的编码器状态,并非都是可以到达的。但是,在网格图的中间部分,也就是 j 位于 $K - 1 \leq j \leq L$ 的范围内时,对应的编码器的所有状态都是可以到达的。值得注意的是,网格图的中间部分表现出其固有的周期性结构。

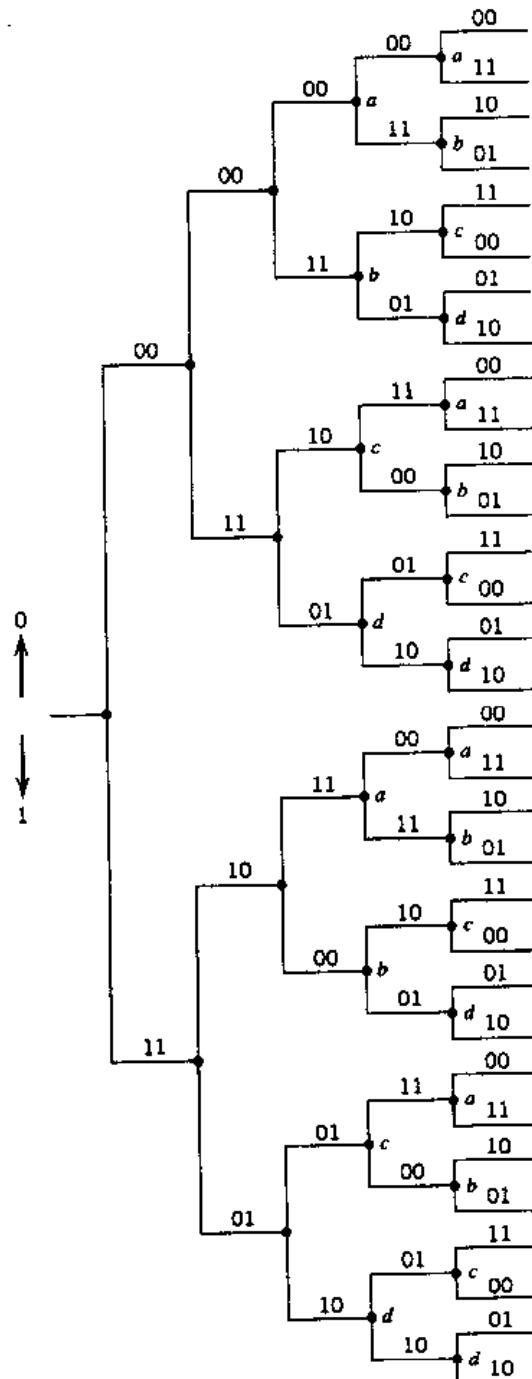


图 10.14 图 10.13(a)所示的卷积码编码器的码树

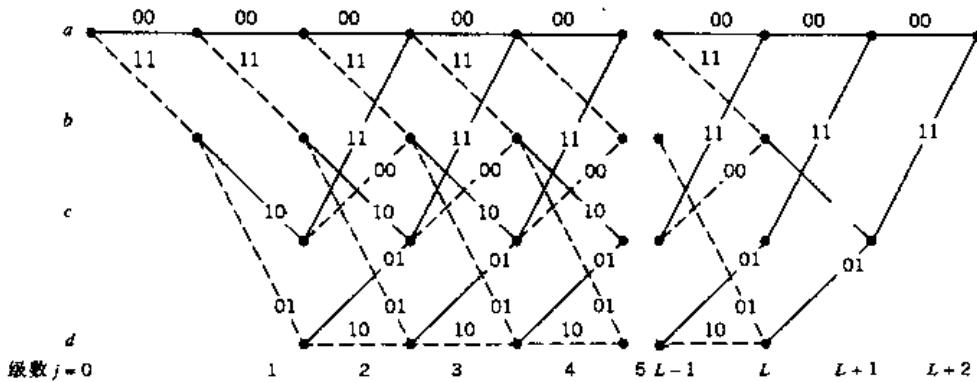
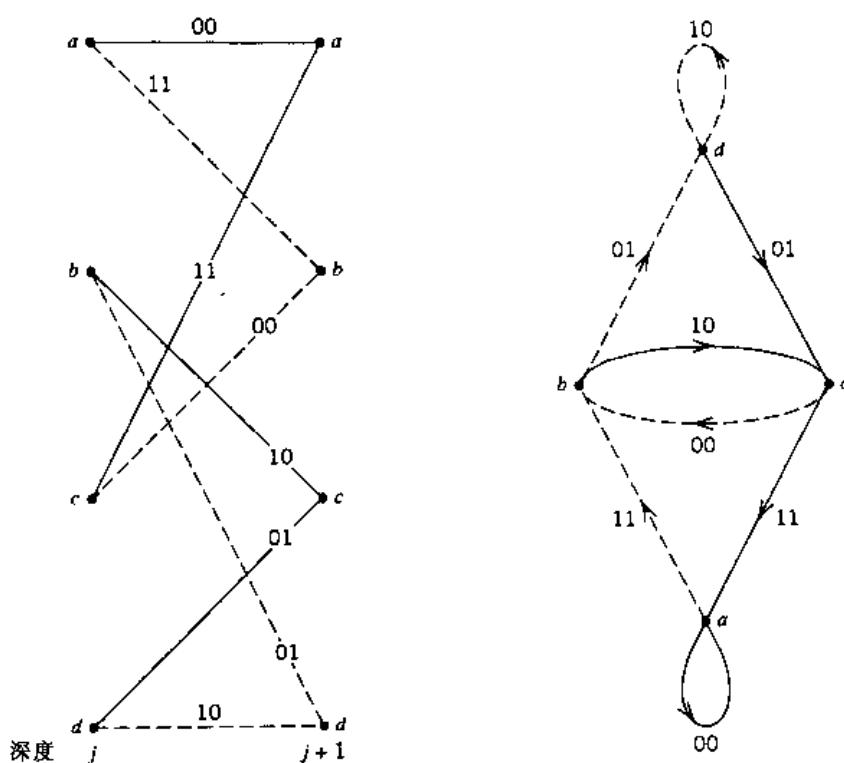


图 10.15 图 10.13(a)所示的卷积码编码器的网格图

表 10.7 图 10.13(a)所示卷积码的状态表

状态	二进制描述
a	00
b	10
c	01
d	11

假设网格图的下一部分对应于时刻 j 和 $j+1$, 本例中设 $j \geq 2$ 。于是, 编码器的当前状态可能是 a, b, c, d 中的任意一个。为了便于叙述, 我们在图 10.16(a)中重新列出这部分。左边的 4 个节点表示编码器的 4 个可能状态, 右边的节点表示相应的下一个状态。显然, 我们可以将左右两边的节点相结合。于是得到编码器的状态图, 如图 10.16(b)所示。图中的节点表示编码器的 4 个可能状态, 每个节点有两个输入分支和两个输出分支。对于输入 0 的状态转变用实线分支表示, 对于输入 1 的状态转变则用虚线表示。每个分支上的二进制标号表示编码器由一个状态转变到另一个状态时的输出。例如, 设编码器的当前状态是 01, 用节点 c 来表示。图 10.13(a)所示编码器的输入为 1, 则状态转变为 10, 且编码器输出为 00。因此, 对于任何输入信息序列, 都可以借助状态图很容易地得出图 10.13(a)所示编码器的输出。从全 0 的初始状态 a 开始, 根据信息序列的不同, 确定不同的状态图前进路径。如果输入是 0, 则走实线分支, 如果输入是 1, 则走虚线分支。每通过一个分支, 分支上相应的二进制标号就是对应的输出。例如, 输入的信息序列为 10011, 状态图通过的路径是 $abca b d$, 则相应的输出序列为 11, 10, 11, 11, 01, 这与前面得到的结果相同。因此, 卷积码的输入 - 输出关系也可以完全用状态图来表述。



(a) 图 10.13(a) 所示编码器的网格图中间段的部分 (b) 图 10.13(a) 所示编码器的状态图

图 10.16 图 10.13(a) 所示编码器网格图的中间部分和状态图

10.6 卷积码的最大似然译码

理解了卷积码的编码过程之后,下面将讨论卷积码的译码。在这一节,首先介绍最大似然译码的基本理论,然后提出一种可实际应用的有效算法。

以 \mathbf{m} 表示信息矢量, \mathbf{c} 表示经过编码器后送入离散无记忆信道的输入码矢量, \mathbf{r} 表示接收矢量。由于信道噪声的原因, \mathbf{r} 有可能与发送的码矢量不一致。译码器根据接收矢量 \mathbf{r} 来得到信息矢量的一个估计值 $\hat{\mathbf{m}}$ 。由于信息矢量 \mathbf{m} 和码矢量 \mathbf{c} 是一一对应的,所以译码器能够等效地求出码矢量的估计值 $\hat{\mathbf{c}}$ 。当且仅当 $\hat{\mathbf{m}} = \mathbf{m}$ 时,有 $\hat{\mathbf{c}} = \mathbf{c}$ 。否则,接收机的译码器必定存在误码。如果给定了接收矢量 \mathbf{r} ,只有当译码的误比特率最小时,用来选择估计值 $\hat{\mathbf{c}}$ 的译码规则才是最优的。由第 6 章可知,对于等概的信息,如果估计值 $\hat{\mathbf{c}}$ 选择最大的对数似然函数,则译码器的误比特率最小。取 $p(\mathbf{r}|\mathbf{c})$ 表示发送矢量 \mathbf{c} 并接收到矢量 \mathbf{r} 的条件概率,则相应的对数似然函数为 $\log p(\mathbf{r}|\mathbf{c})$ 。最大似然译码或判决的准则如下:

$$\text{选取估计值 } \hat{\mathbf{c}}, \text{使得对数似然函数 } \log p(\mathbf{r}|\mathbf{c}) \text{ 为最大。} \quad (10.56)$$

假设信道为二进制对称信道,且发送码矢量 \mathbf{c} 和接收矢量 \mathbf{r} 都是长度为 N 的二进制序列。由于信道噪声造成的误码,这两个序列之间在某些位上会出现相互不同的情况。分别以 c_i 和 r_i 表示 \mathbf{c} 和 \mathbf{r} 的第 i 位元素。于是有

$$p(\mathbf{r}|\mathbf{c}) = \prod_{i=1}^N p(r_i|c_i) \quad (10.57)$$

相应的对数似然函数为

$$\log p(\mathbf{r}|\mathbf{c}) = \sum_{i=1}^N \log p(r_i|c_i) \quad (10.58)$$

定义转移概率 $p(r_i|c_i)$ 为

$$p(r_i|c_i) = \begin{cases} p, & \text{如果 } r_i \neq c_i \\ 1 - p, & \text{如果 } r_i = c_i \end{cases} \quad (10.59)$$

设接收矢量 \mathbf{r} 和发送的码矢量 \mathbf{c} 正好在位置 d 上不同。编号 d 是矢量 \mathbf{r} 和矢量 \mathbf{c} 之间的汉明距离。因此,式(10.58)可写为

$$\begin{aligned} \log p(\mathbf{r}|\mathbf{c}) &= d \log p + (N - d) \log(1 - p) \\ &= d \log\left(\frac{p}{1-p}\right) + N \log(1 - p) \end{aligned} \quad (10.60)$$

通常,当 $p < 1/2$ 时,可认为误比特率已经足够低。注意到 $N \log(1 - p)$ 相对于矢量 \mathbf{c} 为一个常数。因此,可将二进制对称信道的最大似然译码规则重述为:

选取估计值 $\hat{\mathbf{c}}$,使得接收矢量 \mathbf{r} 和发送码矢量 \mathbf{c} 之间的汉明距离最小 (10.61)

也就是说,对于二进制对称信道而言,最大似然译码可简化为最小码距译码。在这种译码器中,将接收矢量 \mathbf{r} 与所有可能的发送矢量 \mathbf{c} 进行对比,选择最接近 \mathbf{r} 的一个作为正确的发送码矢量。这里“最接近”的意思是指矢量之间不同的二进制符号最少(即汉明距离最小)。

10.6.1 维特比算法^[9]

根据二进制对称信道的最大似然译码和最小码距译码的等效性,我们现在可以通过在码的码树中选择一条路径来对卷积码进行译码,其中要求这条路径的编码序列与接收序列之间的差异最小。由于码的码树与网格图之间是等效的,所以,可等效地采用网格图的表示法来限制可能路径的选择。网格图优于码树的原因主要是网格图的状态数不会随着输入比特数的增加而持续增长,而是始终保持在常数 2^{K-1} ,其中, K 是码的约束长度。

例如,设有一码率为 $r = 1/2$ 、约束长度 $K = 3$ 的卷积码的网格图如图 10.15 所示。通过观察,我们发现在 $j = 3$ 的这一级的 4 个节点中,每个节点都有两条进入的路径,并且这两条路径都从这个节点继续向前延伸。显然,最小码距译码器在这一点需要进行判决:要使系统性能不受损失,应该保留哪条路径。在 $j = 4$ 这一级也要进行相似的判决,以此类推。这个顺序的判决就是当卷积码通过网格图时维特比(Viterbi)算法所要做的工作,即计算网格图中每条可能路径之间的距离度量或差异。某一指定路径的距离度量是指这条路径所表示的编码序列和接收序列之间的汉明距离。这样,对图 10.15 中所示的网格图中的每个节点(状态),维特比算法将比较进入它们的两条路径。距离度量小的路径将被保留,另一条路径则被丢弃。如此在网格图的每一级(即第 j 级)上进行重复计算, $M \leq j \leq L$,其中, $M = K - 1$ 是编码器的存储器个数, L 是输入的信息序列的长度。保留下来的那条路径称为留存路径或激活路径。对于约束长度 $K = 3$ 的卷积码,只有 $2^{K-1} = 4$ 条留存路径及其距离度量被保存。 2^{K-1} 条路径的列表总能保证包含有最大似然的选择。

在维特比算法的应用中存在一个难题,就是在对进入同一个状态的两条路径进行比较时,

总有可能出现两条路径的距离度量相等的情况。如果遇到这种情况,就只能通过类似于抛硬币的简单猜测来决定了。

总之,维特比算法是一种最大似然估计的译码器算法,这种算法对于 AWGN 信道来说是最优的选择,维特比算法的步骤如下:

初始化

将网格图的最左边的状态(第 0 级的全零状态)标为 0,因为这一点的计算结果不存在差异。

计算 $j+1$ 步

取 $j=0,1,2,\dots$,并确保在第 j 步时已经完成了以下两个步骤:

- 对所有的留存路径进行鉴别。
- 保存每个状态的留存路径及其距离度量。

然后,在 $j+1$ 级(时钟时间),将输入支路的距离度量加到第 j 级对应的留存路径的距离度量上,从而计算得到网格图上进入每个状态的所有路径的距离度量。然后,对于每个状态,鉴别出具有最小距离度量的路径,作为第 $j+1$ 步的留存路径,并更新计算结果。

最后一步

继续计算,直到完成对整个网格图的前向搜索并到达最终节点(全零状态),就确定了一条最大似然路径。与分组码的译码器类似,最大似然路径对应的符号序列作为接收序列的译码输出被送往信宿。从这个意义上说,将维特比算法视为最大似然序列估计器更为准确。

但是,当接收序列很长(接近无穷)时,维特比算法对存储器的要求就变得非常高,因此必须采取某些折中手段。通常采用的方法是将译码器的路径存储器进行截断处理。先确定一个长度为 ℓ 的译码窗口,并对接收序列中相应的帧进行维特比算法的计算,计算在 ℓ 步后停止。判断出“最好”的路径并将这条路径的第一个分支上的相应符号发给用户。这条路径上最后一个分支对应的符号将被丢弃。接下来,译码窗口向前移动一个时隙,再对下一帧码进行同样的判断,以此类推。按这种方法进行的译码判决,已不再是真正意义上的最大似然估计,但如果译码窗口足够长,这种判决的性能与最大似然估计是一样的。实验和分析都证明:只要译码窗口的长度 ℓ 达到卷积码约束长度的 5 倍或更高,就可以获得令人满意的结果。

例 10.6 接收序列为全零时的正确译码

设图 10.13(a)所示的编码器产生了一个全零序列并送入二进制对称信道,接收序列为 0100010000…。由于信道噪声的原因,接收序列中有两个误码:一个是第 2 比特,另一个是第 6 比特。下面用维特比译码算法来纠正这个双误码图样。

图 10.17 中给出了 $j=1,2,3,4,5$ 时维特比算法的计算结果。从图中可以看出,在 $j=2$ 处有 4 条路径,每条对应于编码器 4 个状态中的一个。图中还给出了各路径在每一级计算得到的距离度量。

在图 10.17 的左边,对于 $j=3$,给出了进入每个状态的路径及各自对应的距离度量。在图的右边,给出了 $j=3,4,5$ 时由维特比算法得到的 4 条留存路径。

观察图 10.17 中 $j=5$ 的 4 条留存路径,可发现全零路径的距离度量最小。于是,从这点开始,向前保留这条距离度量最小的路径。这显然表明全零序列是维特比算法的最大似然估计结果,这恰好与发送序列完全一致。

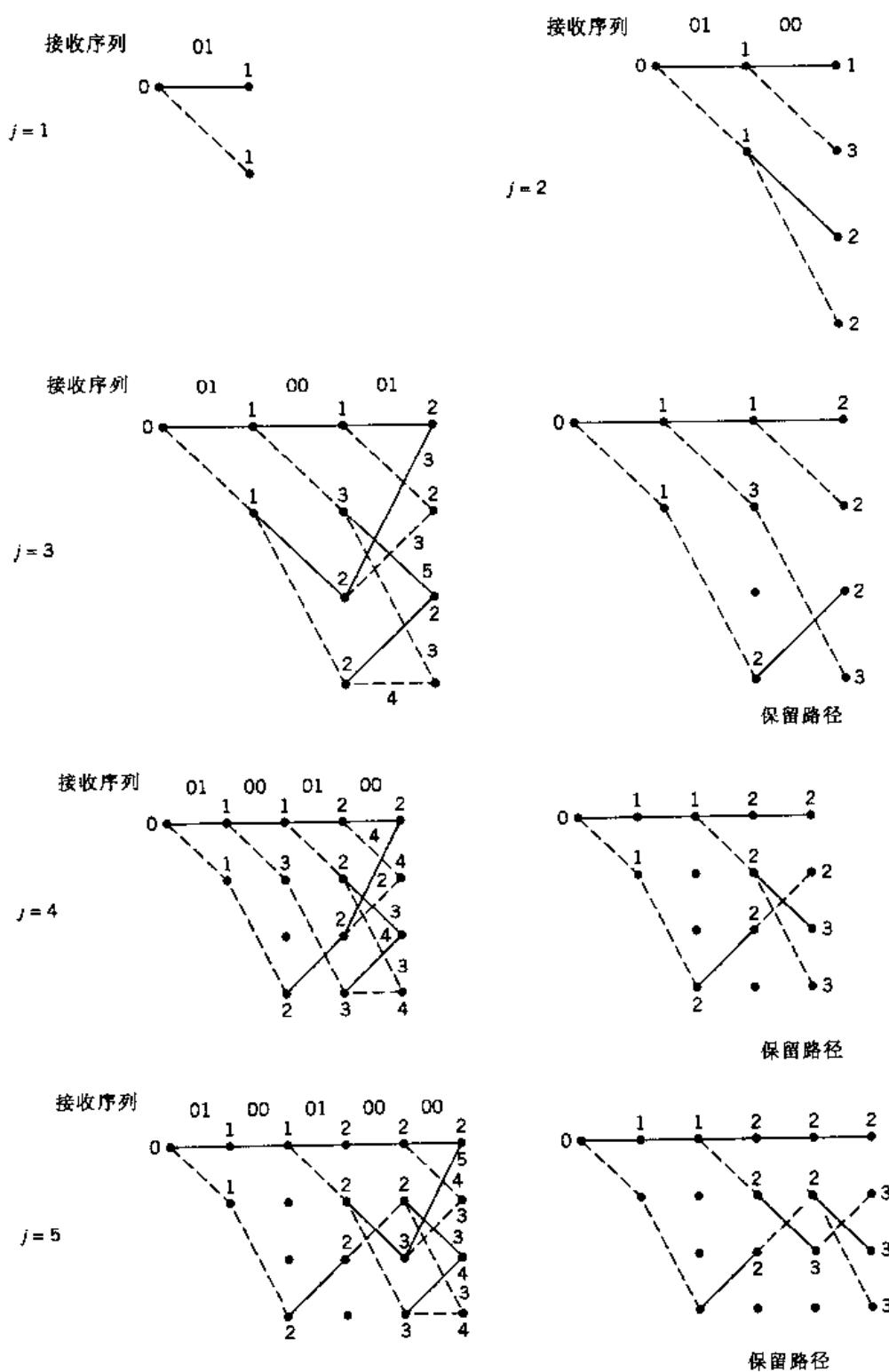


图 10.17 例 10.6 中维特比算法的分步说明

例 10.7 接收序列为全零时的错误译码

在例 10.6 中, 设接收序列为 1100010000…, 与发送的全零序列相比, 接收序列包含三个误码。

图 10.18 给出了 $j = 1, 2, 3, 4$ 时维特比译码的结果。不难发现, 正确的译码路径在 $j = 3$ 被破坏。显然, 当卷积码的码率为 $1/2$, 约束长度 $K = 3$ 时, 用维特比译码是不能纠正有三个错误的图样的。但是, 这个规则有一个例外, 就是当三个错误的分散间隔大于约束长度时, 维特比算法就很有可能正确译码。

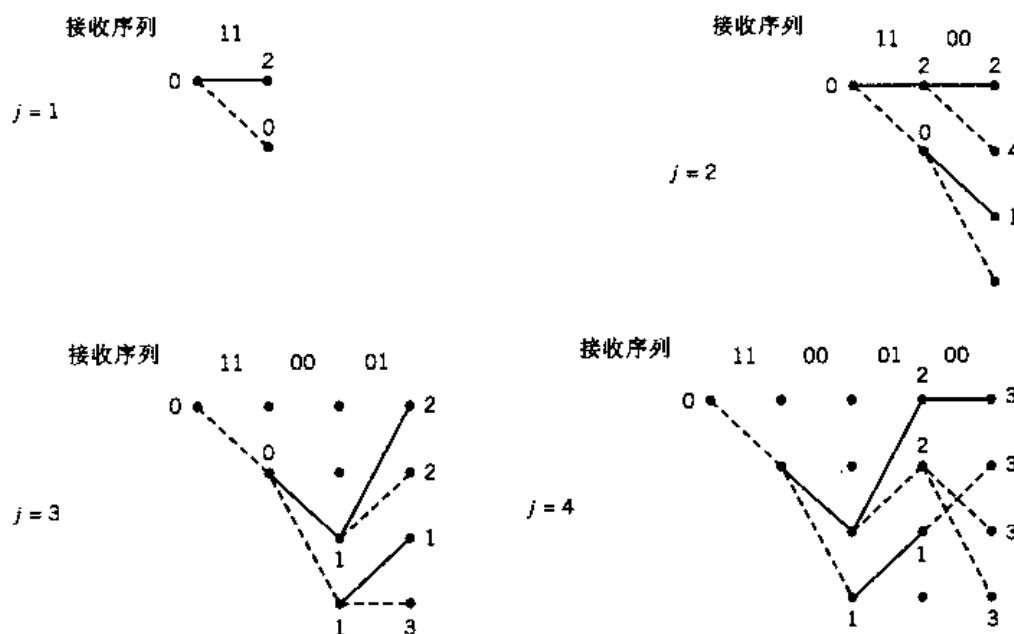


图 10.18 例 10.7 中维特比算法的分步说明

10.6.2 卷积码的自由距离

卷积码的性能不仅取决于所用的译码算法,而且与码自身的距离特性有关。衡量一个卷积码抗信道噪声能力的最重要的指标是自由距离,记作 d_{free} 。卷积码的自由距离被定义为该码中任意两个码字间的最小汉明距离。当且仅当 $d_{\text{free}} > 2t$ 时,卷积码才能够纠正 t 个误码。

自由距离可以很容易从图 10.13(a)所示的卷积码编码器的状态图中求得。任何非零码序列都分别对应于一条完整的路径,这些路径均起始于 00 状态(节点 a)又终止于 00 状态。将这些路径上的节点按图 10.19 所示的方式分开,是很有意义的。该图可以被看成是一个信号流图,其中只有一个输入和一个输出。一个信号流图由节点和带方向的分支组成。其操作规则如下:

1. 在每个分支上,都将输入节点的信号乘上表征该分支的传递系数。
2. 一个有输入分支的节点汇集了由所有分支产生的信号。
3. 一个节点的信号将均等地分配给所有由这个节点输出的分支。
4. 图中的转移函数是输出信号与输入信号的比值。

再回到图 10.19 所示的信号流图,图中分支上 D 的指数表示对应于该分支的编码器输出的汉明重量。 L 的指数总是等于 1,这是因为每个分支的长度均为 1。用 $T(D, L)$ 来表示信号流图的转移函数,其中, D 和 L 都是虚拟变量。对于图 10.19 的例子,可利用规则 1,2 和 3 很容易地得到如下的输入 - 输出关系:

$$\left. \begin{array}{l} b = D^2 L a_0 + L c \\ c = D L b + D L d \\ d = D L b + D L d \\ a_1 = D^2 L c \end{array} \right\} \quad (10.62)$$

其中, a_0 , b , c , d 和 a_1 分别表示图中的节点信号。求解(10.62)的方程组, 可得到 a_1/a_0 的比值, 于是图 10.19 所示图形的转移函数为

$$T(D, L) = \frac{D^5 L^3}{1 - DL(1 + L)} \quad (10.63)$$

可将上式按二项式展开为

$$T(D, L) = D^5 L^3 \sum_{i=0}^{\infty} (DL(1 + L))^i \quad (10.64)$$

在式(10.64)中令 $L = 1$, 可得幕级数形式表示的距离转移函数

$$T(D, 1) = D^5 + 2D^6 + 4D^7 + \dots \quad (10.65)$$

由于自由距离是卷积码中任意两个码字间的最小距离, 而距离转移函数 $T(D, 1)$ 分别列举了给定距离的码字数, 因此 $T(D, 1)$ 中第一项的指数就被定义为自由距离。因此, 根据式(10.65), 可以得出图 10.13(a)所示卷积码的自由距离 $d_{\text{free}} = 5$ 。

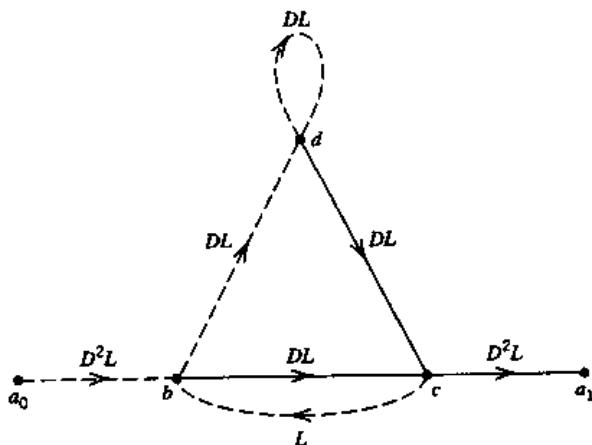


图 10.19 改进的卷积码状态变换图

这个结果表明最多可纠正接收序列中的两个误码, 这是因为两个或两个以下的误码使接收序列与发送序列之间的汉明距离最多为 2, 而与其他可能的发送序列之间的汉明距离至少为 3。换言之, 尽管发送过程中出现了任意的两个误码, 但是与其他可能的发送序列相比, 接收序列仍然最接近于实际发送序列。但是, 如果接收序列中出现 3 个或更多紧靠的传输误码, 这个结论就不再成立了。这个结论进一步证实了前面例 10.6 和例 10.7 的结果。

在使用距离转移函数 $T(D, 1)$ 来计算卷积码的自由距离时, 假定用于表示 $T(D, 1)$ 的单位时延变量 D 的幕级数是收敛的(即和为有限值)。对于图 10.13(a)所示的卷积码, 这个假设要求证明式(10.65)的展开式是有效的。但是, 实际上并不能保证 $T(D, 1)$ 总是收敛的。当 $T(D, 1)$ 不收敛时, 有限的传输误码会导致无限的译码误码。卷积码就很容易导致这种灾难性

的误码传播,这样的卷积码称为恶性码^[10]。值得注意的是,系统的卷积码不会是灾难性的。遗憾的是,对于给定的约束长度 K 来说,用图 10.13 所示方案得到的系统卷积码的自由距离通常小于非系统卷积码,如表 10.8 所示。

表 10.8 码率为 1/2 的系统卷积码和非系统卷积码的最大自由距离

约束长度 K	系 统 码	非 系 统 码
2	3	3
3	4	5
4	4	6
5	5	7
6	6	8
7	6	10
8	7	10

10.6.3 渐近编码增益^[11]

对一个给定的译码方案,编码器状态图的转移函数采用类似于图 10.19 所示的方式进行改进,就可以用来估算误比特率的边界。具体的估算方法已超出了本书讨论的范围。这里仅概括两种特殊信道对应的结果,一个是二进制对称信道,另一个是具有二进制输入的 AWGN 信道。设采用的相干检测方法为二进制 PSK。

1. **二进制对称信道。**二进制对称信道可以看成是以二进制 PSK 作为调制方式,以硬判决进行解调的 AWGN 信道。对于非编码的二进制 PSK 系统来讲,二进制对称信道的转移概率等于其误比特率(BER)。回顾第 6 章的知识,对于大的 E_b/N_0 值,即信号能量与每比特噪声功率谱密度的比值,无编码的二进制 PSK 的误比特率由指数因子 $\exp(-E_b/N_0)$ 决定。另一方面,采用同一调制方案的卷积编码的误比特率由指数因子 $\exp(-d_{\text{free}} r E_b / 2 N_0)$ 决定,其中, r 为码率, d_{free} 是卷积码的自由距离。作为衡量采用硬判决译码的编码方式带来的误码性能改进的指标,可以用指数来定义渐近编码增益(以 dB 表示):

$$G_a = 10 \log_{10} \left(\frac{d_{\text{free}} r}{2} \right) \text{ dB} \quad (10.66)$$

2. **二进制输入 AWGN 信道。**假设一个无输出量化,即输出的幅度在 $(-\infty, \infty)$ 之间的无记忆二进制输入 AWGN 信道。对于这种信道,理论显示,对于大的 E_b/N_0 值而言,卷积码的二进制 PSK 的误比特率由指数因子 $\exp(-d_{\text{free}} r E_b / N_0)$ 决定,其参数如前所述。在这种情况下,可定义渐近编码增益为:

$$G_a = 10 \log_{10}(d_{\text{free}} r) \text{ dB} \quad (10.67)$$

由式(10.66)和式(10.67)可以看出,具有二进制输入的 AWGN 信道的渐近编码增益要比二进制对称信道大 3 dB。换言之,如果要获得相同的误码性能,对于大的 E_b/N_0 值而言,二进制对称信道的发射机必须发送比二进制输入 AWGN 信道的发射机大 3 dB 的信号能量(或功

率)。显然,无量化的解调器输出比硬判决更具有性能上的优势。但是,这是以增加译码器的复杂度为代价的,因为其中要求接收机接收模拟的输入。

在二进制输入 Q 进制输出的离散无记忆信道的情况下,二进制输入的 AWGN 信道的渐近编码增益近似等于 0.25 dB ,其中, $Q = 8$ 。也就是说,可用有限量化输出的软判决译码器(典型的情况是 $Q = 8$)代替模拟译码器,同时还能使性能接近最佳值。

10.7 网格编码调制^[12]

在本章讨论的各种常用的信道编码方法中,发射机中的编码过程与调制过程是分开进行的,接收机中的译码与解调也是如此。而差错控制是通过在发送的码字中加入冗余比特来实现的,这将降低每信道带宽传输的信息比特率。也就是说,功率利用率的提高是以牺牲带宽利用率为代价的。

为了更有效地利用带宽和功率,必须将编码和调制作为一个整体来对待。在这种情况下,就必须重新定义编码为在发射信号上加上某种信号图样的过程。实际上,这个定义包含了传统的奇偶校验编码的思想。

带限信道的网格编码就是将调制和编码作为一个整体,而不是两个分开部分的结果。这种调制和编码的组合就是网格编码调制。这种信号形式具有三个基本特征:

1. 在星座图中使用信号点的数目要大于相同数据速率下所用调制格式所需的信号点数,多余的信号点可以作为不牺牲带宽条件下前向纠错编码的冗余。
2. 卷积码被用来在连续的信号点之间引入某种依赖关系,因而,只有某些信号点的图样或序列才允许采用卷积码。
3. 在接收机中采用软判决译码,这样就可将那些允许的信号序列建成网格结构的模型,因此,这种编码方式称为“网格编码”。

这是使用扩大的星座图的结果。随着星座图的增大,在固定信噪比的情况下,符号差错概率也会增加。因此,如果采用硬判决译码,在译码之前就要面临信息的损失。在结合了编码和调制的网格编码调制上采用软判决译码,就可以改善这种状况。

在 AWGN 信道中,网格编码调制的最大似然估计译码,就是寻找一条与接收序列之间具有最小欧氏平方距离的特殊路径。因此,在设计网格编码调制时,重点应是使两个码矢量(等效于码字)的欧氏距离最大,而不是使纠错码的汉明距离最大。这样做的原因是,除了传统的二进制 PSK 和 QPSK 外,汉明距离的最大化不等同于欧氏平方距离的最大化。下面将阐述的欧氏距离才是我们感兴趣的度量。此外,我们的讨论仅局限于信号点的二维星座图的情况。这样选择的含义就是将网格编码调制限制于多级 AM 或 PM,如 M 进制的 PSK 和 M 进制的 QAM。

用于设计这类网格编码调制的方法,就是将一个 M 进制星座图连续划分为 $2, 4, 8, \dots$ 个大小为 $M/2, M/4, M/8, \dots$ 的子集,此时,各个信号点之间的最小欧氏距离逐渐增大。这种设计方法代表了为带宽受限信道构造有效编码调制技术的关键思想。

下面,我们借助于图 10.20 所示 8-PSK 的圆形星座图来说明划分的过程。图中描述了星座图本身和经过两级划分后得到的 2 个和 4 个子集。这些子集的共同特性就是其中各个点之

间的最小欧氏距离按如下的规律增加: $d_0 < d_1 < d_2$ 。

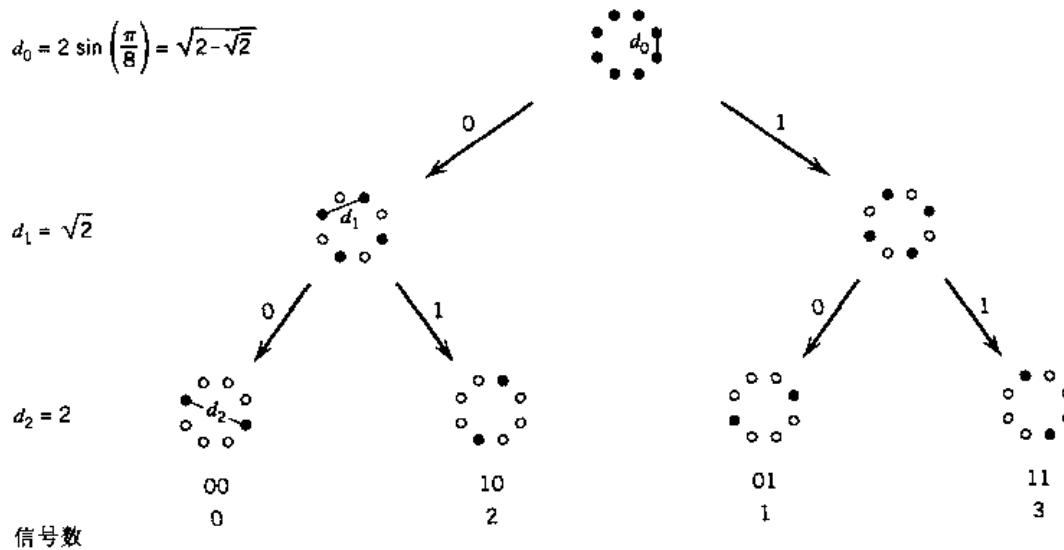


图 10.20 8-PSK 星座图划分, $d_0 < d_1 < d_2$

图 10.21 给出了 16-QAM 对应的方形星座图的划分情况。从这个图同样可以看出, 各子集内部的欧氏距离也是逐渐增加的: $d_0 < d_1 < d_2 < d_3$ 。

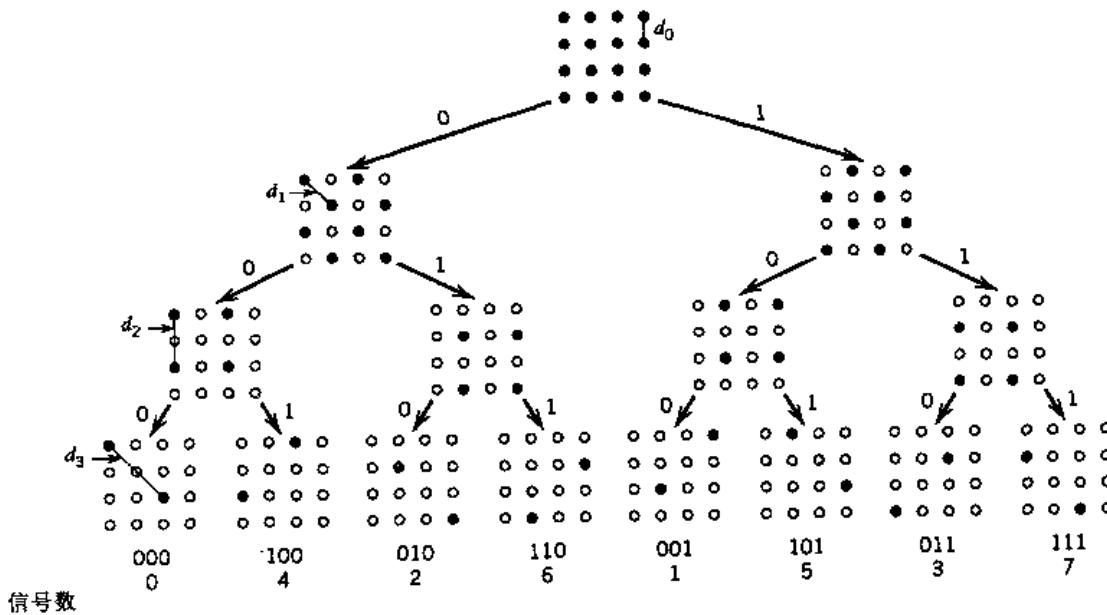


图 10.21 16-QAM 星座图划分, $d_0 < d_1 < d_2 < d_3$

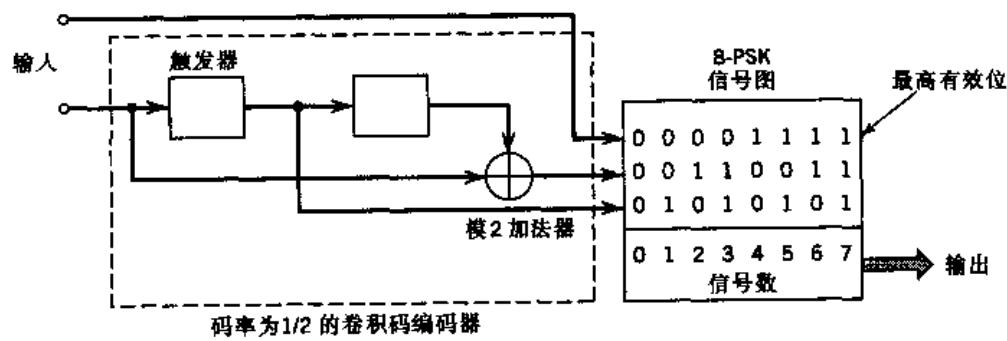
在二维星座图连续划分得到的子集基础上, 可以设计出相对简单但高效的编码方案。特别地, 为了发送正交调制(同步且正交)的 n 比特/符号的信号, 可从有 2^{n+1} 个信号点的二维星座图开始, 将这些信号点用于采用的调制方式。圆形栅格用于 M 进制 PSK, 方形栅格则用于 M 进制的 QAM。星座图被划分为 4(或 8) 个子集。输入的每个符号中的一个(或两个)比特分别被送入码率为 $1/2$ 或 $2/3$ 的二进制卷积码编码器。根据每个符号编码后得到的两个(或 3 个)比特来决定选择某个特定的子集。根据剩下的未编码数据比特来决定将选中的子集中哪个

个信号点作为发射信号。这一类网格编码称为翁格博克(Ungerboeck)码。

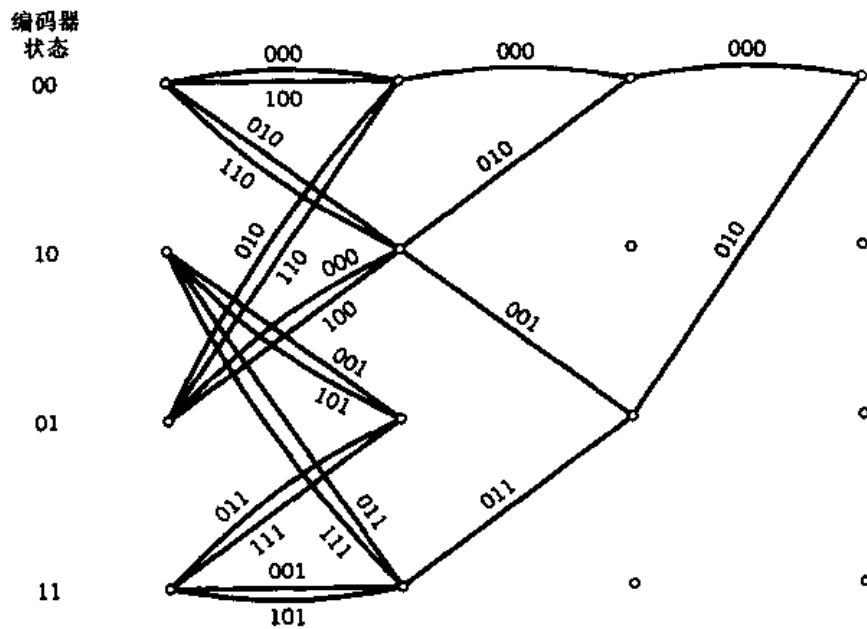
由于调制器中有存储器,因此在接收机中可用维特比算法来进行最大似然序列估计。翁格博克码的网格图中,每一个分支对应于一个子集,而不是一个单独的信号点。检测的第一步,就是确定每个子集中在欧氏距离的意义上最接近接收信号点的信号点。信号点确定后,确定了分支的距离度量(信号点与接收信号点之间的欧氏平方距离),维特比算法才可以正常进行。

10.7.1 8-PSK 的翁格博克码

图 10.22(a)所示的方案是最简单的传输 2 比特/符号的翁格博克 8-PSK 码。该方案中采用了码率为 $1/2$ 的卷积码。相应的网格图如图 10.22(b)所示,其中包括 4 个状态。需要注意的是,输入的二进制码字中最重要的比特未被编码。因此,网格图中每个分支可对应于 8-PSK 调制器中两个不同的输出,或者等效地对应于图 10.20 中 4 个两点子集中的一个。图 10.22(b)所示的网格图还包括了最小距离路径。



(a) 四状态的 8-PSK 翁格博克码编码图;其中信号图对应于图 10.20



(b) (a) 对应的的网格图

图 10.22 四状态的 8-PSK 翁格博克码的编码图和网格图

图 10.23(a)为另一种传输 2 比特/抽样的翁格博克 8-PSK 码方案。这种方案比前一种方案复杂一级。这个方案采用的是码率为 $2/3$ 的卷积码。因此,码的网格图有 8 种状态,如图 10.23(b)所示。在这种情况下,对输入的两个二进制比特都进行了编码。因此,网格图的每个分支都对应于 8-PSK 调制器的一个特定输出值。图 10.23(b)中包含了最小距离路径。

图 10.22(b)和图 10.23(b)中都包含了编码器的状态。在图 10.22 中,编码器的状态被定义为两级移位寄存器的内容。而在图 10.23 中,则被定义为单级(上层)移位寄存器的内容加上两级(下层)移位寄存器的内容。

10.7.2 渐近编码增益

根据 10.6 节的讨论,可定义翁格博克码的渐近编码增益为

$$G_a = 10 \log_{10} \left(\frac{d_{\text{free}}^2}{d_{\text{ref}}^2} \right) \quad (10.68)$$

其中, d_{free} 是该码的自由欧氏距离, d_{ref} 是具有相同比特信号能量的未编码调制方案的最小欧氏距离。例如,在图 10.22(a)所示的 8-PSK 翁格博克码的应用中,信号的星座图包含有 8 个信号点,每个信号点上可发送两个信息比特。因此,在未编码的情况下,需要 4 个信息点的信号星座图。可将未编码的 4-PSK 作为图 10.22(a)所示的 8-PSK 翁格博克码的参考。

图 10.22(a)中的 8-PSK 翁格博克码可获得 3 dB 的渐近编码增益,计算方法如下:

1. 图 10.22(b)中网格图的每个分支对应于有两个对应信号点的子集。因此,自由欧氏距离 d_{free} 不可能大于这个子集中两个对应信号点之间的欧氏距离 d_2 。于是有

$$d_{\text{free}} = d_2 = 2$$

d_2 的定义如图 10.24(a)所示,也可以见图 10.20。

2. 如图 10.24(b),作为参考的具有相同每比特信号能量的未编码 QPSK 的最小欧氏距离等于

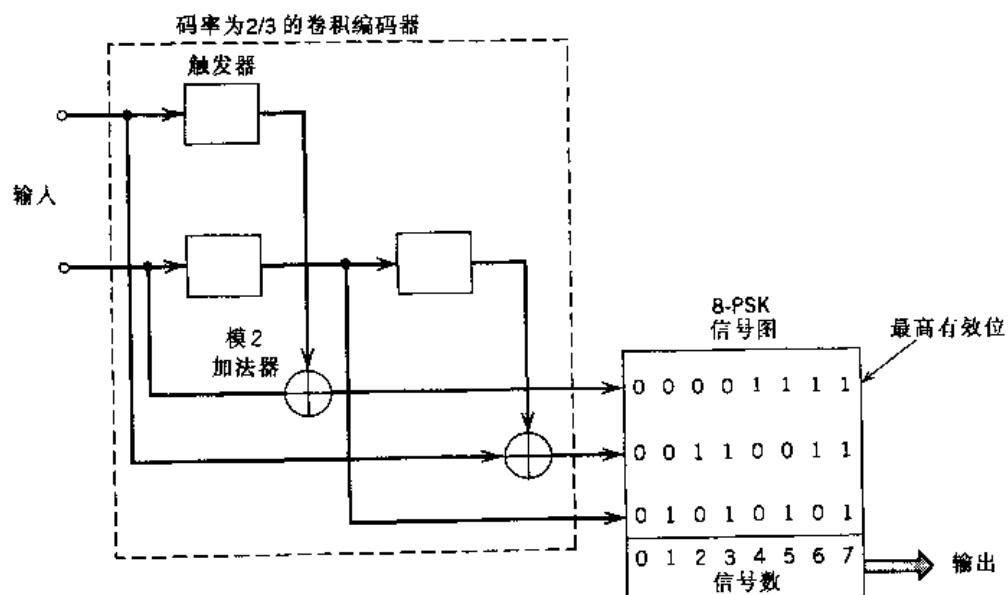
$$d_{\text{ref}} = \sqrt{2}$$

如前所述,由式(10.68)可得渐近编码增益为 $10 \log_{10} 2 = 3$ dB。

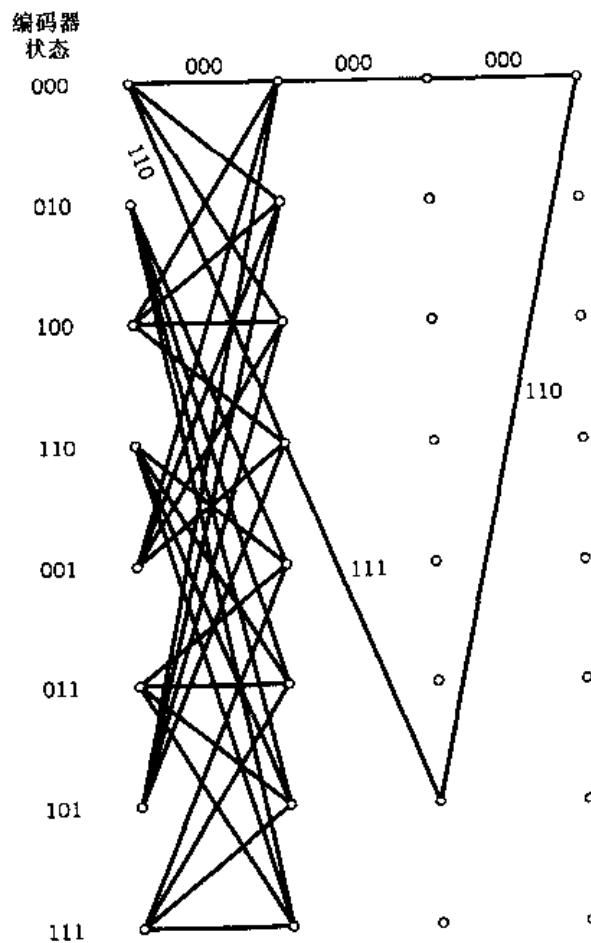
随着卷积码状态数的增加,翁格博克码的渐近编码增益也会增加。表 10.9 给出了 8-PSK 翁格博克码的渐近编码增益(dB)随状态数增加的变化情况,表中以未编码的 4-PSK 为参考。需注意的是,要获得 6 dB 的改善,就需要有很多的状态数。

表 10.9 相对于 4-PSK 的 8-PSK 翁格博克码的渐近编码增益

状态数	4	8	16	32	64	128	256	512
编码增益(dB)	3	3.6	4.1	4.6	4.8	5	5.4	5.7



(a) 八状态的 8-PSK 翁格博克码编码图;其中信号图对应于图 10.20



(b) 八状态的 8-PSK 翁格博克码网格图的部分分支

图 10.23 八状态的 8-PSK 翁格博克码的编码图和部分网格图分类

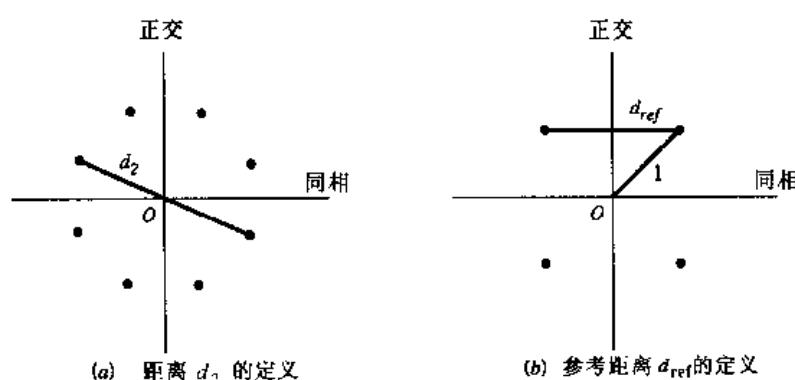


图 10.24 计算 8-PSK 翁格博克码渐近编码增益的信号空间

10.8 Turbo 码^[13]

通常,一个好码的设计是通过构造大量代数结构的码来解决的,这些码都对应于可行的译码方案。这一点在前面的线性分组码和卷积码的讨论中都得到了证实。但是,这些传统的编码方法都存在一个困难,即为了尽量接近香农信道容量的理论极限,需要增加线性分组码字的长度或卷积码的约束长度,这实际上会使最大似然估计译码器的计算复杂度以指数增加,最终会使译码器的复杂度达到无法实现的地步。

因此,人们提出了各种各样的构造具有大“等效”分组长度的有效编码方法。基本思想都是将译码过程分为许多较容易实现的步骤来完成。在这些方法的基础上,Turbo 码和低密度奇偶校验码的发展取得了最大的成功,从而开辟了一条令人振奋的构造好码并以合理的复杂度进行译码的新途径。本节将讨论 Turbo 码,而低密度奇偶校验码将在 10.10 节进行讨论。

10.8.1 Turbo 码的编码

如图 10.25 所示,Turbo 码最基本的形式就是由两个用交织器连接在一起的系统码编码器组成的。

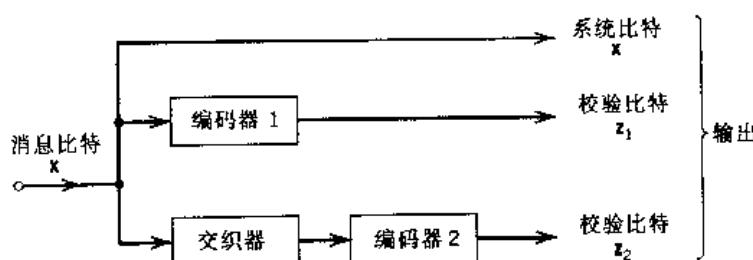


图 10.25 Turbo 码编码器框图

交织器是一种输入输出变换装置,这种装置能够根据一个固定的字符集,以一种完全确定的方式来改变字符序列的顺序。也就是说,根据输入端的符号,在输出端产生相同但排列顺序不同的字符。交织器可以有很多种类型,其中最主要的两种就是周期交织器和伪随机交织器。Turbo 码采用的是伪随机交织器,该交织器只作用于系统比特。在 Turbo 码中,采用交织器的目的有两个:

- 将一半 Turbo 码中容易发生的误码与另一半中特别不容易发生的误码联系起来。实际上,这也是 Turbo 码的性能优于传统码的主要原因。
- 提供充分的关于不匹配译码的性能参数,这是在信道统计量未知或被错误指定的情况下出现的一个问题。

一般的做法是,将相同的码输入图 10.25 的两个编码器,但并非必须这样。Turbo 码推荐使用的组成码是短的限长递归系统卷积码(RSC 码)。使卷积码递归(将一个或多个抽头的输出反馈到移位寄存器的输入端)的目的是为了使移位寄存器的内部状态与其过去的输出有关,这将影响错误图样的状态(即系统码的单个误码会产生无穷多个校验误码),从而获得更好的整体编码策略以提高系统性能。

例 10.8 八状态的 RSC 码

图 10.26 给出了一个八状态 RSC 码编码器的例子。这个递归卷积码的生成多项式为

$$g(D) = \left[1, \frac{1 + D + D^2 + D^3}{1 + D + D^3} \right] \quad (10.69)$$

D 是时延变量。矩阵 $g(D)$ 的第二项是反馈移位寄存器的转移函数,其定义为输出变换式除以输入变换式。令 $M(D)$ 为信息序列 $\{m_i\}_{i=1}^k$ 的变换式, $B(D)$ 为奇偶校验序列 $\{b_i\}_{i=1}^{n-k}$ 的变换式。根据定义可得

$$\frac{B(D)}{M(D)} = \frac{1 + D + D^2 + D^3}{1 + D + D^3}$$

将上式进行交叉相乘,有

$$(1 + D + D^2 + D^3)M(D) = (1 + D + D^3)B(D)$$

将其反变换到时域,得到

$$m_i + m_{i-1} + m_{i-2} + m_{i-3} + b_i + b_{i-1} + b_{i-3} = 0 \quad (10.70)$$

其中,加法是模 2 的。式(10.70)是奇偶校验等式,对于所有 i ,图 10.26 所示卷积码编码器都满足此式。

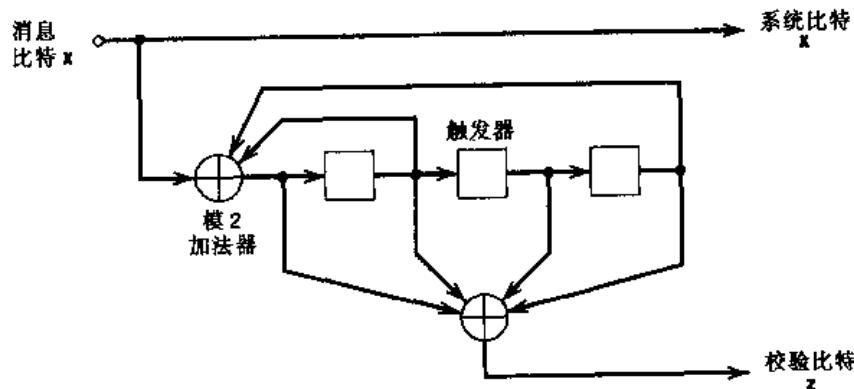


图 10.26 八状态 RSC 码编码器

图 10.25 中,将数据流直接输入到编码器 1,而将该数据流经过伪随机重新排序后的结果输入到编码器 2。系统比特(源信息比特)和由两个编码器生成的两组奇偶校验比特组成了

Turbo 编码器的输出。虽然组成码是卷积的,但 Turbo 码实际上是分组码,且分组长度决定于交织器的长度。进一步,由于图 10.25 中的两个 RSC 码编码器是线性的,因此可认为 Turbo 码是线性分组码。

Turbo 码的分组特性引出了一个现实的问题:应该怎样来确定码字的开头与结尾?通常采用的方法是,先将编码器初始化为全零状态,然后对数据进行编码。在完成对一定数目的数据比特的编码后,再加上一些尾比特,使编码器在每个分组的结尾返回到全零状态,此后再循环往复。Turbo 码的结束方法如下:

- 一种简单的方法是,关闭编码器中的第一个 RSC 码,而留下第二个不关闭。这种方法的缺点是,由第二个 RSC 码生成的分组的最后比特与其他比特相比,更容易受到噪声的影响。实验结果显示:在信噪比提高时,Turbo 码表现出相对平稳的性能。Turbo 码的这种特性并不是误码底限,而只是在低信噪比的情况下,当误码性能急剧下降时所表现出来的类似于误码底限的性质。这种“误码底限”受多种因素影响,其中最主要的是交织器的选择。
- 一种更精确的方法^[14]是同时对称地终止编码器中的两个组成码。通过结合采用好的交织器和双重终止,可使“误码底限”相对于第一种方法(即简单终止)有一个数量级的提高。

在 Turbo 码的最初模型中,为了保持 1/2 的码率,在数据经信道传输之前,需要对图 10.25 所示的两个编码器产生的奇偶校验比特进行压缩。删除掉某些奇偶校验比特可以得到收缩的码,从而提高数据速率。码字的收缩与扩展相反。需要强调的是,对 Turbo 码的生成而言,压缩并不是必须的。

图 10.25 所示并行编码方案的新颖之处在于,其中采用了递归的系统卷积码,并在两个编码器之间引入了伪随机的交织器。因此,由于伪随机交织器的优点,Turbo 码(对信道)实质上表现出随机的特性,同时,其结构又使得译码方案切实可行。根据编码理论可知,如果分组足够大,随机选取的码可接近香农的信道容量。^[15]这是 Turbo 码具有优越性能的真正原因。下面就讨论 Turbo 码的性能。

10.8.2 Turbo 码的性能

图 10.27 给出了一个码率为 1/2 且具有较大分组长度的 Turbo 码,以二进制数据形式经 AWGN 信道传输时的误码性能。^[16]其中采用了一个大小为 65 536 的交织器和一个基于 BCJR 的译码器(该译码器的详细情况将在后面的章节中给出)。在计算过程中,进行了 18 次 Turbo 译码迭代。

为了便于比较,图 10.27 还给出了相同 AWGN 信道条件下的另外两条曲线:

- 未编码数据传输(码率 $r = 1$)。
- 码率为 1/2 时的香农理论极限,参见图 9.18(b)。

由图 10.27 可得如下两个重要结论:

1. 虽然在低 E_b/N_0 时,采用 Turbo 码进行传输时的误比特率明显高于未编码的数据传输,但当 E_b/N_0 达到某一临界值时,Turbo 码的误比特率就会迅速下降。
2. 当误比特率达到 10^{-5} 时,Turbo 码对应的误比特率仅比香农理论极限小 0.5 dB。

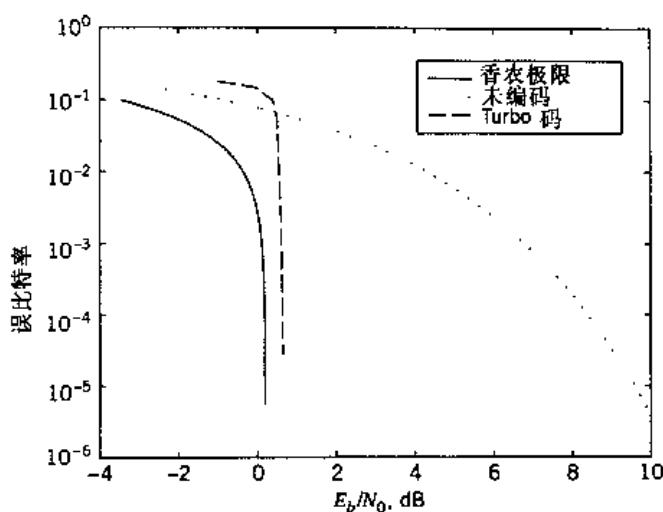


图 10.27 码率为 1/2 的 Turbo 码和未编码数据在 AWGN 信道中传输时的误码性能, 以及码率为 1/2 的码的信道容量的香农理论极限

值得注意的是, 要获得如此高的性能改善, 交织器的大小或 Turbo 码的分组长度必须足够大。此外, 改善性能所需的大量迭代也会增加译码器的等待时间。造成这些缺点的主要原因是信息的数字处理没有对反馈提供帮助, 这也是 Turbo 码译码器的一个与众不同的特征。

现在我们已对 Turbo 码的性能有了一定的了解, 下面将讨论 Turbo 码的译码是如何进行的。

10.8.3 Turbo 码的译码

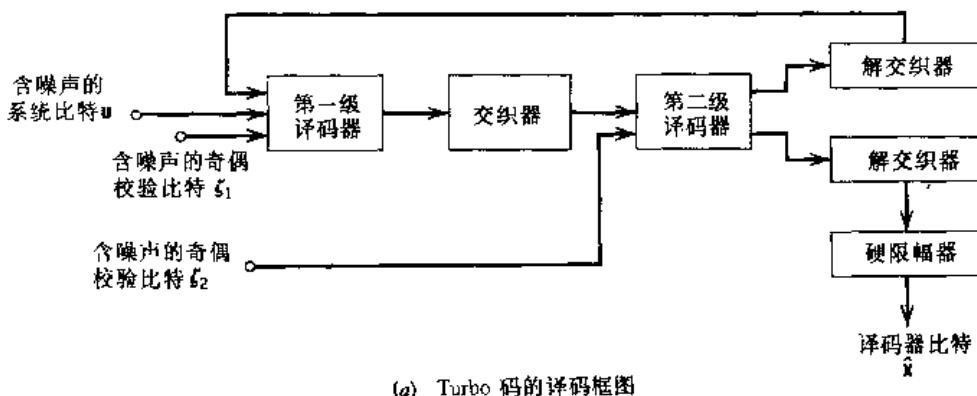
Turbo 码名称的由来是因为其译码算法类似于“涡轮发动机”的原理。图 10.28(a)显示了 Turbo 译码器的基本结构。其中通过对系统比特噪声模型和两个译码段中两组奇偶校验比特进行运算, 产生原始信息比特的估计值。

两个译码段中均用到了 BCJR 算法^[17], 该算法最初是由 Bahl, Cocke, Jelinek 以及 Raviv (以此命名)发明的, 用来解决最大后验概率检测的问题。BCJR 算法与维特比算法的不同表现在两个基本方面:

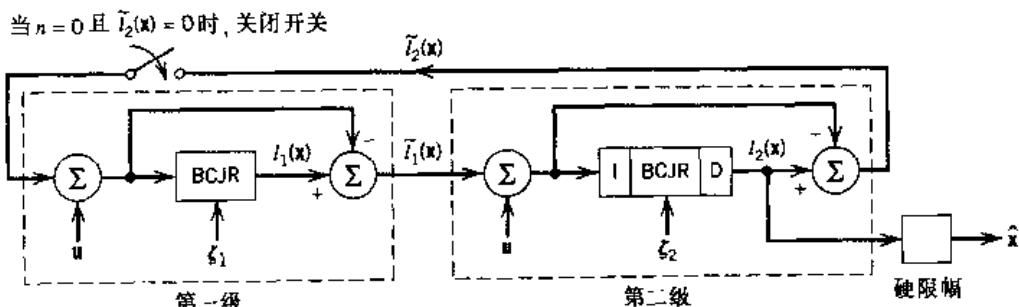
1. BCJR 算法是软输入软输出的译码算法, 需要经过两次递归, 一次向前一次向后, 且两次都是软判决。相反, 维特比算法是软输入硬输出的译码算法, 只有一次包含软判决的前向递归, 在递归结束后采用硬判决, 并以此在多条路径中选定一条特定的路径作为留存路径。因此, 在计算中, BCJR 算法由于要进行后向的递归, 所以要比维特比算法复杂。
2. BCJR 算法是一种 MAP 译码器, 因为它通过估算码字中各比特的最大后验概率来使误比特率达到最小。为了重构原始的数据序列, BCJR 算法的软输出采用硬限幅。另一方面, 维特比算法是一种最大似然序列估计器, 因为其中要使整个码序列的似然函数, 而不是每个比特达到最大。因此, BCJR 算法的平均误比特率可略好于维特比算法, 但决不会更差。

重要的是, BCJR 算法的公式是建立在两个基本假设之上的:(1)发射机采用卷积编码作为信道编码(其数学模型为马尔可夫过程);(2)信道是无记忆的。在目前的讨论中, 马尔可夫假设是

指:如果码可以用网格图来表示,那么,网格图当前的状态取决于其过去的状态和输入比特。在下面的章节里我们将给出 BCJR 算法的数学解释。



(a) Turbo 码的译码框图



(b) Turbo译码器的外部形式。I 表示交织器, D 表示解交织器, BCJR 表示对数最大后验概率译码的 BCJR 算法

图 10.28 Turbo 码的译码

在描述图 10.28(a)所示的两段 Turbo 译码器的执行情况之前,介绍一下外部信息的概念是有必要的。这个概念最方便的表示法是对数似然比值。此时,外部信息就可以用如图 10.29 所示的两个对数似然比的差来计算。从形式上讲,(由系统信息比特译码段产生的)外部信息的定义为(在该译码段的)输出端计算得到的对数似然比与(由反馈到这个译码段输入端的对数似然比表示的)内部信息的差值。实际上,外部信息是通过挖掘所关心的信息比特与译码器产生的原始输入数据比特之间的相关性而得到的增量信息。

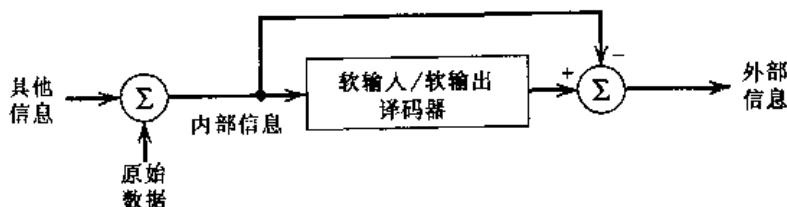


图 10.29 外部信息的概念说明

在这个基础上,就可以用图 10.28(b)所示的对称外部方式来描述图 10.28(a)中两段 Turbo 译码器中的信息流了。第一个译码段运用 BCJR 算法生成系统比特 x_i 的软估计值,该估计值可表示为对数似然比

$$l_1(x_i) = \log \left(\frac{P(x_i = 1 | u, \zeta_1, \tilde{l}_2(x))}{P(x_i = 0 | u, \zeta_1, \tilde{l}_2(x))} \right), \quad i = 1, 2, \dots, k \quad (10.71)$$

其中, \mathbf{u} 是含噪声的系统比特组, ζ_1 是由编码器 1 生成的含噪声的奇偶校验比特组, $\tilde{l}_2(\mathbf{x})$ 是(第二译码段产生并反馈到第一译码段的)信息比特组对应的外部信息。设 k 个信息比特是统计独立的, 则在第一译码段输出端的总的对数似然比为

$$l_1(\mathbf{x}) = \sum_{j=1}^k l_1(x_j) \quad (10.72)$$

因此, 第一译码段生成的信息比特对应的外部信息是

$$\tilde{l}_1(\mathbf{x}) = l_1(\mathbf{x}) - \tilde{l}_2(\mathbf{x}) \quad (10.73)$$

其中, $\tilde{l}_2(\mathbf{x})$ 还未进行定义。

在应用第二译码段之前, 外部信息 $\tilde{l}_1(\mathbf{x})$ 被重新排序, 以补偿在 Turbo 编码器中引入的伪随机交织。另外, 由编码器 2 产生的含有噪声的奇偶校验比特 ζ_2 被用来作为输入。这样, 根据 BCJR 算法, 第二译码段就能够产生更精确的信息比特 \mathbf{x} 的软估计值。将这个估计值重新交织, 得到总的对数似然比 $l_2(\mathbf{x})$ 。因此, 反馈到第一译码段的外部信息 $\tilde{l}_2(\mathbf{x})$ 是

$$\tilde{l}_2(\mathbf{x}) = l_2(\mathbf{x}) - \tilde{l}_1(\mathbf{x}) \quad (10.74)$$

$\tilde{l}_1(\mathbf{x})$ 由式(10.73)定义, $l_2(\mathbf{x})$ 是由第二译码段计算得到的对数似然比。特别地, 对于矢量 \mathbf{x} 的第 i 位元素, 有

$$l_2(x_i) = \log_2 \left(\frac{P(x_i = 1 | \mathbf{u}, \zeta_2, \tilde{l}_1(\mathbf{x}))}{P(x_i = 0 | \mathbf{u}, \zeta_2, \tilde{l}_1(\mathbf{x}))} \right), \quad i = 1, 2, \dots, k \quad (10.75)$$

由于第一译码段中 $\tilde{l}_2(\mathbf{x})$ 的作用, 围绕着两个译码段的反馈回路将闭合。注意, 虽然含有噪声的系统比特组 \mathbf{u} 只应用于图 10.28(a) 所示的第一译码段, 但是, 如果按照图 10.28(b) 所述的外部对称方式来将信息流公式化, 就会发现, \mathbf{u} 实际上也应用到了第二译码段。

信息比特组 \mathbf{x} 的估计值是在第二译码段的输出端, 通过硬限制对数似然比 $l_2(\mathbf{x})$ 进行计算得到的, 可表示为

$$\hat{\mathbf{x}} = \text{sgn}(l_2(\mathbf{x})) \quad (10.76)$$

其中, 符号函数分别对 $l_2(\mathbf{x})$ 的每个元素进行运算。

为了对 Turbo 码的译码算法进行初始化, 需在算法的第一次迭代时令 $\tilde{l}_2(\mathbf{x}) = 0$, 如图 10.28(b) 所示。

在图 10.28 所示的 Turbo 译码器中, 只将外部信息从一个译码段馈入另一个译码段, 这主要是为了尽可能保持本次迭代与下次迭代之间的比特统计独立性。此处描述的反馈译码策略就是依赖于这种假设。如果这种统计独立性的假设严格成立, 则当迭代的次数 $\hat{\mathbf{x}}$ 趋近无穷时, 由式(10.76)定义的估计值就趋近于最大 MAP 的解。^[18] 这个统计独立性的假设在绝大多数现实情况下是近似成立的。

10.8.4 BCJR 算法

为了使 Turbo 译码的讨论更完整, 有必要介绍为最大 MAP 估计而设计的 BCJR 算法的数学

解释。

令 $x(t)$ 为网格编码器在 t 时刻的输入, $y(t)$ 为在接收端观测到的相应输出。注意, 此处 $y(t)$ 可以有多种观测结果, 例如, 对于码率为 $1/n$ 的码, 每比特输入生成 n 比特的码。在这种情况下, 可定义一个 n 维的观察矢量

$$\mathbf{y}_{(1,t)} = [y(1), y(2), \dots, y(t)]$$

令 $\lambda_m(t)$ 表示网格编码器状态 $s(t)$ 为 m 的概率, 其中, $m = 1, 2, \dots, M$ 。于是有

$$\lambda(t) = P[s(t) | \mathbf{y}] \quad (10.77)$$

其中, $s(t)$ 和 $\lambda(t)$ 都是 $M \times 1$ 的矢量。因此, 对于具有与 RSC 码相同反馈的码率为 $1/n$ 的线性卷积码, 信息比特为 1 的概率为

$$P(x(t) = 1 | \mathbf{y}) = \sum_{s \in \mathcal{S}_A} \lambda_s(t) \quad (10.78)$$

其中, \mathcal{S}_A 是在输入端对应于符号 1 的一系列变换, $\lambda_s(t)$ 是 $\lambda(t)$ 的 s 分量。

上述状态概率的前向估计值可定义为一 $M \times 1$ 矢量

$$\alpha(t) = P(s(t) | \mathbf{y}_{(1,t)}) \quad (10.79)$$

状态概率 $\mathbf{y}_{(1,t)}$ 的后向估计值也可定义为一 $M \times 1$ 矢量

$$\beta(t) = P(s(t) | \mathbf{y}_{(t,k)}) \quad (10.80)$$

其中:

$$\mathbf{y}_{(t,k)} = [y(t), y(t+1), \dots, y(k)]$$

矢量 $\alpha(t)$ 和 $\beta(t)$ 分别对应于过去的数据和将来的数据。因此, 可将分离定理定义如下:

t 时刻的状态概率与前向估计值 $\alpha(t)$ 和后向估计值 $\beta(t)$ 的关系是

$$\lambda(t) = \frac{\alpha(t) \cdot \beta(t)}{\|\alpha(t) \cdot \beta(t)\|_1} \quad (10.81)$$

其中, $\alpha(t) \cdot \beta(t)$ 是 $\alpha(t)$ 和 $\beta(t)$ 的矢量乘积。 $\|\alpha(t) \cdot \beta(t)\|_1$ 是矢量乘积的模 L_1 。

矢量乘积 $\alpha(t) \cdot \beta(t)$ (不同于内积) 由 $\alpha(t)$ 和 $\beta(t)$ 的单个元素定义:

$$\alpha(t) \cdot \beta(t) = \begin{bmatrix} \alpha_1(t)\beta_1(t) \\ \alpha_2(t)\beta_2(t) \\ \vdots \\ \alpha_M(t)\beta_M(t) \end{bmatrix} \quad (10.82)$$

$\alpha(t) \cdot \beta(t)$ 的模 L_1 为

$$\|\alpha(t) \cdot \beta(t)\|_1 = \sum_{m=1}^M \alpha_m(t)\beta_m(t) \quad (10.83)$$

分离定理就是指由过去状态决定的 t 时刻的状态分配, 与由将来状态决定的 t 时刻的状态分配是相互独立的。分离定理符合编码的马尔可夫假设, 是 BCJR 算法的基础。同时, 分离定理为用一种简单的方法将前向和后向估计值进行合成以得到状态概率的完整描述奠定了基础。

可进一步定义 t 时刻的状态转移概率为

$$\gamma_{m',m}(t) = P(s(t) = m, y(t) | s(t-1) = m') \quad (10.84)$$

并将 $M \times M$ 的转移概率矩阵表示为

$$\Gamma(t) = \{\gamma_{m',m}(t)\} \quad (10.85)$$

这样,就可以将递归定理用公式表示为:

前向估计值 $\alpha(t)$ 和后向估计值 $\beta(t)$ 可以由下式递归计算得到:

$$\alpha^T(t) = \frac{\alpha^T(t-1)\Gamma(t)}{\|\alpha^T(t-1)\Gamma(t)\|_1} \quad (10.86)$$

和

$$\beta(t) = \frac{\Gamma(t+1)\beta(t+1)}{\|\Gamma(t+1)\beta(t+1)\|_1} \quad (10.87)$$

其中,上标 T 表示矩阵的转置。

在给定观察矢量的条件下,分离与递归定理共同定义了计算网格码状态和状态转移后验概率的 BCJR 算法。这样,就可以通过计算所需状态的被选择子集的总和得到 Turbo 译码所需的似然比。

10.9 计算机实验:Turbo 码的译码

Turbo 码具有以下两个性质:

性质 1

Turbo 码的误码性能随着译码算法迭代次数的增加而得到改善。要实现这一点,可以在前向通道中将外部信息由第一译码段的输出馈入第二译码段的输入,并在后向通道中将外部信息由第二译码段的输出馈入第一译码段的输入,然后让迭代的译码过程根据接收的带噪声信息和奇偶校验比特自然进行。

性质 2

Turbo 译码器在采用一种可行的计算方式后,可接近信道容量的香农理论极限。这一性质已在实验中得到验证,但尚未进行理论验证。

性质 2 要求 Turbo 码的分组长度足够大。遗憾的是,这个性质的证明要用到 Turbo 译码算法的精确计算,这已超出了本书的范围。因此,本实验将主要关注性质 1。

本实验的主要目的,就是用 BCJR 算法中的对数最大后验概率来证明 Turbo 译码的性质 1。实验中仅考虑 AWGN 信道。Turbo 编译码器的具体情况如下:

Turbo 编码器(如图 10.25 所示):

编码器 1:[1,1,1]卷积码

编码器 2:[1,0,1]卷积码

分组(即交织)长度:1200 比特

Turbo 译码器(如图 10.28 所示):

采用最大后验概率译码的 BCJR 算法。

实验分别在 $E_b/N_0 = 1 \text{ dB}, 1.5 \text{ dB}, 2 \text{ dB}, 2.5 \text{ dB}$ 的情况下进行, 每个 E_b/N_0 值都对应于不同的迭代次数。对于每次实验, 当噪声中断的分组数累计达到 20 后, 计算总的比特差错数(分组长度为 1200 比特)。然后, 用比特差错数与被编码的总比特数的比值来估算误码率。值得注意的是, 在计算过程中, 有许多编码比特的分组是正确译码的。

实验结果如图 10.30。从图中可得出以下结论:

1. 对于一个给定的 E_b/N_0 , 随着迭代次数的增加, 误码率下降。这就证明了 Turbo 译码的性质 1。
2. 8 次迭代之后, 译码性能不会有显著的改善。
3. 在固定的迭代次数下, 误码率随着 E_b/N_0 的增加而降低。这正是我们所期望的。

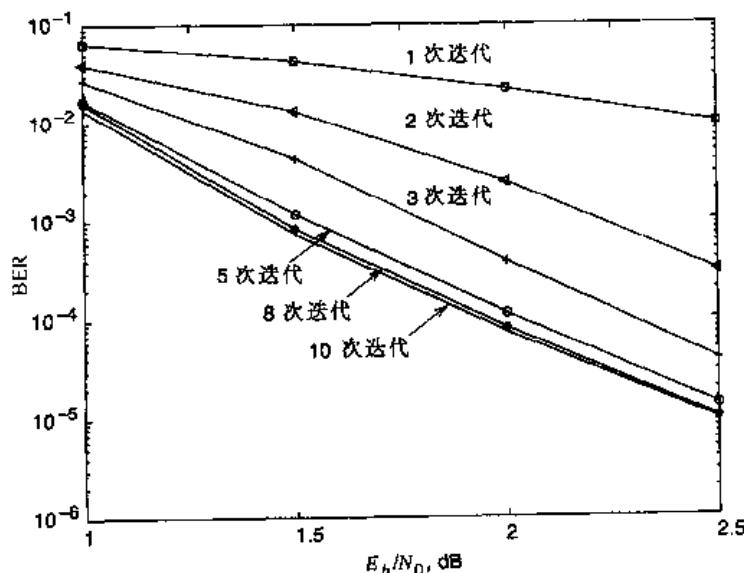


图 10.30 Turbo 译码的计算机实验结果

10.10 低密度奇偶校验码^[19]

10.8 节中讨论的 Turbo 码和本节讨论的低密度奇偶校验码(LDPC)同属于差错控制编码技术的一个大类, 即组合码。与 Turbo 码相比, LDPC 码具有两个重要的优点:

- 没有低码重的码字
- 迭代译码的复杂度较低

低码重码字的现象, 是指在 Turbo 码中常常出现的少数码字与给定码字很接近的情况, 这是我们所不希望的。由于码重相近, 信道噪声的存在有时会使得发送码字被误认为是与其临近的码字。实际上, 这主要是由于前面提到的误码底限(误比特率一般在 $10^{-5} \sim 10^{-6}$ 之间)引起的。相反地, LDPC 码却能够很容易地进行重构, 并且不含有这种低码重的码字, 从而获得很小(近似为零)的误比特率。当然, 在 Turbo 码中, 对交织器进行特殊设计, 也能够使误码底限的问题得到减轻。

下一个问题是译码的复杂度问题。Turbo译码器的计算复杂度由BCJR算法决定,而BCJR算法是在编码器卷积码的网格图上进行的。BCJR算法每次递归的计算次数与网格图中状态数目成线性关系。Turbo码使用的网格图一般有16个或更多的状态。相反,LDPC码使用的是只有两个状态的奇偶校验网格图。所以,与Turbo译码器相比,LDPC译码器要简单得多。同时,LDPC译码器的运行速度也高于Turbo译码器。

但是,由于采用LDPC码的目的是为了用于大的分组长度,所以,其编码复杂度要高于Turbo码。

10.10.1 LDPC码的构造

LDPC码由用 \mathbf{A} 表示的奇偶校验矩阵确定, \mathbf{A} 是一个“稀疏”矩阵。也就是说,其元素多数为0,少数为1。通常所说的 (n, t_c, t_r) LDPC码中, n 表示分组的长度, t_c 表示矩阵 \mathbf{A} 中各列的码重(也就是1的个数), t_r 表示 \mathbf{A} 中各行的码重, $t_r > t_c$ 。该LDPC码的码率为:

$$r = 1 - \frac{t_c}{t_r} \quad (10.88)$$

下面证明上式的正确性。令 ρ 表示奇偶校验矩阵 \mathbf{A} 中1的密度。由10.3节可得

$$t_c = \rho(n - k)$$

和

$$t_r = \rho n$$

其中, $n - k$ 是矩阵 \mathbf{A} 的行数, n 是矩阵 \mathbf{A} 的列数(即分组长度)。用 t_c 除以 t_r ,得到

$$\frac{t_c}{t_r} = 1 - \frac{k}{n}$$

由定义,分组码的码率为 k/n ,因此,式(10.88)成立,其条件是矩阵 \mathbf{A} 中各行必须线性独立。

LDPC码的结构可用二分图来表示,如图10.31所示(以 $n = 10, t_c = 3, t_r = 5$ 的码为例)。图中,左边的节点是可变节点,分别对应于码字中各元素。右边的节点是校验节点,分别对应于码组中各码字满足的一组奇偶校验约束条件。以图10.31为例的这类LDPC码被认为是规则的,因为其中所有相似的节点都对应于相同的等级。在图10.31所示的例子中,可变节点的等级 $t_c = 3$,校验节点的等级 $t_r = 5$ 。当分组长度 n 趋近于无穷时,每个校验节点只对应于可变节点的很小一部分,因此称为是低密度的。

矩阵 \mathbf{A} 是通过在矩阵 \mathbf{A} 中随机地放置1构造出来的,但也需要满足下列约束规则:

- 矩阵中每列包含有固定数目的(即 t_c 个)1
- 矩阵中每行包含有固定数目的(即 t_r 个)1

事实上,为了避免奇偶校验矩阵 \mathbf{A} 中行的线性相关,这些约束规则经常被破坏。

与10.3节中讨论的线性分组码不同,LDPC码的奇偶校验矩阵 \mathbf{A} 不是系统形式的(即奇偶校验比特不是对角线形式的),因此其中采用的符号也与10.3节中不同。可用模2的高斯消去法推导LDPC的生成矩阵(见例10.9)。如10.3节,可将 $1 \times n$ 的码矢量 \mathbf{c} 分割为

$$\mathbf{c} = [\mathbf{b} : \mathbf{m}]$$

其中, \mathbf{m} 是 $1 \times n$ 的信息矢量, \mathbf{b} 是 $1 \times (n - k)$ 的校验矢量, 见式(10.9)。相应地, 奇偶校验矩阵 \mathbf{A} 可以分割为

$$\mathbf{A}^T = \begin{bmatrix} \mathbf{A}_1 \\ \dots \\ \mathbf{A}_2 \end{bmatrix} \quad (10.89)$$

其中, \mathbf{A}_1 是 $(n - k) \times (n - k)$ 的方阵, \mathbf{A}_2 是 $k \times (n - k)$ 的矩阵。为方便起见, 用上标 T 表示分块阵的转置。将式(10.16)的约束条件应用到 LDPC 码, 有

$$[\mathbf{b} : \mathbf{m}] \begin{bmatrix} \mathbf{A}_1 \\ \dots \\ \mathbf{A}_2 \end{bmatrix} = 0$$

等价地, 有

$$\mathbf{b}\mathbf{A}_1 + \mathbf{m}\mathbf{A}_2 = 0 \quad (10.90)$$

由式(10.7)可知, 矢量 \mathbf{m} 和 \mathbf{b} 有如下关系:

$$\mathbf{b} = \mathbf{m}\mathbf{P}$$

其中, \mathbf{P} 是系数矩阵。将此关系式代入式(10.90), 不难发现, 对于任意的非零信息矢量 \mathbf{m} , LDPC 码的系数矩阵满足条件

$$\mathbf{P}\mathbf{A}_1 + \mathbf{A}_2 = 0$$

这个等式适用于所有非零的信息矢量, 尤其是 \mathbf{m} 等于 $[0 \cdots 0 \ 1 \ 0 \cdots 0]$ 的情况, 此时将分离出生成矩阵的单独一行。

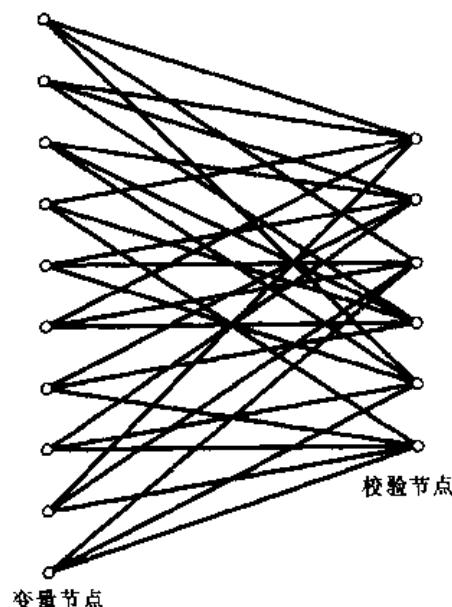


图 10.31 (10,3,5)LDPC 码的二分图

由上式可得

$$\mathbf{P} = \mathbf{A}_2 \mathbf{A}_1^{-1} \quad (10.91)$$

其中,矩阵 \mathbf{A}_1^{-1} 是矩阵 \mathbf{A}_1 的逆矩阵(模 2)。最后,LDPC 码的生成矩阵可定义为

$$\begin{aligned} \mathbf{G} &= [\mathbf{P} : \mathbf{I}_k] \\ &= [\mathbf{A}_2 \mathbf{A}_1^{-1} : \mathbf{I}_k] \end{aligned} \quad (10.92)$$

其中, \mathbf{I}_k 是 $k \times k$ 的单位矩阵,见式(10.12)。

值得注意的是,如果对任意的 LDPC 码取奇偶校验矩阵 \mathbf{A} ,并从中随机地取出 $n - k$ 列构成方阵 \mathbf{A}_1 ,即使 \mathbf{A} 的各行是线性独立的,也不能保证 \mathbf{A}_1 是非退化阵(逆阵 \mathbf{A}_1^{-1} 存在)。实际上,一个典型的 LDPC 码一般具有很大的分组长度 n 。由于随机选择的矩阵 \mathbf{A}_1 中至少有一行为全零的可能性很大,所以, \mathbf{A}_1 为非退化阵的可能性很小。当然,当 \mathbf{A} 中各行线性独立时,其中会存在某些 $(n - k)$ 列能够生成非退化矩阵 \mathbf{A}_1 ,如例 10.9 所述的情况。对于 LDPC 码的某些构造方法来讲, \mathbf{A} 中最前面的 $(n - k)$ 列可以保证(或者说有很大的可能性)生成非退化的矩阵 \mathbf{A}_1 ,但通常这种情况是不存在的。

例 10.9 (10,3,5)LDPC 码

设有一 LDPC 码,可由图 10.31 的二分图来描述。那么,这个码的奇偶校验矩阵可以定义为

$$\mathbf{A} = \left[\begin{array}{cccccc|ccc} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ \hline 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \end{array} \right] \quad \begin{matrix} \mathbf{A}_1^T \\ \mathbf{A}_2^T \end{matrix}$$

上式在满足了约束规则 $t_c = 3$ 和 $t_r = 5$ 的同时,表现出随机性。按照式(10.89)所示的方式将矩阵 \mathbf{A} 分割为

$$\mathbf{A}_1 = \left[\begin{array}{ccccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \end{array} \right]$$

$$\mathbf{A}_2 = \left[\begin{array}{ccccc} 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array} \right]$$

为推导出 \mathbf{A}_1 的逆矩阵,先将式(10.90)写为

$$\underbrace{[b_0, b_1, b_2, b_3, b_4, b_5]}_b \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}_{A_1} = \underbrace{[u_0, u_1, u_2, u_3, u_4, u_5]}_u = m A_2$$

其中,引入矢量 u 来表示矩阵 m 和 A_2 的乘积。通过高斯消去法,可将 A_1 化为下三角阵(即主对角线上的元素均为 0):

$$A_1 \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

这个变换可通过在方阵 A_1 的列之间进行模 2 加运算得到:

- 第 1 列和第 2 列加到第 3 列上。
- 第 2 列加到第 4 列上。
- 第 1 列和第 4 列加到第 5 列上。
- 第 1 列、第 2 列和第 5 列加到第 6 列上。

相应地,矢量 u 转化为:

$$u \rightarrow [u_0, u_1, u_0 + u_1 + u_2, u_1 + u_3, u_0 + u_3 + u_4, u_0 + u_1 + u_4 + u_5]$$

因此,在变换后的 A_1 前面乘上一个校验矢量 b ,再运用反向的逐次消元法(模 2),即可求得与矢量 u 中各元素相对应的矢量 b 的元素:

$$\underbrace{[u_0, u_1, u_2, u_3, u_4, u_5]}_u \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}_{A_1^{-1}} = \underbrace{[b_0, b_1, b_2, b_3, b_4, b_5]}_b$$

因此,矩阵 A_1 的逆矩阵为

$$A_1^{-1} = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

于是,矩阵乘积 $A_2 A_1^{-1}$ 为(使用给定的 A_2 的值和刚求得的 A_1^{-1} 的值):

$$A_2 A_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

最后,由式(10.92)可得(10,3,5)LDPC 码的生成矩阵为

$$G = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\underbrace{A_2 A_1^{-1}}_{I_k}$$

本例只是为了说明 LDPC 码的生成过程。在实际应用中,分组的长度 n 要比本例中大几个数量级。在构造矩阵 A 时,可强制规定矩阵交叠(即矩阵 A 任意两列之间的内积)不超过 1。这个限制条件虽然超出了约束规则,但可以改善 LDPC 码的性能。遗憾的是,本例中采用的小的分组长度,是很难满足这个附加要求的。

10.10.2 LDPC 码的最小码距

在实际应用中,LDPC 码的分组长度是很大的,可以从 $10^3 \sim 10^6$,也就是说,在特定的码组中,可用码字的数目相应地会很大。所以,对 LDPC 码进行代数分析是相当困难的。但是,对一个 LDPC 码集进行统计分析却是相对容易的。这些分析使得我们能够在统计意义上得到码集中成员码的某些特性。此外,我们还发现:在对该 LDPC 码集进行随机抽取时,具有这些特性的码被选中的概率较高。

在这些特性中,成员码的最小码距是我们最为关心的。回顾 10.3 节,线性分组码最小码距的定义是该码组中任意两个码矢量之间的最小汉明距离。对于一个 LDPC 码集来说,成员码的最小码距显然是一个随机变量。另外,当分组长度 n 增加时,对于固定的 $t_c \geq 3$ 且 $t_r > t_c$,最小码距的概率分布可能会近似等于一个在分组长度 n 的固定分段点 $\Delta_{t_c t_r}$ 处的单位阶跃函数,从而超界。^[20]因此,对于大的 n 值,几乎码集中所有的 LDPC 码都有至少为 $n\Delta_{t_c t_r}$ 的最小码距。表 10.10 给出了不同的 t_c 和 t_r 值对应的 LDPC 码的码率 r 和 $\Delta_{t_c t_r}$ 。从表中可以看出:当 $t_c = 3$ 且 $t_r = 6$ 时,码率取最大值 $1/2$,在分段点处码率取最小值。因此,在设计 LDPC 码时,我们倾向于选择 $t_c = 3, t_r = 6$ 。

表 10.10* 不同码重 t_c 和 t_r 对应的 LDPC 码的码率和分段项 $\Delta_{t_c t_r}$

t_c	t_r	码率 r	$\Delta_{t_c t_r}$
5	6	0.167	0.255
4	5	0.2	0.210
3	4	0.25	0.122
4	6	0.333	0.129
3	5	0.4	0.044
3	6	0.5	0.023

* 经 IEEE 允许摘自 Gallager(1962)。

10.10.3 LDPC 的概率译码

在发射机中,信息矢量 \mathbf{m} 被编码为码矢量 $\mathbf{c} = \mathbf{m}\mathbf{G}$,其中, \mathbf{G} 是采用指定码重对(t_c, t_r)和最小码距 d_{\min} 的编码对应的生成矩阵。码矢量 \mathbf{c} 经噪声信道传输后产生接收矢量

$$\mathbf{r} = \mathbf{c} + \mathbf{e}$$

其中, \mathbf{e} 是由信道噪声引起的误码矢量,见式(10.17)。经构造可得矩阵 \mathbf{A} ,它是 LDPC 码的奇偶校验矩阵,且 $\mathbf{AG}^T = 0$ 。假设接收矢量为 \mathbf{r} ,逐比特译码的问题就是要找到一个最可能的矢量 $\hat{\mathbf{c}}$,使之满足条件 $\hat{\mathbf{c}}\mathbf{A}^T = 0$ 。

在下文中,一个比特是指接收矢量 \mathbf{r} 中的一个元素,一个校验是指矩阵 \mathbf{A} 中的一行。令 $\mathcal{S}(i)$ 表示校验 i 中的一组比特, $\mathcal{S}(j)$ 表示含有比特 j 的一组校验。在比特组 $\mathcal{S}(i)$ 中去掉比特 j 用 $\mathcal{S}(i) \setminus j$ 来表示。同样,在比特组 $\mathcal{S}(j)$ 中去掉校验 i 用 $\mathcal{S}(j) \setminus i$ 来表示。

译码算法由相互交替的两个步骤组成:水平步骤和垂直步骤。这两个步骤分别沿矩阵 \mathbf{A} 的行和列进行。在这两个步骤中,与矩阵 \mathbf{A} 中非零元素相关的统计量交替更新。其中, P_i^x 表示在水平步骤中,通过除 i 以外的校验,从而得到的比特 j 等于符号 x (x 为 0 或 1)的概率。 Q_j^x 表示比特 j 为固定值 x ,且其他比特的概率为 $P_{ij}:j' \in \mathcal{S}(i) \setminus j$ 时,满足校验 i 的概率。

LDPC 译码算法的步骤如下:^[21]

初始化

令变量 P_j^0 和 P_j^1 分别等于符号 0 和 1 的先验概率 p_j^0 和 p_j^1 ,有 $p_j^0 + p_j^1 = 1$ 。

水平步骤

水平步骤通过校验 i 来执行。定义

$$\Delta P_{ij} = P_{ij}^0 - P_{ij}^1$$

对于每个码重对(i, j),计算

$$\Delta Q_{ij} = \prod_{j' \in \mathcal{S}(i) \setminus j} \Delta P_{ij'}$$

因此,假设

$$Q_j^0 = \frac{1}{2} (1 + \Delta Q_{ij})$$

$$Q_j^1 = \frac{1}{2} (1 - \Delta Q_{ij})$$

垂直步骤

在垂直步骤中, P_j^0 和 P_j^1 的概率值用水平步骤中计算得到的量来更新。特别地,对每个比特 j ,计算

$$P_{ij}^0 = \alpha_{ij} p_j^0 \prod_{i' \in \mathcal{S}(j) \setminus i} Q_{i'}^0,$$

$$P_{ij}^1 = \alpha_{ij} p_j^1 \prod_{i' \in \mathcal{S}(j) \setminus i} Q_{i'}^1$$

其中,范围因子 α_{ij} 的选择是为了使得

$$P_i^0 + P_i^1 = 1$$

在垂直步骤中,我们也可以更新伪后验概率

$$P_i^0 = \alpha_j p_i^0 \prod_{i \in s(j)} Q_i^0$$

$$P_i^1 = \alpha_j p_i^1 \prod_{i \in s(j)} Q_i^1$$

其中, α_j 的选择是为了使得

$$P_i^0 + P_i^1 = 1$$

由垂直步骤得到的结果被用于计算临时的估计值 $\hat{\mathbf{c}}$ 。如果条件 $\hat{\mathbf{c}}\mathbf{A}^T = 0$ 得到满足,译码算法就会终止。否则,算法就回到水平步骤继续进行。如果经过最大次数(100 或 200 次)的迭代后仍不能得到有效的译码,则可宣告译码失败。这里描述的译码过程是一般低复杂度和积算法(sum-product algorithm)的特殊情况。

简单地说,和积算法用来传递二分图的校验节点与变量节点之间的统计量。由于奇偶校验约束条件可以用一个简单的单比特存储卷积码来表示,所以,如前所述,LDPC译码器实现起来比Turbo译码器更为简单。

但从性能上来讲,根据文献报道的实验结果,常规的 LDPC 码不能同对应的Turbo码一样接近香农极限。

10.11 不规则码

第 10.8 节讨论的 Turbo 码和第 10.10 节讨论的 LDPC 码都分别符合各自的规则。这两种码经过噪声信道后的纠错性能可以通过采用各自的不规则形式来得到提高。

在图 10.25 所示的带编码器的标准 Turbo 码中,交织器将每个系统比特映射为卷积码编码器 2 唯一的一个输入比特。相反,不规则的 Turbo 码^[22]则用特殊设计的交织器将几个系统比特映射为卷积码编码器的多个输入比特。例如,每 10% 的系统比特可以被映射为卷积编码器的 8 个输入,而不仅仅是一个。如图 10.32 所示,在两条编码路径中均采用了类似的不规则交织器,用来产生对应于信息比特 x 的奇偶校验比特 z_1 和 z_2 。不规则 Turbo 码的译码采用与规则 Turbo 码类似的方式。

为了构造一个不规则的 LDPC 码,^[23]我们可以根据某些分布特性来选择二分图中的可变节点和校验节点的级数。例如,可按照下述方法来构造不规则的 LDPC 码:

- 一半的可变节点为 5 级,另一半可变节点为 3 级。
- 一半的校验节点为 6 级,另一半校验节点为 8 级。

对给定的分组长度和等级序列,可随机选取边界(可变节点和校验节点间的连接)来定义一个码集。特别地,由可变节点发出的边界次序是任意的,对于校验节点也同样如此。

图 10.33 显示了下列码的误码性能:^[24]

- 不规则的 LDPC 码: $k = 50\,000, n = 100\,000$, 码率 = 1/2
- 规则的 Turbo 码: $k = 65\,536, n = 131\,072$, 码率 = 1/2

- 不规则 Turbo 码: $k = 65\,536, n = 131\,072$, 码率 = 1/2

其中, k 是信息比特数, n 是分组长度。规则和不规则的 Turbo 码的卷积编码器的生成多项式为

$$\text{编码器 } 1: g(D) = 1 + D^4$$

$$\text{编码器 } 2: g(D) = 1 + D + D^2 + D^3 + D^4$$

图 10.33 中还给出了码率为 $r = 1/2$ 的相应的信道容量理论极限。

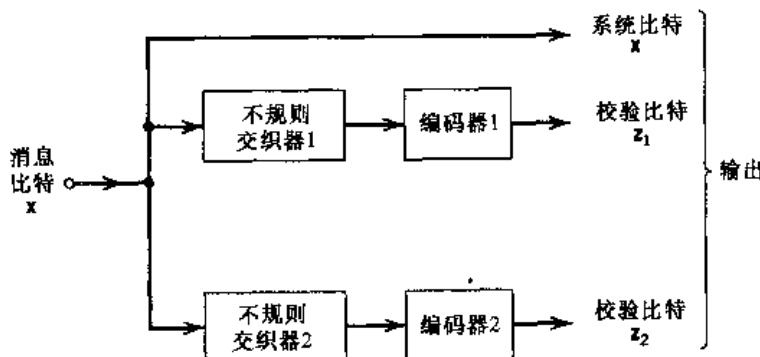


图 10.32 不规则 Turbo 码的框图

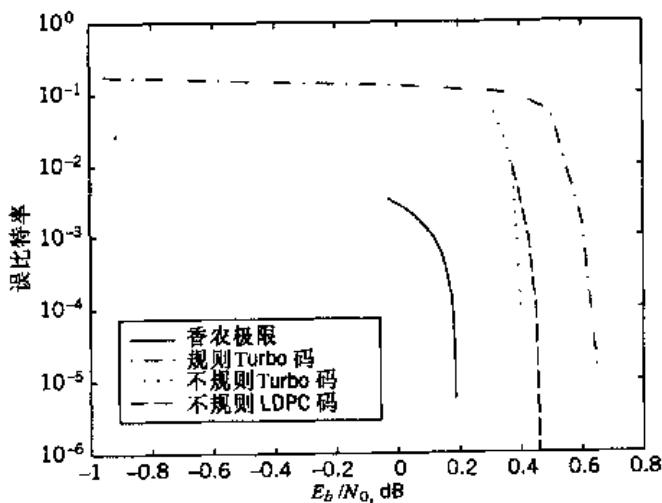


图 10.33 规则 Turbo 码、不规则 Turbo 码和不规则 LDPC 码的噪声性能

基于图 10.33 给出的结果, 我们可以得到如下结论:

- 不规则的 LDPC 码性能优于规则的 Turbo 码, 因为它比规则 Turbo 码更接近于香农理论极限 0.175 dB;
- 三个码中, 不规则 Turbo 码的性能最好, 因为它仅与香农理论极限相差 0.213 dB。

10.12 总结与讨论

本章介绍了差错控制编码技术, 这些技术是在噪声信道上进行可靠数字通信的必不可少的工具。通过在数据发送之前有控制地加入冗余, 可减小传输过程中误码的影响。这些冗余

也使接收机中的译码器能够进行误码检测和纠错。

差错控制编码技术可以分为两大类：

1. 代数码。在接收端,依靠在设计码时建立的抽象代数结构来进行译码。代数码包括汉明码,最大长度码,BCH 码和 RS 码。这些码具有两个共同特性:

线性特性,即码组中任意两个码字的和也是一个码字。

循环特性,即码组中任意一个码的循环移位仍然是码组中的码字。

RS 码是非常有效的码,它能同时纠正随机和突发的误码。RS 码已在一些恶劣环境如深空(deep-space)通信中,以及 CD 盘片中得到广泛的应用。

2. 概率码。在接收机中,主要依靠统计方法来译码。概率码包括网格编码调制、Turbo 码和低密度奇偶校验码。特别地,译码是基于两种基本方式中的某一种进行的,这两种基本方式为:

软输入硬输出。网格编码调制的译码中采用的最大似然序列估计的维特比算法就是一个例子。

软输入软输出。Turbo 译码过程中基于逐比特最大后验概率估计的 BCJR 算法,以及低密度奇偶校验码译码时采用的特定形式的和积算法,都是软输入软输出译码的例子。

网格编码调制将线性卷积码的编码与调制相结合,使之与传统的未编码的多电平调制相比,在不牺牲带宽利用率的基础上,显著提高了编码增益。Turbo 码与 LDPC 码具有如下的共同特性:

- 都是线性分组的随机编码。
- 在可实现的方式下,误码性能可非常接近香农信道容量的理论极限。

在实际情况下,Turbo 码和低密度奇偶校验码都使得 10 dB 数量级的编码增益成为可能,这在过去是不可能达到的。这些编码增益可以被用来显著提高数字通信接收机的动态范围,主要是增加数字通信系统的比特率,或使每符号的发送能量有明显降低。这些优点对于数字通信的两种重要应用——无线通信和深空通信的设计具有重大意义。事实上,Turbo 码在深空通信和无线通信系统中的应用已经标准化。

注释与参考

- [1] 关于纠错编码的介绍性讨论可参阅 Lucky(1989)的第 2 章,也可参阅 Adámek(1991)和 Bhargava(1983)。差错控制编码的经典书籍由 Peterson 和 Weldon(1972)所著。在 Gallager(1968)的经典著作中也有关于差错控制编码的讨论。Lin 和 Costello(1983),Micheleson 和 Levesque(1985),MacWilliams 和 Sloane(1977)以及 Wilson(1998)都对差错控制编码的研究做出了贡献。关于编码理论的关键论文集可参阅 Berlekamp(1974)。
- [2] 要了解不同的 ARQ 方案,可参阅 Lin, Costello 以及 Miller(1984)。
- [3] 在医学上,综合症状(syndrome)一词是用来描述那些有助于疾病诊断的症状。在编码学中,错误图样扮演了疾病的角色而奇偶校验的出错扮演了症状的角色。校正子(syndrome)是由 Hagelbarger 发明的。

- [4] 第一个纠错编码(汉明码)几乎是在香农提出信息论的同时由 Hamming 发明的。详细情况可参阅 Hamming(1950)的经典论文。
- [5] 对 BCH 码及其译码算法的描述可参阅 Lin 和 Costello(1983, pp.141 – 183)以及 MacWilliams 和 Sloane(1977, pp.257 – 293)。关于二进制 BCH 码的表 10.6 改编自 Lin 和 Costello(1983)。
- [6] 里德 – 索罗蒙码是为了对其发明者表示敬意而命名的, 可参阅他们 1960 年的经典论文。里德 – 索罗蒙码的详细情况可参阅 MacWilliams 和 Sloane(1977, pp.294 – 306)。由 Wicker 和 Bhargava(1994)编写的书中包含了对里德 – 索罗蒙码的介绍, 对其发明者 Irving S. Reed 和 Gustave Solomon 的历史回顾, 里德 – 索罗蒙码在太阳系探索、CD 光盘、ARQ 协议和扩频通信中的应用, 以及其他相关问题。
- [7] 将卷积码作为分组码的替代是由 P. Elias(1955)第一次提出的。
- [8] 网格图一词是由 Forney(1973)提出的。
- [9] 在 Viterbi(1967)的经典论文中, 提出了一种卷积码的译码算法, 即著名的维特比算法。这种算法由 Forney(1972, 1973)证明是一种最大似然译码算法。维特比算法的通俗说明由 Lin 和 Costello(1983), 以及 Blahut(1990)和 Adámek(1991)给出。
- [10] 恶性卷积码在 Benedetto, Biglieri 以及 Castellani(1987)的著作中进行了讨论。表 10.8 改编自该书。
- [11] 二进制对称信道和二进制输入的 AWGN 信道的渐近编码增益估算的详细情况, 可参阅 Viterbi 和 Omura(1979, pp.242 – 252), 以及 Lin 和 Costello(1983, pp.322 – 329)。
- [12] 网格编码调制是由 G. Ungerboeck 发明的。其发展史在 Ungerboeck(1982)中有述及。表 10.9 改编自 Ungerboeck 后来的论文。
网格编码调制可以看成是编码信号空间的一种形式, 这个观点在 Lee 和 Messerschmitt(1994)的著作第 14 章中进行了初步的讨论。网格编码调制的进一步论述可参阅 Biglieri, Divsalar, McLane, Simon(1991)以及 Schlegel(1997)的著作。
- [13] Turbo 码是由 C. Berrou 和 A. Glavieux 发明的。Battail(1987)以及 Hagenauer 和 Hoecher(1989)的两篇论文推动了后续工作的展开。Turbo 码的首次提出采用的是启发式的讨论, 它是在 Berrou, Glavieux 和 Thitimajshima(1993)的一篇会议论文中被提出的, 也可参阅 Berrou 和 Glavieux(1996)的著作。关于 Turbo 码早期工作的回顾以及今后的发展, 可参阅 Berrou 和 Glavieux(1998)的著作。
关于 Turbo 码的基础书籍可参阅 Heegard 和 Wicker(1999)的著作。
通过对随机编码(见注释[15])的回顾, Benedetto 和 Montorosi(1996)给出了 Turbo 码突出性能的部分解释。
在 McEliece, MacKay 和 Cheng(1998)以及 Kschischang 和 Frey(1998)的论文中提出了两个相互独立的研究报告。这两个研究报告都证明了 Turbo 译码是 Pearl(1982)提出的人工智能中一种算法的翻版, 这种算法涉及到可信度的传播, 这里“可信度”一词是后验概率的另一种说法。这两篇论文开辟了一条 Turbo 码研究的新途径, 即将 Turbo 译码与人工智能联系起来。关于 Turbo 码有的其他见解可参阅 Frey(1998)的书。
伪随机交织器是 Turbo 码的基础。Denenshgaran 和 Mondin(1999)给出了为 Turbo 码设计交织器(如置换器)的系统过程。

- [14] Turbo 码的双重终止在 Guinand 和 Lodge(1996)的著作中进行了讨论。
- [15] 随机编码在 Cover 和 Thomas(1991)的著作中第 8.7 节进行了讨论。
- [16] 图 10.27 源自 Frey(1998)著作中的图 6.8。
- [17] 20 世纪 60 年代初, Baum 和 Welch 就导出了解决参数估计问题的迭代过程, 称为 Baum-Welch 算法, 参见 Baum 和 Petrie(1966), 以及 Baum 等人(1970)的著作。在 Bahl, Cocke, Jelinek 以及 Raviv(1974)的著作里被命名为 BCJR 的算法中, Baum-Welch 算法用于解决软输出和卷积码的最大似然译码的问题。
- [18] 当迭代次数趋近无穷时, 式(10.76)中估计值 \hat{x} 近似等于最大后验概率解。其证明过程在 Moher 和 Gulliver(1998)的论文中进行了讨论。
- [19] LDPC 码最初是由 Gallager(1962, 1963)发现的。后来又由 MacKey 和 Neal(1995)重新发现, 可参阅 MacKey(1999)。
20 世纪 60 年代及此后很长一段时间里, 由于当时的计算机还没有足够的能力来处理获得性能优异的 LDPC 码所需的分组长度, 因此 LDPC 码并没有得到很好的应用。
- [20] 对于码重对 (t_c, t_r) , LDPC 码最小码距的概率分布近似等于分组长度的单位阶跃函数。要了解这个论述的详细情况, 可参阅 Gallager(1962, 1963)。
- [21] 这里所述的 LDPC 码的译码算法是由 MacKey 和 Neal(1996, 1997)给出的。
- [22] 不规则 Turbo 码是由 Frey 和 MacKey(1999)发明的。
- [23] 不规则 LDPC 码是由 MacKey 等人(1999)以及 Richardson 等人(1999)分别发明的。
- [24] 性能如图 10.34 所示的码字的发明者分别为:
 • 规则 Turbo 码: Berrou 和 Glavieux(1996), 以及 Berrou 等人(1995)。
 • 不规则 Turbo 码: Frey 和 MacKey(1999)。
 • 不规则 LDPC 码: Richardson 等人(1999)。

习题

软判决编码

10.1 设有一个二进制输入 Q 进制输出的离散无记忆信道。如果信道的转移概率 $p(j|i)$ 满足条件

$$p(j|0) = p(Q-1-j|1), \quad j = 0, 1, \dots, Q-1$$

我们就称该信道为对称的。假设信道输入符号为 0 或 1 的概率相等。证明信道的输出符号也等概地出现, 即

$$p(j) = \frac{1}{Q}, \quad j = 0, 1, \dots, Q-1$$

10.2 二进制 PSK 信号的量化解调器如图 10.3(a)所示, 其中, 量化器是一个四级量化器, 如图 P10.2 所示。请估算二进制输入四进制输出的离散无记忆信道的转移概率。并由此证明该信道是对称信道。假设发射信号的每比特能量为 E_b , 加性高斯白噪声的均值为零, 功率谱密度为 $N_0/2$ 。

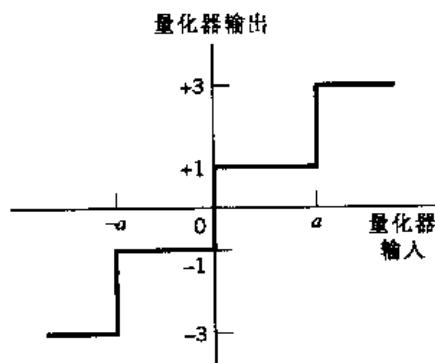


图 P10.2

- 10.3 考虑二进制输入的 AWGN 信道, 其中, 二进制符号 0 和 1 的概率相等。二进制符号以 PSK 方式在信道中传输。码的符号能量为 E , AWGN 信道的均值为零, 功率谱密度为 $N_0/2$ 。证明信道的转移概率为

$$p(y|0) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(y + \sqrt{\frac{2E}{N_0}}\right)^2\right], \quad -\infty < y < \infty$$

线性分组码和循环码

- 10.4 在单奇偶校验码中, 将一个单独的奇偶校验比特加到一组 k 个信息比特(m_1, m_2, \dots, m_k)之后。单奇偶校验比特 b_1 的选取要使码字满足偶校验规则

$$m_1 + m_2 + \dots + m_k + b_1 = 0, \quad \text{mod } 2$$

如 $k=3$, 请给出此码的 2^k 种可能出现的码字。

- 10.5 试比较例 10.2 中讨论的(7,4)汉明码的奇偶校验矩阵与(4,1)重复码的奇偶校验矩阵。
 10.6 考虑例 10.2 中的(7,4)汉明码, 其生成矩阵为 \mathbf{G} , 奇偶校验矩阵为 \mathbf{H} , 证明这两个矩阵满足

$$\mathbf{HG}^T = \mathbf{0}$$

- 10.7 (a) 对于例 10.2 中的(7,4)汉明码, 构造其对偶码的 8 个码字。
 (b) 求(a)中对偶码的最小码距。
 10.8 考虑例 10.1 的(5,1)重复码, 试估算下面两种误码图样的校正子:
 (a) 5 个所有的单错误图样
 (b) 10 个所有的双错误图样
 10.9 当只要求进行错误检测时, 可使用非系统码。本题研究这样一种循环码的生成。令 $g(X)$ 表示生成多项式, $m(X)$ 表示信息多项式。则码多项式 $c(X)$ 为

$$c(X) = m(X)g(X)$$

因此,如果给定了生成多项式,就可以方便地求出码组中的各个码字。为了说明这个过程,设(7,4)汉明码的生成多项式为

$$g(X) = 1 + X + X^3$$

试确定此码的16个码字,并证明此码的非系统性。

- 10.10 多项式 $1 + X^7$ 的本原因子有 $1 + X + X^3$ 和 $1 + X^2 + X^3$ 。在例10.3中,用 $1 + X + X^3$ 作为(7,4)汉明码的生成多项式。这里,我们考虑采用 $1 + X^2 + X^3$ 作为生成多项式。这样产生的(7,4)汉明码与例10.3中将有所不同。试设计对应于生成多项式

$$g(X) = 1 + X^2 + X^3$$

的编码器和校正子计算器。并将结果与例10.3进行比较。

- 10.11 设有一(7,4)汉明码,其生成多项式为

$$g(X) = 1 + X + X^3$$

码字0111001被送入一噪声信道,接收码字为0101001,其中有一个误码。试求这个接收码字的校正子多项式 $s(X)$,并证明该多项式与误码多项式 $e(X)$ 相同。

- 10.12 已知(15,11)汉明码的生成多项式为

此码采用系统形式,试设计相应的编码器和校正子计算器。

$$g(X) = 1 + X + X^4$$

- 10.13 如果(15,4)最大长度码是题10.12中(15,11)汉明码的对偶码。试完成:

- (a) 找出编码器的反馈连接,并与第7章中表7.1的最大长度码相比较。
- (b) 求出生成多项式 $g(X)$ 。并由此求出初始状态为0001时的输出序列。通过初始状态在编码器中的循环移动,验证结果的正确性。

- 10.14 设有(31,15)RS码。求:

- (a) 此码的一个符号包含有多少比特?
- (b) 按比特计算,分组长度是多少?
- (c) 码的最小码距是多少?
- (d) 此码可纠多少个符号差错?

卷积码

- 10.15 已知卷积码编码器包含一个两级单移位寄存器(约束长度 $K=3$),3个模2加法器和一个输出复接器。编码器的生成序列为

$$g^{(1)} = (1, 0, 1)$$

$$g^{(2)} = (1, 1, 0)$$

$$g^{(3)} = (1, 1, 1)$$

试画出编码器的方框图。

注意:为了便于比较不同编码器在同输入的情况下不同的输出,习题10.16~10.23中均采用信息序列10111…作为编码器的输入序列。

- 10.16 设有一如图P10.16所示的卷积码编码器,其码率 $r=1/2$,约束长度 $K=2$ 。码是系统形式的。试求出信息序列10111…经编码器后产生的输出。

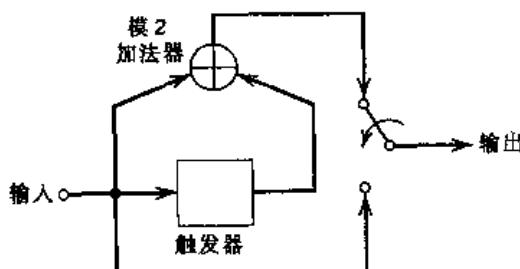


图 P10.16

10.17 设有一如图 P10.17 所示的卷积码编码器, 其码率 $r = 1/2$, 约束长度 $K = 4$ 。试求出信息序列 10111…经编码器后产生的输出。

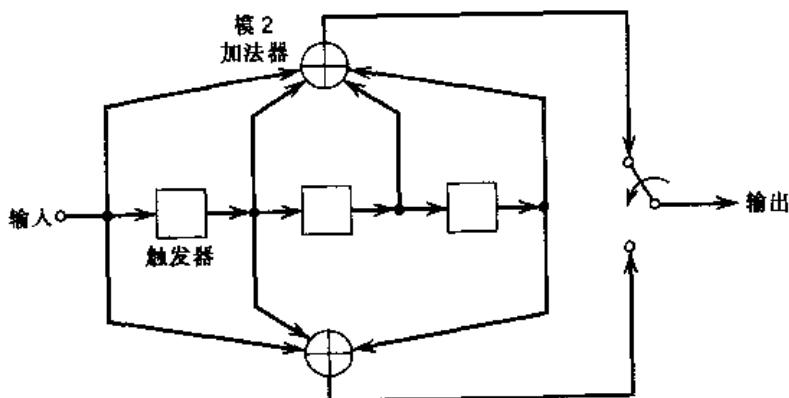


图 P10.17

10.18 设有一如图 10.13(b)所示的卷积码编码器, 其码率 $r = 2/3$, 约束长度 $K = 2$ 。试求出信息序列 10111…经编码器后产生的输出码序列。

10.19 画出图 P10.16 中所示卷积码编码器的码码树。并找出信息序列 10111…对应的码树中的路径, 将由此得出的编码器输出序列与题 10.16 的结果进行比较。

10.20 画出图 P10.17 中所示卷积码编码器的码码树。并找出信息序列 10111…对应的码树中的路径, 将由此得出的编码器输出序列与题 10.17 的结果进行比较。

10.21 画出图 P10.17 中所示卷积码编码器的网格图。设信息序列的长度为 5, 试找出信息序列 10111…对应的网格图中的路径, 将由此得出的编码器输出序列与习题 10.17 的结果进行比较。

10.22 画出图 P10.17 中所示卷积码编码器的状态图。设起始状态为全 0 状态, 试找出信息序列 10111…对应的状态图中的路径, 并将由此得出的编码器输出序列与习题 10.17 的结果进行比较。

10.23 设有一编码器如图 10.13(b)所示。

(a) 画出此编码器的状态图。

(b) 由全零状态开始, 画出信息序列 10111…对应的路径。并将由此得出的编码器输出序列与习题 10.18 的结果进行比较。

10.24 将 MSK 看成是有限状态机, 试画出 MSK 的网格图。对 MSK 的详细描述见第 6 章。

10.25 码率为 $1/2$, 约束长度为 3 的卷积码的网格图如图 P10.25 所示。设发送全零码, 而接收序列为 100010000…, 请用维特比算法计算译码后的输出序列。

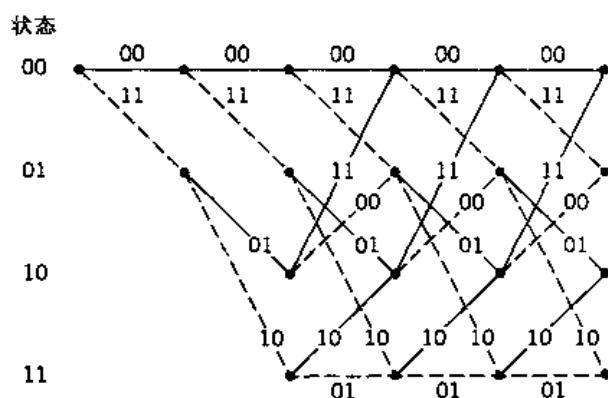


图 P10.25

10.26 设有一码率为 $1/2$, 约束长度为 7 的卷积码, 其自由距离 $d_{\text{free}} = 10$, 试计算下面两个信道的渐近编码增益:

- (a) 二进制对称信道
- (b) 二进制输入的 AWGN 信道

10.27 在 10.6 节中, 我们就卷积码的最大似然译码介绍了维特比算法。维特比算法的另一个应用是对色散信道中由于 ISI 引起中断的接收序列进行最大似然译码。对于二进制数据序列, 图 P10.27 中给出了 ISI 的网格图。信道是离散的, 其有限冲激响应可表示为 $(1, 0.1)$ 。如果接收序列是 $(1.0, -0.3, -0.7, 0, \dots)$ 。试用维特比算法来确定这个序列的最大似然译码的结果。

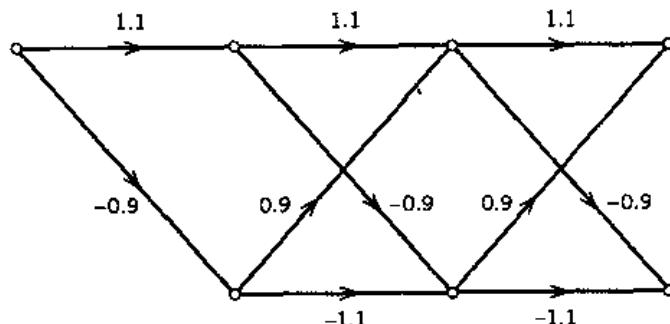


图 P10.27

10.28 图 P10.28 显示了一个 32-QAM 的交叉星座图。试将这个星座图分割为 8 个子集。要求在每一步分割中, 指明子集内部的(最小)欧氏距离。

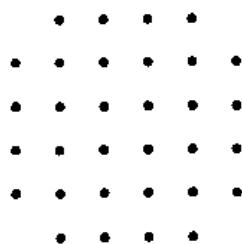


图 P10.28

10.29 由本章的引言可知,信道编码能够在给定的误码性能下降低系统对 E_b/N_0 的要求,或是在给定的 E_b/N_0 下减小接收天线的尺寸。回顾第 8 章例 8.2。现在假设下行链路中采用的是码率为 1/2 的卷积编码器,约束长度 $K=7$ 的维特比译码。如采用软量化,这个方案的编码增益为 5.1 dB。试完成:

- 重新计算系统所需的 E_b/N_0 。
- 假设系统要求的 E_b/N_0 不变,计算由于在下行链路中采用了这种编码方案,可使接收的碟形天线的尺寸减小多少。

10.30 与本章讨论的卷积码不同,第 6 章中用于 V.32 音频调制解调器中的卷积码是非线性的。图 P10.30 就是用在这个调制解调器中的卷积码的电路图。其中,除了模 2 加法器和延迟器之外,还采用了模 2 乘法器和门电路。试解释图 P10.30 中编码器非线性的原因,并用实例来证明你的解释。

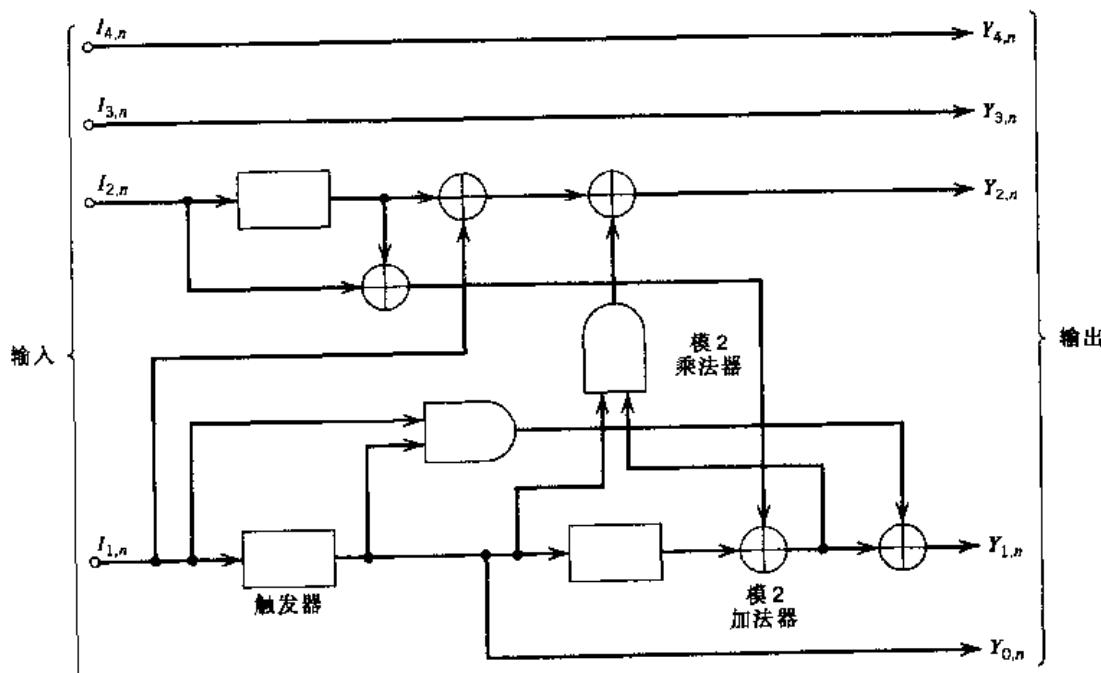


图 P10.30

Turbo 码

10.31 令 $r_c^{(1)} = p/q_1$ 和 $r_c^{(2)} = p/q_2$ 分别为图 10.25 所示 Turbo 编码器中的 RSC 编码器 1 和编码器 2 的码率。试计算 Turbo 码的码率。

10.32 图 10.25 所示的 Turbo 编码器中组成码的反馈特性有如下含义:一个比特的误码对应于一个无限的信道误码序列。试用一个由符号 1 和无穷多个符号 0 组成的信息序列来说明这个现象。

10.33 设码率为 1/2 的 Turbo 码的生成矩阵分别为

$$\text{四状态编码器: } g(D) = \left[1, \frac{1+D+D^2}{1+D^2} \right]$$

八状态编码器: $g(D) = \left[1, \frac{1 + D^2 + D^3}{1 + D + D^2 + D^3} \right]$

十六状态编码器: $g(D) = \left[1, \frac{1 + D^4}{1 + D + D^2 + D^3 + D^4} \right]$

(a) 试画出这些 RSC 编码器的方框图。

(b) 求出每个编码器对应的奇偶校验等式。

10.34 图 10.25 所示的 Turbo 编码器中采用了两个 RSC 编码器。

(a) 扩展此编码器,使之包含 M 个交织器。

(b) 画出 Turbo 译码器的方框图,要求译码器采用由扩展生成的 M 组奇偶校验比特。

10.35 Turbo 译码需要依靠外部信息的反馈。在 Turbo 译码器中采用的基本原则是避免反馈某一译码段自身生成的译码段信息。试从概念上解释这个原则的正确性。

10.36 假设一个通信接收机由两部分组成,一个解调器和一个译码器。解调器基于由调制器和信道组成的马尔可夫模型,而译码器基于前向纠错码的马尔可夫模型。试讨论 Turbo 原理如何为此系统构造一个联合解调器和译码器。

计算机实验

10.37 本实验将评估分组长度对译码器噪声性能的影响,以继续研究 10.9 节给出的 Turbo 码。

和前面一样,Turbo 编码器的两个卷积码编码器是:

编码器 1: [1, 1, 1]

编码器 2: [1, 0, 1]

传输 E_b/N_0 是 1 dB, 导致终止的分组误码数不超过 15。

试在这些背景条件下,对两个不同分组(交织器)长度 200 和 400,给出 Turbo 译码器的误比特率与迭代次数的关系图。

附录 1 概 率 论

A1.1 概率的概念

概率论起源于某种现象,通过实验我们可以直接或间接地建立起该现象的模型。其中,实验的结果取决于偶然性。并且,如果实验重复进行,受随机现象或偶然机制的影响,结果也会不同。这样的实验就称为随机实验。例如,抛掷硬币得到的观察结果就属于这种情况。在这抛掷硬币的实验中,可能出现的结果为“正面”或“反面”。

随机实验具有三个基本特征:

1. 实验可以在相同的条件下重复进行。
2. 实验中任何一次的结果都是不可预测的。
3. 对于实验中大量的尝试,结果呈现出一定的统计规律性。也就是说,如果实验重复进行很多次,就可以观测到一个确定的实验结果的平均图样。

A1.1.1 相对频率近似

令事件 A 代表随机实验的一个可能结果。例如在抛硬币实验中,用事件 A 代表“正面”。假设在 n 次实验后,事件 A 发生了 $N_n(A)$ 次。可将比值 $N_n(A)/n$ 赋给事件 A ,这个比值就称为事件 A 的相对频率。显然,相对频率是一个小于等于 1 的非负实数。即

$$0 \leq \frac{N_n(A)}{n} \leq 1 \quad (\text{A1.1})$$

如果事件 A 在实验中没有发生,则 $N_n(A)/n = 0$ 。另一方面,如果在 n 次实验中事件 A 都发生,则 $N_n(A)/n = 1$ 。

如果 n 次实验的次序可以任意排列,且当 n 很大时,相对频率 $N_n(A)/n$ 收敛于同一极限,则称这个实验具有统计规律性。相应地,可定义事件 A 的概率为

$$P(A) = \lim_{n \rightarrow \infty} \left(\frac{N_n(A)}{n} \right) \quad (\text{A1.2})$$

式(A1.2)所示的极限不应仅从数学的角度来看待,而应看成是一个关于事件概率的论述,即一个事件的概率是指在长期的实验中,事件 A 发生的次数的长期比例关系。例如,在抛掷硬币的实验中,我们认为在大量的抛掷中,大约有一半是正面朝上。

一个事件的概率代表了在实验的各次尝试中该事件发生的可能性。在许多工程应用和机会性的游戏中,用式(A1.2)来定义事件的概率是合理的。但是,在其他的许多应用场合,这样的定义是不够的。例如,要进行股票市场的统计分析,应怎样获得实验的可重复性呢?一个比较令人满意的办法是,先提出那些我们需要的概率的度量,将其视为公理,然后再用相对频率来证明这些公理。

A1.1.2 概率的公理

我们很容易想到将实验看成是一个定义空间，而将实验的可能结果看成是定义空间中的点。实验中第 k 个结果对应于定义空间中的一个点，这个点称为样本点，用 s_k 表示。对应于所有可能实验结果的全部样本点的总和称为样本空间，用 S 表示。一个事件可以对应一个或一组样本点。特别地，整个样本空间 S 称为必然事件，空集 \emptyset 称为无效或不可能事件，单个的样本点称为基本事件。

例如一个掷骰子的实验。在这个实验中，有六个可能出现的结果：1, 2, 3, 4, 5 或 6 点朝上。为这些可能的结果分配样本点，就可以得到一个一维的样本空间，这个样本空间由六个点组成，如图 A1.1 所示。结果为点数 6 的基本事件对应于样本点 {6}。另一方面，结果为偶数的基本事件对应于样本空间的一个子集 {2, 4, 6}。注意，事件一词可用来描述子集。

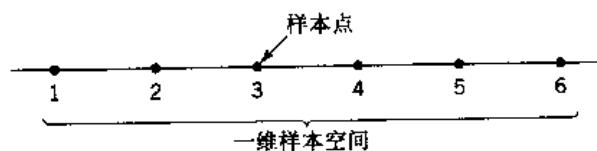


图 A1.1 掷骰子实验的样本空间

现在就可以给出概率的正式定义了。一个概率系统由三部分组成：

1. 一个基本事件(结果)构成的样本空间 S 。
2. 由 S 的子集中的事件组成的类 \mathcal{E} 。
3. 类 \mathcal{E} 中的每个事件 A 都对应于一个概率度量 $P(\cdot)$ 。概率量度 $P(\cdot)$ 具有如下性质：

$$(i) \quad P(S) = 1 \quad (A1.3)$$

$$(ii) \quad 0 \leq P(A) \leq 1 \quad (A1.4)$$

(iii) 如果 $A + B$ 是类 \mathcal{E} 中两个互斥的事件的联合，那么

$$P(A + B) = P(A) + P(B) \quad (A1.5)$$

性质(i), (ii) 和 (iii) 称为概率的公理。公理(i)说明必然事件的概率为 1。公理(ii)说明事件的概率是小于或等于 1 的非负实数。公理(iii)说明两个互斥事件的联合概率等于两个单独事件概率的和。这三个公理用来处理有限样本空间的实验已经足够。

虽然这种概率论的公理化方法是在事实的基础上抽象出来的，但三个公理都有各自的相对频率解释。公理(ii)对应于式(A1.1)。公理(i)对应于 n 次实验中每次都发生事件 A 的式(A1.1)的极限情况。下面解释公理(iii)。如果 n 次实验中事件 A 发生的次数为 $N_n(A)$ ，事件 B 发生的次数为 $N_n(B)$ ，则联合事件“ A 或 B ”发生了 $N_n(A) + N_n(B)$ 次 (A 和 B 不可能在同一次实验中发生)。因此， $N_n(A + B) = N_n(A) + N_n(B)$ ，于是有

$$\frac{N_n(A + B)}{n} = \frac{N_n(A)}{n} + \frac{N_n(B)}{n}$$

其数学形式与公理(iii)相似。

公理(i), (ii) 和 (iii) 组成了概率的隐含定义。我们可以利用这些公理推出概率的其他基本性质。

性质 1

$$P(\bar{A}) = 1 - P(A) \quad (\text{A1.6})$$

其中, \bar{A} (代表“非 A ”)是事件 A 的补集。

这个性质有助于研究未发生的事件。为了证明这个性质, 将样本空间 S 用两个互斥事件 A 和 \bar{A} 的联合来表示:

$$S = A + \bar{A}$$

然后, 利用公理(i)和(iii)可得

$$1 = P(A) + P(\bar{A})$$

由此, 可直接得出式(A1.6)。

性质 2

如果 M 个互斥的事件 A_1, A_2, \dots, A_M 具有完备性:

$$A_1 + A_2 + \dots + A_M = S \quad (\text{A1.7})$$

那么

$$P(A_1) + P(A_2) + \dots + P(A_M) = 1 \quad (\text{A1.8})$$

为了证明这个性质, 首先在式(A1.7)中利用公理(i):

$$P(A_1 + A_2 + \dots + A_M) = 1$$

接下来, 将公理(iii)归纳为

$$P(A_1 + A_2 + \dots + A_M) = P(A_1) + P(A_2) + \dots + P(A_M)$$

因此, 式(A1.8)成立。当 M 个事件具有相等的可能性(即相同的发生概率)时, 式(A1.8)可简化为

$$P(A_i) = \frac{1}{M}, \quad i = 1, 2, \dots, M$$

性质 3

如果事件 A 和 B 不是互斥的, 则联合事件“ A 或 B ”的概率为

$$P(A + B) = P(A) + P(B) - P(AB) \quad (\text{A1.9})$$

其中, $P(AB)$ 是相交事件“ A 与 B ”的概率。

概率 $P(AB)$ 称为联合概率。其相对频率解释如下:

$$P(AB) = \lim_{n \rightarrow \infty} \left(\frac{N_n(AB)}{n} \right) \quad (\text{A1.10})$$

其中, $N_n(AB)$ 是 n 次实验中 A 和 B 同时发生的次数。公理(iii)是式(A1.9)的特例。当 A 和 B 互斥时, $P(AB) = 0$, 式(A1.9)简化为与式(A1.5)完全相同的形式。

A1.1.3 条件概率

设实验中包含一对事件 A 和 B 。令 $P(B|A)$ 代表事件 A 已发生的条件下, B 发生的概

率。概率 $P(B|A)$ 称为给定条件 A 下 B 的条件概率。假设 A 具有非零概率，则条件概率可定义为

$$P(B|A) = \frac{P(AB)}{P(A)} \quad (\text{A1.11})$$

其中， $P(AB)$ 是 A 和 B 的联合概率。

可用式(A1.11)的相对频率解释来证明上式所定义的条件概率。假设进行一个实验并检验一对事件 A 和 B 的发生情况。令 $N_n(AB)$ 代表在 n 次实验中相交事件 AB 发生的次数。并假设在 n 次实验中事件 A 发生了 $N_n(A)$ 次。由于相交事件 AB 代表 A 和 B 都发生，所以 $N_n(A)$ 必然包含在 $N_n(AB)$ 之内。换言之，有

$$\frac{N_n(AB)}{N_n(A)} \leq 1$$

比值 $N_n(AB)/N_n(A)$ 表示 A 发生的情况下 B 发生的相对频率。对于很大的 n ，这个比值就等于条件概率 $P(B|A)$ 。即

$$P(B|A) = \lim_{n \rightarrow \infty} \left(\frac{N_n(AB)}{N_n(A)} \right)$$

等价地，有

$$P(B|A) = \lim_{n \rightarrow \infty} \left(\frac{N_n(AB)/n}{N_n(A)/n} \right)$$

且

$$P(AB) = \lim_{n \rightarrow \infty} \left(\frac{N_n(AB)}{n} \right)$$

和

$$P(A) = \lim_{n \rightarrow \infty} \left(\frac{N_n(A)}{n} \right)$$

所以，式(A1.11)的结果是成立的。

我们可将式(A1.11)写为

$$P(AB) = P(B|A)P(A) \quad (\text{A1.12})$$

很明显，上式还可以写为

$$P(AB) = P(A|B)P(B) \quad (\text{A1.13})$$

由此可得出结论：两个事件的联合概率可以表示成其中某一事件给定的情况下，另一事件的条件概率与这个事件的基本概率的乘积。值得注意的是，条件概率 $P(A|B)$ 和 $P(B|A)$ 与前面定义的各种概率具有相同的性质。

在已知条件概率 $P(A|B)$ 以及概率 $P(A)$ 和 $P(B)$ 时，可以很容易地求得条件概率 $P(B|A)$ 。由式(A1.12)和式(A1.13)，如果 $P(A) \neq 0$ ，则

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \quad (\text{A1.14})$$

这个关系式是贝叶斯公式的特殊形式。

假设条件概率 $P(B|A)$ 仅等于事件 B 的基本概率, 即

$$P(B|A) = P(B)$$

在这个条件下, 相交事件 AB 的发生概率就等于事件 A 和 B 的基本概率的乘积:

$$P(AB) = P(A)P(B)$$

于是有

$$P(A|B) = P(A)$$

也就是说, 假设事件 B 发生的情况下, 事件 A 发生的概率等于事件 A 的基本概率。所以, 了解某个事件的情况并不能增加对另一事件发生概率的了解。满足上述条件的事件 A 和事件 B 称为是相互统计独立的。

A1.2 随机变量

使用了样本空间的概念后, 可将实验结果视为一个可以在一组样本点周围漂移的变量。变量的值由实验决定。我们将一个定义域在样本空间、值域为实数组的函数称为实验的随机变量。但是, 随机变量这个词有时会引起混淆。首先, 随机一词不是用在等概意义上的。其次, 变量一词的本义是与实验结果无关的。但随机变量一词已经在有关概率的文献中被广泛使用, 因此沿用至今。

当实验结果为 s 时, 随机变量可用 $X(s)$ 或 X 来表示。例如, 掷骰子的结果所对应的是一组由六个样本点组成的样本空间, 这些样本点可用整数 $1, 2, \dots, 6$ 来表示。如果认为样本点 k 等同于掷骰子时得到 k 个点, 那么函数 $X(k) = k$ 就是一个随机变量, 这样 $X(k)$ 就不等于掷骰子时得到的点数。在这个例子中, 随机变量取一组离散的值, 因此称为离散随机变量。更准确地讲, X 只能在有限的观测区域内取有限个值。如果随机变量 X 可以在整个观测区域内取任意的值, 则称这个随机变量 X 为连续随机变量。例如, 当用随机变量表示某一特定瞬间的噪声电压的幅度时, 它就是一个连续随机变量, 因为它可以取 $-\infty$ 到 $+\infty$ 之间的任意值。

进一步, 我们需要一个随机变量的统计描述, 该统计描述对离散随机变量和连续随机变量是等效的。假设有随机变量 X 和事件 $X \leq x$ 的概率, 用 $P(X \leq x)$ 表示事件 $X \leq x$ 的概率。显然, 这个概率是虚变量 x 的函数。为了简化表示, 令

$$F_X(x) = P(X \leq x) \tag{A1.15}$$

函数 $F_X(x)$ 称为随机变量 X 的累积分布函数(cdf), 或简称为分布函数。注意, 这里的 $F_X(x)$ 是 x 的函数, 而不是 X 的函数, 但它与随机变量 X 的分布有关, 这一点用下标来表明。对于任意一点 x , 分布函数 $F_X(x)$ 仅表示一个概率。

由式(A1.15)可直接得到分布函数 $F_X(x)$ 的性质:

1. 分布函数 $F_X(x)$ 是有界的, 它分布在 0 和 1 之间;
2. 分布函数 $F_X(x)$ 是一个 x 的非减函数(单调上升)。即

$$F_X(x_1) \leq F_X(x_2) \quad x_1 < x_2 \tag{A1.16}$$

随机变量概率的另一种表示方法也是常常用到的。这就是分布函数的导数

$$f_X(x) = \frac{d}{dx} F_X(x) \quad (\text{A1.17})$$

称为随机变量的概率密度函数(pdf)。注意,式(A1.17)中的微分是关于虚变量 x 的。密度函数这个词起源于事件 $x_1 < X \leq x_2$ 的概率

$$\begin{aligned} P(x_1 < X \leq x_2) &= P(X \leq x_2) - P(X \leq x_1) \\ &= F_X(x_2) - F_X(x_1) \\ &= \int_{x_1}^{x_2} f_X(x) dx \end{aligned} \quad (\text{A1.18})$$

因此,一个区间上的概率就等于这个区间内概率密度函数下的面积。在式(A1.18)中,令 $x_1 = -\infty$,并做适当改变,就可以得到由概率密度函数定义的分布函数:

$$F_X(x) = \int_{-\infty}^x f_X(\xi) d\xi \quad (\text{A1.19})$$

由于 $F_X(\infty) = 1$ 对应于必然事件的概率,且 $F_X(-\infty) = 0$ 对应于不可能事件的概率,则由式(A1.18)有

$$\int_{-\infty}^{\infty} f_X(x) dx = 1 \quad (\text{A1.20})$$

前面我们已提到,一个分布函数必须是单调递增的。也就是说,分布函数的导数(或概率密度函数)必须是非负的。因此,我们可以称概率密度函数是一个总面积为 1 的非负函数。

到目前为止,我们只研究了单个随机变量的情况。但实际上为了描述一个实验结果,常常需要几个随机变量。下面研究两个随机变量的情况。以这种方式得到的概率表示方法,可以很容易地推广到任意个随机变量的情况。

设有随机变量 X 和 Y 。定义联合分布函数 $F_{X,Y}(x, y)$ 为随机变量 $X \leq x$ 且随机变量 $Y \leq y$ 的概率,其中 x 和 y 分别为两个指定的值。变量 X 和 Y 可以是两个单独的一维随机变量,也可以是一个二维随机变量。无论是哪种情况,联合样本空间都是 xy 平面。联合概率分布函数 $F_{X,Y}(x, y)$ 是实验结果使样本点位于样本空间的某一区域($-\infty < X \leq x, -\infty < Y \leq y$)的概率。即

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y) \quad (\text{A1.21})$$

设联合分布函数 $F_{X,Y}(x, y)$ 处处连续,则偏导数

$$f_{X,Y}(x, y) = \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y} \quad (\text{A1.22})$$

存在且处处连续。 $f_{X,Y}(x, y)$ 称为随机变量 X 和 Y 的联合概率密度函数。联合分布函数 $F_{X,Y}(x, y)$ 对于 x 和 y 都是单调递增的。因此,由式(A1.22)可推导出联合概率密度函数 $f_{X,Y}(x, y)$ 总是非负的。同样,在联合概率密度函数曲面下的柱体的总体积必须为 1,即

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(\xi, \eta) d\xi d\eta = 1 \quad (\text{A1.23})$$

单个随机变量(X)的概率密度函数可以由对应的带有第二个随机变量(Y)的联合概率密

度函数求得。首先有

$$F_X(x) = \int_{-\infty}^{\infty} \int_{-\infty}^x f_{X,Y}(\xi, \eta) d\xi d\eta \quad (\text{A1.24})$$

在式(A1.24)两端对 x 求微分, 可得

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, \eta) d\eta \quad (\text{A1.25})$$

这样, 概率密度函数 $f_X(x)$ 就可以由联合概率密度函数 $f_{X,Y}(x, y)$ 对另一个随机变量 Y 的所有可能值求简单积分得到。同理, 对随机变量 X 进行相同的运算可以得到 $f_Y(y)$ 。概率密度函数 $f_X(x)$ 和 $f_Y(y)$ 称为边缘密度。因此, 联合概率密度函数 $f_{X,Y}(x, y)$ 中包含了关于联合随机变量 X 和 Y 的所有可能信息。

设连续随机变量 X 和 Y 具有联合概率密度函数 $f_{X,Y}(x, y)$ 。在给定 $X = x$ 的条件下, Y 的条件概率密度函数可定义为

$$f_Y(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)} \quad (\text{A1.26})$$

设 $f_X(x) > 0$, 其中, $f_X(x)$ 是 X 的边缘密度。函数 $f_Y(y|x)$ 可以看成是变量 y 的一个函数, 其中, 变量 x 是任意的, 但必须是固定的。因此, 函数 $f_Y(y|x)$ 满足普通概率函数的所有要求, 即

$$f_Y(y|x) \geq 0$$

和

$$\int_{-\infty}^{\infty} f_Y(y|x) dy = 1$$

如果随机变量 X 和 Y 是统计独立的, 那么, 对 X 的结果的了解不会影响 Y 的分布。于是条件概率密度函数 $f_Y(y|x)$ 就等于边缘密度 $f_Y(y)$, 即

$$f_Y(y|x) = f_Y(y)$$

在这种情况下, 可将随机变量 X 和 Y 的联合概率分布密度函数表示为各自边缘密度的乘积:

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

换言之, 如果随机变量 X 和 Y 的联合概率密度函数等于它们的边缘密度的乘积, 那么, 就可以认为 X 和 Y 是相互统计独立的。

A1.3 统计平均

讨论完概率和一些相关问题后, 下面就给出确定随机实验结果均值的办法。随机变量的数学期望或均值的定义为

$$\mu_X = E[X] = \int_{-\infty}^{\infty} xf_X(x) dx \quad (\text{A1.27})$$

E 代表统计期望算子。也就是说, 均值 μ_X 位于随机变量 X 的概率曲线以下面积的重心处。为了解释数学期望值 μ_X 的含义, 将定义式(A1.27)的积分改写成如下形式的近似求和的极限。

令 $|x_k| k = 0, \pm 1, \pm 2, \dots$ 代表实轴上一组间隔均匀的点:

$$x_k = \left(k + \frac{1}{2} \right) \Delta, \quad k = 0, \pm 1, \pm 2, \dots \quad (\text{A1.28})$$

其中, Δ 是相邻两点间的间隔。这样, 我们就可以将式(A1.27)写为和的极限形式:

$$\begin{aligned} E[X] &= \lim_{\Delta \rightarrow 0} \sum_{k=-\infty}^{\infty} \int_{k\Delta}^{(k+1)\Delta} x_k f_X(x) dx \\ &= \lim_{\Delta \rightarrow 0} \sum_{k=-\infty}^{\infty} x_k P\left(x_k - \frac{\Delta}{2} < X \leq x_k + \frac{\Delta}{2}\right) \end{aligned} \quad (\text{A1.29})$$

为了给出式(A1.29)中等式右边的求和的物理解释, 假设对随机变量 X 进行了 n 次独立的观测。令 $N_n(k)$ 代表随机变量 X 落入第 k 个区域内的次数:

$$x_k - \frac{\Delta}{2} < X \leq x_k + \frac{\Delta}{2}, \quad k = 0, \pm 1, \pm 2, \dots$$

于是, 当观测次数 n 很大时, 比值 $N_n(k)/n$ 趋近于概率 $P(x_k - \Delta/2 < X \leq x_k + \Delta/2)$ 。因此, 可近似地求得随机变量 X 的数学期望:

$$\begin{aligned} E[X] &\approx \sum_{k=-\infty}^{\infty} x_k \left(\frac{N_n(k)}{n} \right) \\ &= \frac{1}{n} \sum_{k=-\infty}^{\infty} x_k N_n(k), \quad \text{当 } n \text{ 很大时} \end{aligned} \quad (\text{A1.30})$$

式(A1.30)中等式右边的值就是样本平均值。对所有的 x_k 值求和, 其中每个 x_k 都以对应的发生的次数加权。实际上, 式(A1.30)提供了一种计算数学期望值 $E[X]$ 的基本方法。

下面我们来考虑更一般的情况。令 X 代表一个随机变量, $g(X)$ 代表定义在实轴上的 X 的一个函数。如果令 $g(X)$ 的自变量为一个随机变量, 则

$$Y = g(X) \quad (\text{A1.31})$$

也是随机变量。要得到随机变量 Y 的数学期望值, 可先求出概率密度函数 $f_Y(y)$, 然后再用标准公式

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy$$

然而, 更简单的方法是

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx \quad (\text{A1.32})$$

实际上, 式(A1.32)可以看成是随机变量 X 的任意函数 $g(X)$ 的数学期望概念的推广。

A1.3.1 矩

对于 $g(X) = X^n$ 的特殊情况, 由式(A1.32), 可得到随机变量 X 的概率分布的 n 阶矩, 即

$$E[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx \quad (\text{A1.33})$$

显然,对于 X 来说,最重要的矩就是最前面的两阶矩。在式(A1.33)中令 $n=1$,可得到式(A1.27)所示的随机变量的均值。而当 $n=2$ 时, X 的均方值为

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx \quad (A1.34)$$

我们还可以定义中心矩,即随机变量 X 与其均值之差 μ_X 的矩。也就是 n 阶中心矩

$$E[(X - \mu_X)^n] = \int_{-\infty}^{\infty} (x - \mu_X)^n f_X(x) dx \quad (A1.35)$$

对于 $n=1$,中心矩显然为 0。 $n=2$ 时,二阶矩被看成是随机变量 X 的方差,记为

$$\text{var}[X] = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx \quad (A1.36)$$

随机变量 X 的方差一般用 σ_X^2 来表示。其平方根称为随机变量 X 的标准偏差,记为 σ_X 。

随机变量的方差 σ_X^2 在某种意义上是变量的随机性的度量。确定了方差 σ_X^2 后,基本上就能够限制随机变量 X 的概率密度函数 $f_X(x)$ 在均值左右的有效宽度 μ_X 。切比雪夫(Chebyshev)对此给出了精确描述。切比雪夫不等式规定:对任意的正整数 ϵ ,有

$$P(|X - \mu_X| \geq \epsilon) \leq \frac{\sigma_X^2}{\epsilon^2} \quad (A1.37)$$

从这个不等式可以看出,随机变量的均值和方差决定了概率分布的部分特征。因此,均值和方差通常可实际应用。

由式(A1.34)和式(A1.36),不难发现方差 σ_X^2 和均方值 $E[X^2]$ 有如下的关系:

$$\begin{aligned} \sigma_X^2 &= E[X^2 - 2\mu_X X + \mu_X^2] \\ &= E[X^2] - 2\mu_X E[X] + \mu_X^2 \\ &= E[X^2] - \mu_X^2 \end{aligned} \quad (A1.38)$$

上式第二行中应用了统计期望算子的线性特性。式(A1.38)说明,如果 μ_X 均值为零,那么 σ_X^2 ,随机变量 X 的方差以及 $E[X^2]$ 的均方值就相等。

A1.3.2 特征函数

另一个重要的统计平均量是随机变量 X 的概率分布的特征函数 $\phi_X(v)$ 。其定义为复指数函数的数学期望:

$$\begin{aligned} \phi_X(v) &= E[\exp(jvX)] \\ &= \int_{-\infty}^{\infty} f_X(x) \exp(jvx) dx \end{aligned} \quad (A1.39)$$

其中, v 是实数, $j=\sqrt{-1}$ 。换言之,特征函数 $\phi_X(v)$ 是概率密度函数 $f_X(x)$ 的傅里叶变换(除了指数的符号不同此外),傅里叶变换可参见附录 2。在这个关系式中,我们采用了 $\exp(jvx)$,而不是 $\exp(-jvx)$,是为了与概率论中的用法一致。值得注意的是,在这里 v 和 x 分别扮演了与傅里叶变换中的 $2\pi f$ 和 t 类似的作用。根据傅里叶反变换,我们可以推导出与之相反的关系式

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_X(v) \exp(-jvx) dv \quad (\text{A1.40})$$

这个关系式可以用于通过随机变量 X 的特征函数 $\phi_X(v)$ 来计算概率密度函数 $f_X(x)$ 。

A1.3.3 联合矩

假设有一对随机变量 X 和 Y 。在这种情况下, 很重要的一组统计平均量就是联合矩, 即 $X^i Y^k$ 的期望值。其中, i 和 k 为任意正整数。于是有

$$E[X^i Y^k] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^i y^k f_{X,Y}(x, y) dx dy \quad (\text{A1.41})$$

最重要的联合矩是由 $E[XY]$ 定义的相关值。 $E[XY]$ 对应于式(A1.41)中 $i = k = 1$ 的情况。

中心化的随机变量 $X - E[X]$ 和 $Y - E[Y]$ 的相关值, 即联合矩

$$\text{cov}[XY] = E[(X - E[X])(Y - E[Y])] \quad (\text{A1.42})$$

称为 X 和 Y 的协方差。令 $\mu_X = E[X]$, $\mu_Y = E[Y]$, 可将式(A1.42)展开为

$$\text{cov}[XY] = E[XY] - \mu_X \mu_Y \quad (\text{A1.43})$$

令 σ_X^2 和 σ_Y^2 分别代表 X 和 Y 的方差, 然后对 X 和 Y 的协方差关于 $\sigma_X \sigma_Y$ 进行标准化, 得到 X 和 Y 的相关系数

$$\rho = \frac{\text{cov}[XY]}{\sigma_X \sigma_Y} \quad (\text{A1.44})$$

当且仅当随机变量 X 和 Y 的协方差为零, 即

$$\text{cov}[XY] = 0$$

时, X 和 Y 不相关。当且仅当随机变量 X 和 Y 对应的

$$E[XY] = 0$$

时, X 和 Y 正交。

由式(A1.43)可知, 当随机变量 X 和 Y 相互正交, 且其中至少有一个均值为零时, X 和 Y 不相关, 反之亦然。如果 X 和 Y 统计独立, 则它们也不相关, 但反之不一定成立。

附录 2 信号与系统简述

A2.1 傅里叶分析

令 $g(t)$ 代表非周期确定性信号,且可表示为时间 t 的函数。根据定义,信号 $g(t)$ 的傅里叶变换可以由下面的积分给出:

$$G(f) = \int_{-\infty}^{\infty} g(t) \exp(-j2\pi ft) dt \quad (\text{A2.1})$$

其中, $j = \sqrt{-1}$, 变量 f 代表频率。如果给定了傅里叶变换 $G(f)$, 则原始信号 $g(t)$ 可用傅里叶反变换准确地得到:

$$g(t) = \int_{-\infty}^{\infty} G(f) \exp(j2\pi ft) df \quad (\text{A2.2})$$

注意,在式(A2.1)和式(A2.2)中,我们用小写字母表示时域函数,而用大写字母表示频域函数。函数 $G(f)$ 和 $g(t)$ 组成了傅里叶变换对。

为了使信号 $g(t)$ 的傅里叶变换存在,要求 $g(t)$ 满足三个充分但非必要的条件,这三个条件统称为狄利克雷(Dirichlet)条件:

1. 函数 $g(t)$ 是在任何有限时间区间上都存在有限个最大和最小值的单值函数。
2. 函数 $g(t)$ 在任何有限区间上存在有限个不连续点。
3. 函数 $g(t)$ 绝对可积,即

$$\int_{-\infty}^{\infty} |g(t)| dt < \infty$$

当时域函数 $g(t)$ 是一个物理可实现信号的特定准确描述时,就可以认为傅里叶变换是存在的。换言之,物理上的可实现对于傅里叶变换的存在性来说是一个充分条件。实际上,可以进一步认为,当所有的能量信号,即信号 $g(t)$ 满足

$$\int_{-\infty}^{\infty} |g(t)|^2 dt < \infty$$

时,其傅里叶变换是存在的。

傅里叶变换为信号频率成分(或频谱)的测量提供了数学工具。因此,傅里叶变换和频谱这两个词常常可以互换。这样,当给定一个信号 $g(t)$,其傅里叶变换为 $G(f)$,则可将 $G(f)$ 看成是信号 $g(t)$ 的频谱。同理,可将 $|G(f)|$ 看成是信号 $g(t)$ 的幅度谱,将 $\arg|G(f)|$ 看成是信号的相位谱。

A2.1.1 傅里叶变换的性质

时域函数 $g(t)$ 及其傅里叶变换 $G(f)$ 的关系,以及时域函数 $g(t)$ 的不同运算对频域函数 $G(f)$ 的影响,都可以通过考察傅里叶变换的某些性质来得到。表 A6.2 总结了这些性质。

A2.1.2 狄拉克 δ 函数

严格地讲,傅里叶变换的理论只能应用于满足狄利克雷条件的时域函数。能量信号就属于这类函数。但是,由于实际应用的需要,人们往往希望将傅里叶变换在以下两个方面得到扩展:

1. 将傅里叶级数与傅里叶变换包含于一个统一的理论。这样,就可以将傅里叶级数看成是傅里叶变换的一种特殊情况;
2. 将功率信号(即平均功率为有限值的信号)列入可以进行傅里叶变换的信号之列。

上述两个目标涉及到“正确使用”狄拉克(Dirac) δ 函数(或单位脉冲)的问题。

狄拉克 δ 函数(或 δ 函数)可以用 $\delta(t)$ 来表示,其定义为,函数在 $t=0$ 处幅度为无穷大,而在其余各处幅度均为零,并且函数曲线下的面积为1,即

$$\delta(t) = 0, \quad t \neq 0 \quad (\text{A2.3})$$

和

$$\int_{-\infty}^{\infty} \delta(t) dt = 1 \quad (\text{A2.4})$$

这两个关系式说明函数 $\delta(t)$ 必定是一个关于时间 t 的偶函数,其中心点在 $t=0$ 处。

但是,要使 δ 函数有意义,首先必须将其作为关于时间积分的被积函数的因子。其次,严格地讲,只有当被积函数的其他因子是时域连续的函数时, δ 函数才有意义。假设 $g(t)$ 就是这样一个时域连续函数,下面讨论 $g(t)$ 与时移的 δ 函数 $\delta(t-t_0)$ 的乘积。根据上面的两个定义式即式(A2.3)和式(A2.4),可以将这个乘积积分表示为

$$\int_{-\infty}^{\infty} g(t) \delta(t - t_0) dt = g(t_0) \quad (\text{A2.5})$$

等式左边所示的运算筛选出了函数 $g(t)$ 在 $t=t_0$ 处的值 $g(t_0)$,其中 $-\infty < t < \infty$ 。式(A2.5)可看成是 δ 函数的筛选特性。这个特性有时也作为 δ 函数的定义式。实际上,它是将式(A2.3)和式(A2.4)合并成了一个关系式。

注意到 $\delta(t)$ 函数是一个关于 t 的偶函数,我们可以重写式(A2.5),以突出该式与卷积积分的共同之处:

$$\int_{-\infty}^{\infty} g(\tau) \delta(t - \tau) d\tau = g(t) \quad (\text{A2.6})$$

也就是说,任何函数与 δ 函数的卷积都等于函数本身。这就称为 δ 函数的重复性。

值得注意的是,一般来讲,不可能找到同时具备式(A2.3)和式(A2.4)所示的两个特性,或等效地具备式(A2.5)所示筛选特性的函数。但是,可以想像有这样一系列的函数,它们在 $t=0$ 这一点的峰值越来越高,越来越细,同时函数曲线下的面积保持为1,并且在除 $t=0$ 以外的点上函数值趋近于0,且这些点一直延伸到无穷。也就是说,我们可以将脉冲持续时间趋近于0的单位面积脉冲的有限形式看成是 δ 函数。至于采用哪种形状的脉冲并不重要。

A2.1.3 周期信号的傅里叶变换

众所周知,通过傅里叶级数,一个周期信号可以用复指数的和来表示。同样,在某种意义

上,傅里叶变换也可以用复指数来定义。因此,如果变换中可以包含 δ 函数,则将周期信号的傅里叶变换表示成周期形式也是合理的。

假设有一个周期为 T_0 的周期信号 $g_{T_0}(t)$,根据复指数的傅里叶级数,我们可以将该周期信号 $g_{T_0}(t)$ 表示成

$$g_{T_0}(t) = \sum_{n=-\infty}^{\infty} c_n \exp(j2\pi n f_0 t) \quad (\text{A2.7})$$

其中, c_n 是复傅里叶系数,其定义为

$$c_n = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} g_{T_0}(t) \exp(-j2\pi n f_0 t) dt \quad (\text{A2.8})$$

f_0 为周期 T_0 的倒数的基频,即

$$f_0 = \frac{1}{T_0} \quad (\text{A2.9})$$

令 $g(t)$ 为一个脉冲形的函数,取值等于 $g_{T_0}(t)$ 的一个周期,且在其他各处为 0,即

$$g(t) = \begin{cases} g_{T_0}(t), & -\frac{T_0}{2} < t \leq \frac{T_0}{2} \\ 0 & \text{其他} \end{cases} \quad (\text{A2.10})$$

根据信号 $g(t)$,周期信号 $g_{T_0}(t)$ 可以表示成无穷和的形式

$$g_{T_0}(t) = \sum_{m=-\infty}^{\infty} g(t - mT_0) \quad (\text{A2.11})$$

在此基础上,可将 $g(t)$ 看成是一个生成函数,由它可以生成周期信号 $g_{T_0}(t)$ 。

函数 $g(t)$ 是可以进行傅里叶变换的。因此,可将复傅里叶系数的公式写为

$$\begin{aligned} c_n &= f_0 \int_{-\infty}^{\infty} g(t) \exp(-j2\pi n f_0 t) dt \\ &= f_0 G(n f_0) \end{aligned} \quad (\text{A2.12})$$

其中, $G(n f_0)$ 是 $g(t)$ 的傅里叶变换在频率 $n f_0$ 处的估计值。于是得到周期信号 $g_{T_0}(t)$ 的重构公式为

$$g_{T_0}(t) = f_0 \sum_{n=-\infty}^{\infty} G(n f_0) \exp(j2\pi n f_0 t) \quad (\text{A2.13})$$

或者根据式(A2.11)得到

$$\sum_{m=-\infty}^{\infty} g(t - mT_0) = f_0 \sum_{n=-\infty}^{\infty} G(n f_0) \exp(j2\pi n f_0 t) \quad (\text{A2.14})$$

式(A2.14)是泊松和式的一种形式。

观察函数 $g(t)$ 不难发现, $g(t)$ 是由周期函数 $g_{T_0}(t)$ 的一个周期构成的,并具有连续的频谱函数 $G(f)$,而周期函数 $g_{T_0}(t)$ 本身的频谱却是离散的。由此可以得出结论:时域的周期性可以将信号的频域形式(或频谱)变为定义在基频整数倍频上的离散形式。

A2.1.4 傅里叶变换对

表 A6.3 给出了常用傅里叶变换对的列表, 这些变换对可以由上述公式推得。

A2.1.5 信号通过线性系统传输

系统是指可以根据不同的输入信号产生输出信号的物理装置。习惯上, 将输入信号称为激励, 输出信号称为响应。在线性系统中, 叠加原理是有效的。也就是说, 线性系统对同时的多个激励的响应等于系统对各个激励的响应之和。

在时域中, 线性系统是根据冲激响应来描述的。冲激响应就是在系统的输入端送入单位脉冲或 δ 函数时系统的响应(系统的初始状态为 0)。如果系统是时不变的, 那么, 无论单位脉冲何时送入系统, 系统的冲激响应都是相同的。假设在 $t = 0$ 时向系统送入单位脉冲或 δ 函数, 我们可以用 $b(t)$ 来表示线性时不变系统的冲激响应。若系统的输入为任意的激励 $x(t)$, 则系统的响应 $y(t)$ 可由冲激响应 $b(t)$ 定义为

$$y(t) = \int_{-\infty}^{\infty} x(\tau)b(t - \tau) d\tau \quad (\text{A2.15})$$

这就是卷积积分。上式也可以记为

$$y(t) = \int_{-\infty}^{\infty} b(\tau)x(t - \tau) d\tau \quad (\text{A2.16})$$

因此说, 卷积是可交换的。

在卷积积分中, 要涉及到三个不同的时间范围: 激励时间 τ 、响应时间 t 和系统记忆时间 $t - \tau$ 。卷积积分是时域线性时不变系统的基础。由式(A2.15)可知, 线性时不变系统响应的当前值是输入信号的过去值的加权积分, 加权值由系统的冲激响应决定。因此, 冲激响应是一个系统的记忆函数。

A2.1.6 线性时不变系统的频率响应

假设向一个冲激响应为 $b(t)$ 的线性时不变系统输入一个幅度为 1、频率为 f 的复指数信号, 即

$$x(t) = \exp(j2\pi ft)$$

将这个激励用于式(A2.16), 可得系统的响应为

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} b(\tau) \exp[j2\pi f(t - \tau)] d\tau \\ &= \exp(j2\pi ft) \int_{-\infty}^{\infty} b(\tau) \exp(-j2\pi f\tau) d\tau \end{aligned} \quad (\text{A2.17})$$

定义系统的频率响应为系统冲激响应的傅里叶变换, 即

$$H(f) = \int_{-\infty}^{\infty} b(t) \exp(-j2\pi ft) dt \quad (\text{A2.18})$$

除了要将 τ 用 t 代替外, 上式与式(A2.17)最后一行的积分是完全相同的。因此, 式(A2.17)可写为

$$y(t) = H(f) \exp(j2\pi ft) \quad (\text{A2.19})$$

因此,线性时不变系统对频率为 f 的复指数函数的响应等于在该复指数函数上乘上一个常系数 $H(f)$ 。

通常,频率响应 $H(f)$ 是一个复数,可表示为

$$H(f) = |H(f)| \exp[j\beta(f)] \quad (\text{A2.20})$$

其中, $|H(f)|$ 为幅度响应, $\beta(f)$ 为相位响应。在线性系统冲激响应 $b(t)$ 为实数的特殊情况下,频率响应 $H(f)$ 为共轭对称,即

$$|H(f)| = |H(-f)|$$

且

$$\beta(f) = -\beta(-f)$$

也就是说,冲激响应为实数的线性系统的幅度响应 $|H(f)|$ 是频率的偶函数,而相位响应 $\beta(f)$ 是频率的奇函数。

在某些应用场合,人们更倾向于采用 $H(f)$ 对数的极坐标形式来进行表示,而不只是用 $H(f)$ 本身。定义 $H(f)$ 的自然对数为

$$\log H(f) = \alpha(f) + j\beta(f) \quad (\text{A2.21})$$

其中

$$\alpha(f) = \log |H(f)| \quad (\text{A2.22})$$

函数 $\alpha(f)$ 称为系统增益,其单位为奈培(neper)。 $\beta(f)$ 的单位为弧度。式(A2.21)说明增益 $\alpha(f)$ 和相位 $\beta(f)$ 分别是频率响应 $H(f)$ 的自然对数的实部和虚部。增益也可以用 dB 来表示:

$$\alpha'(f) = 20 \log_{10} |H(f)| \quad (\text{A2.23})$$

$\alpha(f)$ 和 $\alpha'(f)$ 两个增益函数之间的关系为

$$\alpha'(f) = 8.69\alpha(f) \quad (\text{A2.24})$$

其中,1 奈培等于 8.69 dB。

A2.2 带宽

信号的时域和频域表示为互逆的关系。特别地,我们可以得到如下重要结论:

- 当信号的时域形式发生变化时,其频域形式按相反的方向发生变化,反之亦然。这个可逆的关系可防止两个域内信号的随意指定。换言之,我们可以任意指定一个时域函数或一个频域函数,但不能将它们对应起来。
- 如果一个信号在频率上是严格受限的,那么,这个信号的时域描述就是无穷的,即使其幅度逐渐变小。如果在有限的频率带宽之外,信号的傅里叶变换精确等于 0,则称这个信号是严格频率受限或严格带宽受限的。sinc 脉冲

$$\text{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$$

就是严格带宽受限信号的例子。sinc 脉冲在时域上也是渐近受限的,这就证实了以上

对严格带宽受限信号的论述。反过来,如果一个信号是时域严格受限的(信号在有限的时间间隔外精确等于0),那么,信号的频谱是无限延伸的,即使其频谱幅度逐渐变小。矩形脉冲就是这样一个例子。因此,我们可以认为,信号不可能在频域或时域同时受限。

信号的带宽提供了正频率范围内信号主要频谱成分扩展的度量。当信号是严格带宽受限的,其带宽就可以准确定义。例如,sinc脉冲 $\text{sinc}(2\pi f_0 t)$ 的带宽为 W 。然而,在通常情况下,当信号不是严格受限时,在定义带宽时就会遇到困难。这是由于附加在信号频谱成分上的“主要”一词,其数学含义是不准确的,因此不存在一个普遍接受的带宽定义。但是,我们仍然能够给出一些常用的带宽定义。

当信号的频谱关于以明确定义的零点为边界的主瓣(即频谱为零的频率点)对称时,可以用主瓣作为信号带宽定义的基础。特别地,如果信号是低通的(其频谱成分集中于原点),则其带宽被定义为主瓣总宽度的一半,这是因为主瓣频谱中只有一半位于正频率范围内。例如,持续时间为 T 秒的矩形脉冲的集中于原点的主瓣频谱宽度为 $2/T$ 赫。因此,我们可以定义这个矩形脉冲的带宽为 $1/T$ 赫。另一方面,如果信号是带通的,其主瓣频谱集中在 $\pm f_c$ 附近(假设 f_c 足够大),此时,带宽就定义为正频率的主瓣频谱宽度。这样的带宽定义称为零-零(null-to-null)带宽。例如,一个持续时间为 T 秒的频率为 f_c 的射频脉冲,具有集中在 $\pm f_c$ 附近的宽度为 $2/T$ 赫的主瓣,这里假设 f_c 大于 $1/T$ 。因此,我们可以定义射频脉冲的零-零带宽为 $2/T$ 赫。在这个定义的基础上,我们可以认为,将低通信号的频谱成分平移足够大的频率值,就可以使信号的带宽加倍。这个频率转换可以通过调制来实现。

另一个通用的带宽定义是 3 dB 带宽。特别地,如果信号是低通的,3 dB 带宽是指频谱幅度为峰值的零频率点与频谱幅度下降为峰值的正频率点之间的间隔 $1/\sqrt{2}$ 。例如,指数衰减的信号 $\exp(-at)$ 的 3 dB 带宽为 $a/2\pi$ 赫。另一方面,如果信号是带通的,其中心在 $\pm f_c$,则 3 dB 带宽是指在正频率轴上频谱幅度降为 f_c 峰值的两个频率点之间的间隔 $1/\sqrt{2}$ 。3 dB 带宽的优点是可以直接在频谱图上读出,缺点在于如果幅度谱有缓慢下降的尾巴时会引起误导。

此外,还有一种信号带宽的度量,即均方根(rms)带宽,其定义为适当信号平方幅度谱的归一化形式的二阶矩的平方根。假设信号是低通的,则二阶矩可以在原点附近得到。为了求得幅度谱平方的归一化形式,我们可以用到非负函数

$$\frac{|G(f)|^2}{\int_{-\infty}^{\infty} |G(f)|^2 df}$$

其中,分母使用正确的标准化形式,即此比值在整个频率轴上的积分值等于 1。这样就可以用傅里叶变换 $G(f)$ 来定义低通信号 $g(t)$ 的 rms 带宽为

$$W_{\text{rms}} = \left(\frac{\int_{-\infty}^{\infty} f^2 |G(f)|^2 df}{\int_{-\infty}^{\infty} |G(f)|^2 df} \right)^{1/2} \quad (\text{A2.25})$$

均方根带宽 W_{rms} 最具吸引力的特征是,相对于其他两种带宽的定义,可以更方便地进行数学估

算。但是,均方根带宽在实验中的可测量性不强。

A2.2.1 时间-带宽积

对于脉冲信号的任何分类,其时域因子各不相同。但是,信号的持续时间与其带宽的乘积却总是一个常数,即

$$(持续时间 \times 带宽) = \text{常数}$$

这个乘积称为时间-带宽积或带宽持续时间积。时间-带宽积的恒定性是一个信号的时域和频域形式之间互逆关系的另一种表示形式。特别地,如果脉冲信号的持续时间以因子 a 随时间范围的减小而减少,则根据傅里叶变换的时域缩放特性,信号的频谱范围和信号的带宽以相同的因子 a 增加,从而信号的时间-带宽积为一恒定的常数,见表 A6.2 的第二条。例如,持续时间为 T 秒的矩形脉冲具有 $1/T$ 赫的带宽(以主瓣的正频率部分为基础定义),于是脉冲的时间-带宽积等于 1。无论我们选用信号带宽的哪种定义,对应的脉冲信号的时间-带宽积都是常数。选用不同的定义只会改变这个常数的值。

为了更清楚地说明这一点,我们来讨论式(A2.25)所示的 rms 带宽。信号 $g(t)$ 的相应 rms 持续时间为

$$T_{\text{rms}} = \left(\frac{\int_{-\infty}^{\infty} t^2 |g(t)|^2 dt}{\int_{-\infty}^{\infty} |g(t)|^2 dt} \right)^{1/2} \quad (\text{A2.26})$$

其中,假设信号 $g(t)$ 集中在原点附近。由式(A2.25)和式(A2.26),可得时间-带宽积为

$$T_{\text{rms}} W_{\text{rms}} \geq \frac{1}{4\pi} \quad (\text{A2.27})$$

其中,常数为 $1/4\pi$ 。当取高斯脉冲 $\exp(-\pi t^2)$ 时,上式中等号成立。

A2.2.2 噪声等效带宽

上面给出的关于信号带宽的定义(即 3 dB 带宽、零-零带宽和 rms 带宽)都是基于确定性信号的。此外,还有一种基于随机信号和系统的带宽定义,即噪声等效带宽。假设有一功率谱密度为 $N_0/2$ 的白噪声源,将其输入图 A2.1 所示的简单 RC 低通滤波器,则相应的平均输出噪声功率的值等于 $N_0/(4RC)$ 。这个滤波器的 3 dB 带宽为 $1/2\pi RC$ 。其中,滤波器的平均输出噪声功率与带宽成正比。

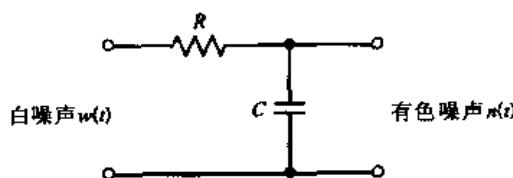


图 A2.1 RC 低通滤波器

通过下面定义的噪声等效带宽,我们可以概括这一结论,使之适用于所有种类的低通滤波

器。假设有一均值为零、功率谱密度为 $N_0/2$ 的白噪声源, 将其输入任意一个转移函数为 $H(f)$ 的低通滤波器。因此, 得到平均输出噪声功率为

$$\begin{aligned} N_{\text{out}} &= \frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 df \\ &= N_0 \int_0^{\infty} |H(f)|^2 df \end{aligned} \quad (\text{A2.28})$$

其中, 最后一行中用到了幅度响应 $|H(f)|$ 是关于频率的偶函数这一性质。

下面再考虑同样的白噪声源连接到零频响应为 $H(0)$ 、带宽为 B 的理想低通滤波器的情况。此时, 平均输出噪声功率为

$$N_{\text{out}} = N_0 B H^2(0) \quad (\text{A2.29})$$

这个平均输出噪声功率等于式(A2.28), 于是, 我们可以正式定义噪声等效带宽为

$$B = \frac{\int_0^{\infty} |H(f)|^2 df}{H^2(0)} \quad (\text{A2.30})$$

因此, 计算噪声等效带宽的过程就是用等效的零频响应为 $H(0)$ 、带宽为 B 的理想低通滤波来代替任意的转移函数为 $H(f)$ 的低通滤波器的过程, 如图 A2.2 所示。用同样的方法, 我们可以定义带通滤波器的噪声等效带宽。

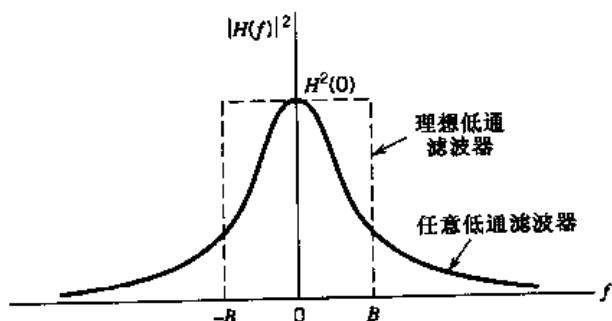


图 A2.2 低通滤波器的噪声等效带宽定义的说明

A2.3 希尔伯特变换

傅里叶变换对于估算能量信号或功率有限信号的频率成分特别有用。因此, 傅里叶变换为分析和设计选频滤波器(用于在频域上分离信号)提供了数学基础。另一种分离信号的方法是基于相位选择的。其中, 在相干信号间采用相位偏移, 以获得所需的信号分离。最简单的相位偏移是 180° 的相移, 在正弦信号的情况下, 这个相位偏移只是信号极性的反转。要将给定信号的所有成分的相角偏移 180° , 就需要用到理想转换器。我们感兴趣的另一个相移是 $\pm 90^\circ$ 的相移。特别地, 当给定信号的所有成分的相角都偏移 $\pm 90^\circ$ 时, 得到的时域函数称为信号的希尔伯特变换。

假设有一信号 $g(t)$, 其傅里叶变换为 $G(f)$ 。则 $g(t)$ 的希尔伯特变换可用 $\hat{g}(t)$ 表示为

$$\hat{g}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{g(\tau)}{t - \tau} d\tau \quad (\text{A2.31})$$

显然, 信号 $g(t)$ 的希尔伯特变换是线性运算。反希尔伯特变换 $\hat{g}(t)$, 即恢复出原来信号 $g(t)$ 的变换的定义为

$$g(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\hat{g}(\tau)}{t - \tau} d\tau \quad (\text{A2.32})$$

函数 $g(t)$ 和 $\hat{g}(t)$ 组成了希尔伯特变换对。表 A6.4 给出了希尔伯特变换对的短表。

从希尔伯特变换的定义中不难发现, $\hat{g}(t)$ 可以解释成 $g(t)$ 与时域函数 $1/\pi t$ 的卷积。由卷积定理可知, 两个时域函数的卷积可变换为频域上两个函数傅里叶变换的乘积, 见表 A6.2 中第 12 项。对时域函数 $1/\pi t$, 有(由表 A6.3):

$$\frac{1}{\pi t} \Leftrightarrow -j \operatorname{sgn}(f) \quad (\text{A2.33})$$

其中, $\operatorname{sgn}(f)$ 是频域的符号函数, 其定义为

$$\operatorname{sgn}(f) = \begin{cases} 1, & f > 0 \\ 0, & f = 0 \\ -1, & f < 0 \end{cases} \quad (\text{A2.34})$$

于是, $\hat{g}(f)$ 的傅里叶变换 $\hat{G}(f)$ 为

$$\hat{G}(f) = -j \operatorname{sgn}(f) G(f) \quad (\text{A2.35})$$

式(A2.35)说明, 将信号 $g(t)$ 输入一个频率响应为 $-j \operatorname{sgn}(f)$ 的线性双端设备, 即可得到其希尔伯特变换 $\hat{g}(t)$ 。这个设备可以看成对正频率的输入信号产生 -90° 相移, 而对负频率的输入信号产生 $+90^\circ$ 相移的设备, 如图 A2.3。经过这个设备, 所有频率成分的幅度不受影响。这种理想的设备可以看成是希尔伯特变换器。

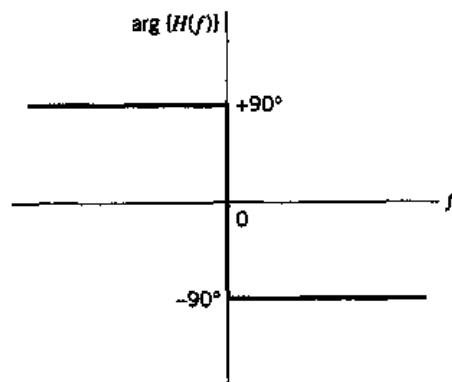


图 A2.3 可获得实数信号的希尔伯特变换的线性双端设备的相位特性

A2.3.1 希尔伯特变换的性质

希尔伯特变换与傅里叶变换有所不同, 其运算完全是在时域进行的。希尔伯特变换有很

多有用的性质,下面列出了其中的一些。设信号 $g(t)$ 是实数值的,这是希尔伯特变换常用的域。对于这类信号,有:

1. 信号 $g(t)$ 及其希尔伯特变换 $\hat{g}(t)$ 具有相同的幅度谱。
2. 如果 $\hat{g}(t)$ 是信号 $g(t)$ 的希尔伯特变换,则 $\hat{g}(t)$ 的希尔伯特变换为 $-g(t)$ 。
3. 信号 $g(t)$ 及其希尔伯特变换 $\hat{g}(t)$ 在整个时间区间 $(-\infty, +\infty)$ 是正交的,即

$$\int_{-\infty}^{\infty} g(t)\hat{g}(t)dt = 0$$

这些性质的证明留给读者进行练习。这些证明可由式(A2.31),式(A2.32)和式(A2.35)得到。

A2.4 信号与系统的复数表示

A2.4.1 预包络

假设有一个实数值的信号 $g(t)$ 。定义信号的预包络信号或解析信号 $g_+(t)$ 为复数值的函数

$$g_+(t) = g(t) + j\hat{g}(t) \quad (\text{A2.36})$$

其中, $\hat{g}(t)$ 是信号 $g(t)$ 的希尔伯特变换。注意到给定信号 $g(t)$ 是预包络 $g_+(t)$ 的实部,该信号的希尔伯特变换是预包络 $g_+(t)$ 的虚部。这种类似相图的方式,简化了电压和电流的交替处理,使得预包络在处理带通的信号与系统时特别有用。

预包络 $g_+(t)$ 最重要的特性之一就是其傅里叶变换。令 $G_+(f)$ 代表 $g_+(t)$ 的傅里叶变换,则有

$$G_+(f) = G(f) + \text{sgn}(f)G(f)$$

由此,不难发现

$$G_+(f) = \begin{cases} 2G(f), & f > 0 \\ G(0), & f = 0 \\ 0, & f < 0 \end{cases} \quad (\text{A2.37})$$

$G(0)$ 是 $G(f)$ 在频率 $f=0$ 时的值。也就是说,信号的预包络在所有负频率上没有频率成分(其傅里叶变换为零)。

从前面的分析易见,对于给定的信号 $g(t)$,可以用两种等效方法中的任何一种来确定该信号的预包络 $g_+(t)$ 。这两种方法是:

1. 先确定信号 $g(t)$ 的希尔伯特变换 $\hat{g}(t)$,然后再用式(A2.36)来计算预包络 $g_+(t)$ 。
2. 先确定信号 $g(t)$ 的傅里叶变换 $G(f)$,接着用式(A2.37)来确定 $G_+(f)$,最后估算 $G_+(f)$ 的傅里叶反变换来求得

$$g_+(t) = 2 \int_0^{\infty} G(f) \exp(j2\pi ft) df \quad (\text{A2.38})$$

对于傅里叶变换为 $G(f)$ 的特殊信号 $g(t)$,这两种方法中的任何一种都优于其他方法。

式(A2.36)定义的是正频率的预包络 $g_+(t)$ 。对称地,可以定义负频率的预包络为

$$g_-(t) = g(t) - j\hat{g}(t) \quad (\text{A2.39})$$

这两个预包络 $g_+(t)$ 和 $g_-(t)$ 互为复共轭,即

$$g_-(t) = g_+^*(t) \quad (\text{A2.40})$$

其中, * 代表复共轭。正如式(A2.37)所强调的,预包络 $g_+(t)$ 的频谱只在正频率处为非零,因此,用一个加号作为其下标。相反,预包络 $g_-(t)$ 的频谱只在负频率处为非零,其傅里叶变换为

$$G_-(f) = \begin{cases} 0, & f > 0 \\ G(0), & f = 0 \\ 2G(f), & f < 0 \end{cases} \quad (\text{A2.41})$$

因此,预包络 $g_+(t)$ 和 $g_-(t)$ 组成了复值信号的互补对。 $g_+(t)$ 和 $g_-(t)$ 之和正好等于原信号 $g(t)$ 的两倍。

A2.4.2 带通信号的规范表示法

设有一带通信号 $g(t)$,其傅里叶变换 $G(f)$ 只在中心频率为 $\pm f_c$ 且频率带宽为 $2W$ 的范围内需要考虑,如图 A2.4(a)所示。我们称 f_c 为载频。在大多数的通信信号中,与 f_c 相比, $2W$ 的带宽是很小的,因此,我们称这样的信号为窄带信号。信号的带宽具体是多少才可以认为是窄带信号,对此我们不加以论述。

假设有一窄带信号 $g(t)$,其傅里叶变换为 $G(f)$,中心频率为 $\pm f_c$,则其预包络可表示为

$$g_+(t) = \bar{g}(t) \exp(j2\pi f_c t) \quad (\text{A2.42})$$

我们称 $\bar{g}(t)$ 为信号的复包络。式(A2.42)可看成是由预包络来定义复包络 $\bar{g}(t)$ 的基础。如图 A2.4(b)所示, $g_+(t)$ 的频谱是局限在频带 $f_c - W \leq f \leq f_c + W$ 内的。因此,对式(A2.42)应用频移特性(表 A6.2 第五项),不难发现复包络 $\bar{g}(t)$ 的频谱被限制在 $-W \leq f \leq W$ 之间,且以原点为中心,如图 A2.4(c)所示。也就是说,带通信号 $g(t)$ 的复包络 $\bar{g}(t)$ 是一个低通信号,这是一个重要的结论。

由定义,给定信号 $g(t)$ 是其预包络 $g_+(t)$ 的实部。我们可以用复包络 $\bar{g}(t)$ 来表示原始的带通信号 $g(t)$:

$$g(t) = \operatorname{Re}[\bar{g}(t) \exp(j2\pi f_c t)] \quad (\text{A2.43})$$

通常, $\bar{g}(t)$ 的值为一个复数。为突出这一特性,可将 $\bar{g}(t)$ 用下式表示:

$$\bar{g}(t) = g_I(t) + jg_Q(t) \quad (\text{A2.44})$$

其中, $g_I(t)$ 和 $g_Q(t)$ 都是实数值的低通函数,其低通特性是从复包络 $\bar{g}(t)$ 继承而来的。因此,可以利用式(A2.43)和式(A2.44),按标准形式来表示原始带通信号 $g(t)$:

$$g(t) = g_I(t) \cos(2\pi f_c t) - g_Q(t) \sin(2\pi f_c t) \quad (\text{A2.45})$$

我们称 $g_I(t)$ 为带通信号 $g(t)$ 的同相分量, $g_Q(t)$ 为带通信号 $g(t)$ 的正交分量。这种命名法承认 $g_Q(t)$ 的乘积因子 $\sin(2\pi f_c t)$ 与 $g_I(t)$ 的乘积因子 $\cos(2\pi f_c t)$ 有 90° 的相位差,且以 $\cos(2\pi f_c t)$ 作为参考。

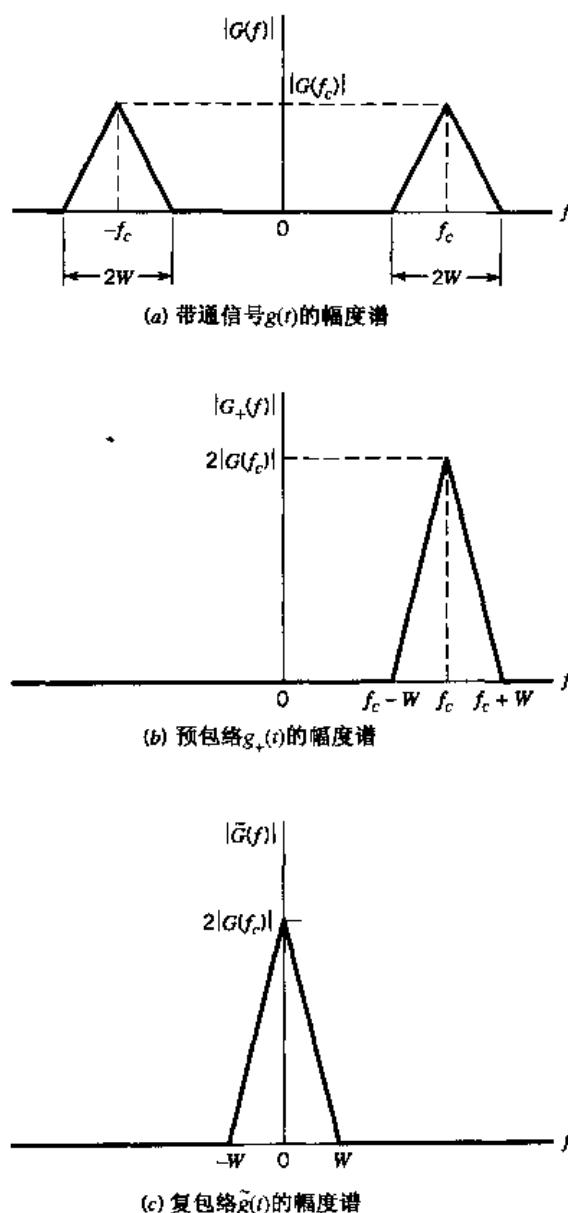


图 A2.4 带通信号、预包络和复包络的幅度谱

根据式(A2.44), 复包络 $\tilde{g}(t)$ 可以画成一个位于 (g_I, g_Q) 平面原点的时变相量, 如图 A2.5(a) 所示。随着时间 t 的变化, 相量的末端在平面中移动。图 A2.5(b) 显示了复指数 $\exp(j2\pi f_c t)$ 的相量表示法。在式(A2.43)给出的定义中, 复包络 $\tilde{g}(t)$ 乘上了复指数 $\exp(j2\pi f_c t)$ 。因此, 将这两个相量的角度相加, 长度相乘, 如图 A2.5(c) 所示。此外, 在图 A2.5(c) 中, (g_I, g_Q) 平面以每秒 $2\pi f_c$ 弧度的角速度旋转。于是, 代表复包络 $\tilde{g}(t)$ 的相量在 (g_I, g_Q) 平面上移动, 同时平面本身也围绕着原点旋转。原始的带通信号 $g(t)$ 是这个时变相量在代表实轴的固定线上的投影, 如图 A2.5(c) 所示。

由于 $g_I(t)$ 和 $g_Q(t)$ 都是限制在频带 $-W \leq f \leq W$ 内的低通信号, 所以可由带通信号 $g(t)$ 采用图 A2.6(a) 所示的方法来得到。图中, 两个低通滤波器是相同的, 带宽均为 W , 为了根据 $g(t)$ 的同相分量和正交分量来重构 $g(t)$, 可以采用图 A2.6(b) 所示的方案。

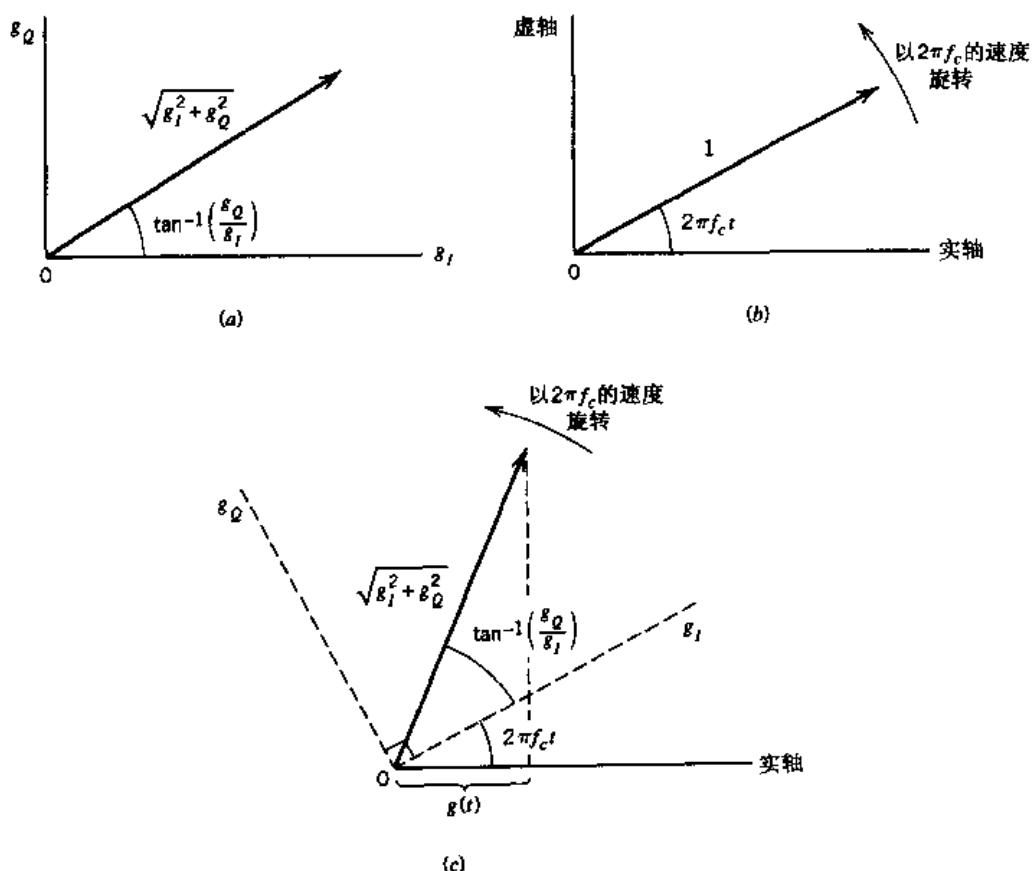
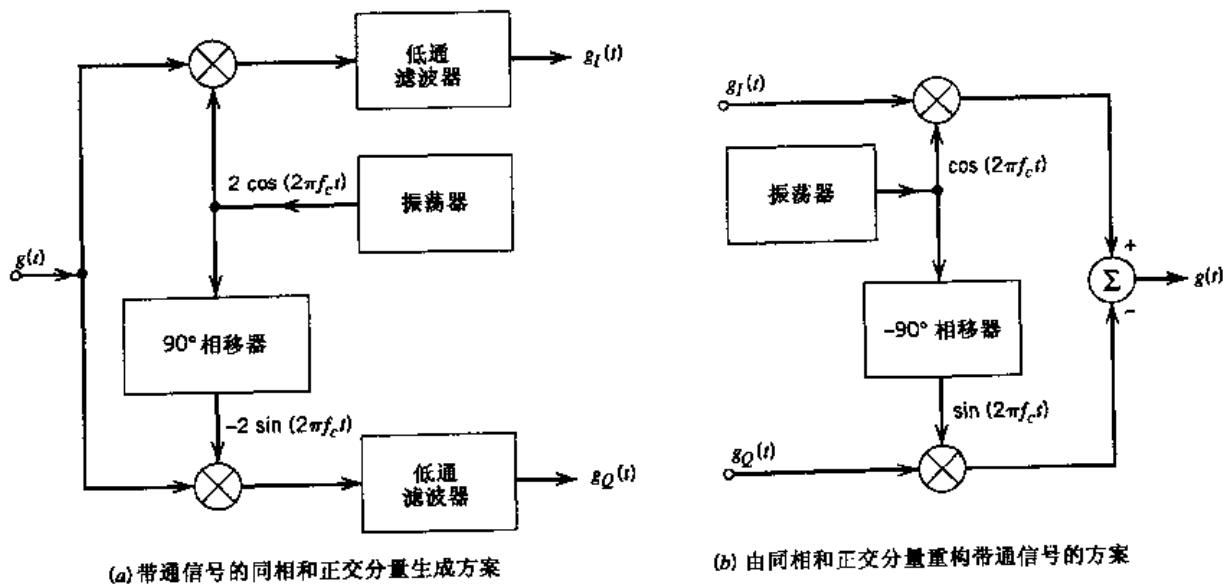
图 A2.5 复包络 $\tilde{g}(t)$ 及其与 $\exp(j2\pi f_c t)$ 乘积的图解

图 A2.6 带通信号的同相和正交分量

图 A2.6 所示的两种方案是研究线性调制系统的基础。低通同相分量 $g_I(t)$ 与 $\cos(2\pi f_c t)$ 的乘积, 以及低通正交分量 $g_Q(t)$ 与 $\sin(2\pi f_c t)$ 的乘积代表调制的线性形式。如果载频 f_c 足够大, 由式(A2.45)定义的带通函数 $g(t)$ 可以看成是一个带通信号波形。相应地, 由 $g_I(t)$ 和

$g_Q(t)$ 到 $g(t)$ 的映射称为通带调制。

式(A2.44)是复包络 $\tilde{g}(t)$ 的笛卡儿形式。也可以用极坐标形式来表示该式：

$$\tilde{g}(t) = a(t) \exp[j\phi(t)] \quad (\text{A2.46})$$

其中, $a(t)$ 和 $\phi(t)$ 都是实数值的低通函数。基于这种极坐标表示法, 原始的带通信号 $g(t)$ 可定义为

$$g(t) = a(t) \cos[2\pi f_c t + \phi(t)] \quad (\text{A2.47})$$

可将 $a(t)$ 看成是带通信号 $g(t)$ 的自然包络(简称包络), 而 $\phi(t)$ 是信号的相位。式(A2.47)代表幅度调制和角度调制的混合调制。实际上, 混合调制包括 AM, FM 和 PM, 这三种调制均为混合调制特例。

通过这些讨论可知, 无论采用式(A2.45)中的同相和正交分量, 还是采用式(A2.47)中的包络和相位来表示带通信号 $g(t)$ (调制后的), 信号 $g(t)$ 的信息分量都完全保存在复包络 $\tilde{g}(t)$ 中。

A2.4.3 术语

应当注意到前面介绍的用来描述带通信号 $g(t)$ 的三种包络的不同之处。现将这三种包络的定义归纳如下：

1. 正频率的预包络 $g_+(t)$ 的定义为

$$g_+(t) = g(t) + j\hat{g}(t)$$

其中, $\hat{g}(t)$ 是信号 $g(t)$ 的希尔伯特变换。根据这种表示法, $\hat{g}(t)$ 可以看成是 $g(t)$ 的正交函数。相应地, 在频域有

$$G_+(f) = \begin{cases} 2G(f), & f > 0 \\ G(0), & f = 0 \\ 0, & f < 0 \end{cases}$$

2. 复包络等于 $\tilde{g}(t)$ 的预包络 $g_+(t)$ 的频移形式, 即

$$\tilde{g}(t) = g_+(t) \exp(-j2\pi f_c t)$$

其中, f_c 是带通信号 $g(t)$ 的载频。

3. 包络 $a(t)$ 等于复包络的幅度, $\tilde{g}(t)$ 等于预包络 $g_+(t)$ 的幅度, 即

$$a(t) = |\tilde{g}(t)| = |g_+(t)|$$

值得注意的是, 对于带通信号 $g(t)$, 预包络 $g_+(t)$ 是一个复带通信号, 其值取决于载频 f_c 。另一方面, 包络 $a(t)$ 总是一个实数低通信号, 且复包络 $\tilde{g}(t)$ 通常是一个复数的低通信号。 $a(t)$ 和 $\tilde{g}(t)$ 的值是与选择的载频 f_c 是相互独立的。这个特性使复包络 $\tilde{g}(t)$ 相对于原始信号 $g(t)$ 具有易分析的优点。

$g(t)$ 的包络 $a(t)$ 和相位 $\phi(t)$, 与图 A2.5(a) 所示的 $g(t)$ 的相互正交的分量 $g_I(t)$ 和 $g_Q(t)$ 之间有如下关系:

$$a(t) = \sqrt{g_I^2(t) + g_Q^2(t)}$$

$$\phi(t) = \tan^{-1}\left(\frac{g_Q(t)}{g_I(t)}\right)$$

反过来,有

$$g_I(t) = a(t) \cos[\phi(t)]$$

$$g_Q(t) = a(t) \sin[\phi(t)]$$

因此,每个带通信号的相互正交的分量都含有幅度和相位信息。两个分量都有惟一的相位 $\phi(t)$, $\phi(t)$ 是模 2π 的。

A2.4.4 带通系统

了解了怎样运用带通信号的复低通表示法,就有可能开发一种处理带通系统分析过程的相应程序。特别地,我们希望证明:带通系统的分析,可以通过建立一个介于低通和带通系统之间的模型(更准确地说是同构的)得到大大的简化。这个模型是建立在希尔伯特变换的基础上的。

假设有一窄带信号 $x(t)$,其傅里叶变换为 $X(f)$ 。信号 $x(t)$ 的频谱被限制在载频 f_c 的 $\pm W$ 赫之间。同样,假设 $W < f_c$ 。该信号由其同相和正交分量表示为

$$x(t) = x_I(t) \cos(2\pi f_c t) - x_Q(t) \sin(2\pi f_c t) \quad (\text{A2.48})$$

其中, $x_I(t)$ 是同相分量, $x_Q(t)$ 是正交分量。用 $\tilde{x}(t)$ 来代表 $x(t)$ 的复包络,则

$$\tilde{x}(t) = x_I(t) + jx_Q(t) \quad (\text{A2.49})$$

将信号 $x(t)$ 应用于冲激响应为 $h(t)$ 、频率响应为 $H(f)$ 的线性时不变带通系统。假设此系统的频率响应被限制在载频 f_c 的 $\pm B$ 赫之间。系统的带宽 $2B$ 通常小于或等于输入信号的带宽 $2W$ 。我们希望根据两个相互正交的分量 $h_I(t)$ 和 $h_Q(t)$ 来表示带通冲激响应 $h(t)$ 。类似于带通信号的表示法,可将 $h(t)$ 表示为

$$h(t) = h_I(t) \cos(2\pi f_c t) - h_Q(t) \sin(2\pi f_c t) \quad (\text{A2.50})$$

定义带通系统的复冲激响应为

$$\tilde{h}(t) = h_I(t) + jh_Q(t) \quad (\text{A2.51})$$

因此得到复数表示法

$$h(t) = \operatorname{Re}[\tilde{h}(t) \exp(j2\pi f_c t)] \quad (\text{A2.52})$$

注意 $h_I(t)$ 、 $h_Q(t)$ 和 $\tilde{h}(t)$ 都是频带被限制在 $-B \leq f \leq B$ 的低通函数。

由式(A2.51),可根据带通冲激响应 $h(t)$ 的两个正交分量 $h_I(t)$ 和 $h_Q(t)$ 来确定复冲激响应 $h(t)$ 。此外,也可以按照以下方法,由带通频率响应 $H(f)$ 来确定冲激响应 $h(t)$ 。首先,由式(A2.52)得到

$$2h(t) = \tilde{h}(t) \exp(j2\pi f_c t) + \tilde{h}^*(t) \exp(-j2\pi f_c t) \quad (\text{A2.53})$$

其中, $\tilde{h}^*(t)$ 的是 $\tilde{h}(t)$ 的复共轭。对式(A2.53)进行傅里叶变换,并由傅里叶变换的复共轭的

性质(表 A6.2 第 10 项), 得到

$$2H(f) = \tilde{H}(f - f_c) + \tilde{H}^*(-f - f_c) \quad (\text{A2.54})$$

其中, $H(f)$ 是 $b(t)$ 的傅里叶变换, $\tilde{H}(f)$ 是 $\tilde{b}(t)$ 的傅里叶变换。因为 $b(t)$ 是实冲激响应, 所以式(A2.54)满足 $H^*(f) = H(-f)$ 的要求。又由于 $\tilde{H}(f)$ 代表的是一个限制于 $|f| \leq B$ ($B < f_c$) 的低通频率响应, 因此可由式(A2.54)推得

$$\tilde{H}(f - f_c) = 2H(f), \quad f > 0 \quad (\text{A2.55})$$

式(A2.55)说明, 对于指定的带通频率响应 $H(f)$, 可通过将对应于正频率的 $H(f)$ 部分平移到原点, 并乘上放大因子 2 来确定 $\tilde{H}(f)$ 。为了确定复冲激响应 $b(t)$, 可用 $\tilde{H}(f)$ 的傅里叶反变换来得到

$$\tilde{b}(t) = \int_{-\infty}^{\infty} \tilde{H}(f) \exp(j2\pi ft) df \quad (\text{A2.56})$$

上述带通的信号与系统的表示法, 为确定(输入为带通信号的)带通系统的输出提供了一种基础的有效方法。假设输入信号 $x(t)$ 的频谱和系统的频率响应 $H(f)$ 都具有相同的中心频率 f_c 。实际上, 由于在选择载频或中心频率时有相当的自由度, 所以, 没有必要去考虑输入信号的载频是否与带通系统的中心频率一致。这样一来, 将输入信号的载频改变 Δf_c , 其实就只是在输入信号的复包络或带通系统的复冲激响应上加上(或减去)了一个因子 $\exp(\pm j2\pi\Delta f_c t)$ 。因此, 可以证明, 前面假设的 $X(f)$ 和 $H(f)$ 具有相同的中心频率是有效的。用 $y(t)$ 代表系统的输出信号, 显然, $y(t)$ 也是一个带通信号, 可根据低通复包络将 $\tilde{y}(t)$ 表示为

$$y(t) = \operatorname{Re}[\tilde{y}(t) \exp(j2\pi f_c t)] \quad (\text{A2.57})$$

输出信号 $y(t)$ 与输入信号 $x(t)$ 和系统的冲激响应 $b(t)$ 的卷积有关:

$$y(t) = \int_{-\infty}^{\infty} b(\tau)x(t - \tau) d\tau \quad (\text{A2.58})$$

根据预包络, 有 $b(t) = \operatorname{Re}[b_+(t)]$ 和 $x(t) = \operatorname{Re}[x_+(t)]$ 。因此, 可根据预包络 $x_+(t)$ 和 $b_+(t)$ 将式(A2.58)写为

$$y(t) = \int_{-\infty}^{\infty} \operatorname{Re}[b_+(\tau)] \operatorname{Re}[x_+(t - \tau)] d\tau \quad (\text{A2.59})$$

更进一步, 可以利用下式所示的预包络的基本性质(此处不加证明):

$$\int_{-\infty}^{\infty} \operatorname{Re}[b_+(\tau)] \operatorname{Re}[x_+(\tau)] d\tau = \frac{1}{2} \operatorname{Re} \left[\int_{-\infty}^{\infty} b_+(\tau)x_+^*(\tau) d\tau \right] \quad (\text{A2.60})$$

其中, 与式(A2.59)一样, 用 τ 作为积分变量。用 $x(-\tau)$ 代替 $x(\tau)$, 就可以将式(A2.60)右边的复共轭移走。因此, 由式(A2.59)中 $x_+(t - \tau)$ 与式(A2.60)中 $x_+(\tau)$ 的幅度的代数差, 并利用预包络与带通函数复包络的关系, 可得到

$$\begin{aligned}
 y(t) &= \frac{1}{2} \operatorname{Re} \left[\int_{-\infty}^{\infty} h_+(\tau) x_+(t - \tau) d\tau \right] \\
 &= \frac{1}{2} \operatorname{Re} \left[\int_{-\infty}^{\infty} \tilde{h}(\tau) \exp(j2\pi f_c \tau) \tilde{x}(t - \tau) \exp(j2\pi f_c(t - \tau)) d\tau \right] \\
 &= \frac{1}{2} \operatorname{Re} \left[\exp(j2\pi f_c t) \int_{-\infty}^{\infty} \tilde{h}(\tau) \tilde{x}(t - \tau) d\tau \right]
 \end{aligned} \quad (\text{A2.61})$$

对比式(A2.57)和式(A2.61)的右边部分,可以很容易推导出:对于足够大的载频 f_c ,输出信号的复包络 $\bar{y}(t)$ 、输入信号的复包络 $\bar{x}(t)$ 和带通系统的复冲激响应 $\tilde{h}(t)$ 有如下关系:

$$2\bar{y}(t) = \int_{-\infty}^{\infty} \tilde{h}(\tau) \tilde{x}(t - \tau) d\tau \quad (\text{A2.62})$$

也可简写为

$$2\bar{y}(t) = \tilde{h}(t) \star \tilde{x}(t) \quad (\text{A2.63})$$

其中,★代表卷积。换言之,除了放大因子 2 以外,带通系统输出信号的复包络 $\bar{y}(t)$ 可以由系统的复冲激响应 $\tilde{h}(t)$ 和输入带通信号的复包络 $\tilde{x}(t)$ 经卷积求得。对于卷积来说,对于带通函数和相应的低通函数,式(A2.63)得到的是同构的结果。

此结果的重要性在于,在处理带通信号与系统时,我们只需关心低通函数 $\tilde{x}(t)$ 、 $\bar{y}(t)$ 和 $\tilde{h}(t)$,即激励、响应和系统。也就是说,带通系统的分析由于乘上了因子 $\exp(j2\pi f_c t)$ 而变得复杂,但可以用一个等效且简单的低通分析来代替,这种分析完全保留了滤波过程的本质。这个过程的图示说明见图 A2.7。

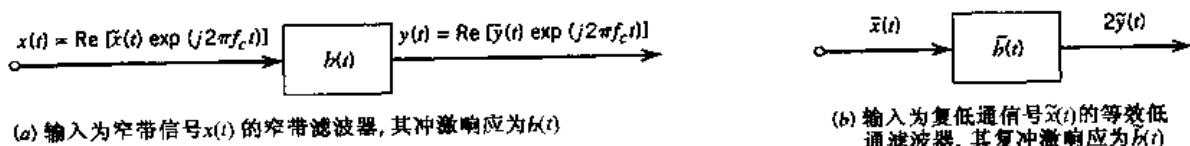


图 A2.7 窄带滤波器和等效低通滤波器

输入信号的复包络 $\tilde{x}(t)$ 和带通系统的复冲激响应 $\tilde{h}(t)$, 分别由其同相和正交分量来定义,见式(A2.49)和式(A2.51)。将这些关系式代入到式(A2.63),得到

$$2\bar{y}(t) = [h_I(t) + jh_Q(t)] \star [x_I(t) + jx_Q(t)] \quad (\text{A2.64})$$

由于卷积具有分配律,式(A2.64)可以等效地写成

$$2\bar{y}(t) = [h_I(t) \star x_I(t) - h_Q(t) \star x_Q(t)] + j[h_Q(t) \star x_I(t) + h_I \star x_Q(t)] \quad (\text{A2.65})$$

响应的复包络 $\bar{y}(t)$ 可以用其同相和正交分量定义为

$$\bar{y}(t) = y_I(t) + jy_Q(t) \quad (\text{A2.66})$$

对比式(A2.65)和式(A2.66)的实部和虚部,可得同相分量 $y_I(t)$ 的关系式

$$2y_I(t) = h_I(t) \star x_I(t) - h_Q(t) \star x_Q(t) \quad (\text{A2.67})$$

以及正交分量 $y_Q(t)$ 的关系式

$$2y_Q(t) = h_Q(t) \star x_I(t) + h_I(t) \star x_Q(t) \quad (\text{A2.68})$$

因此,为了估算系统输出的复包络 $\tilde{y}(t)$ 的同相分量和正交分量,可以采用图 A2.8 所示的低通等效模型。模型中所有的信号和冲激响应都是实数值的低通函数。因此,该等效模型为有效地模拟低通滤波器或用于数字计算机的通信信道提供了实践基础。

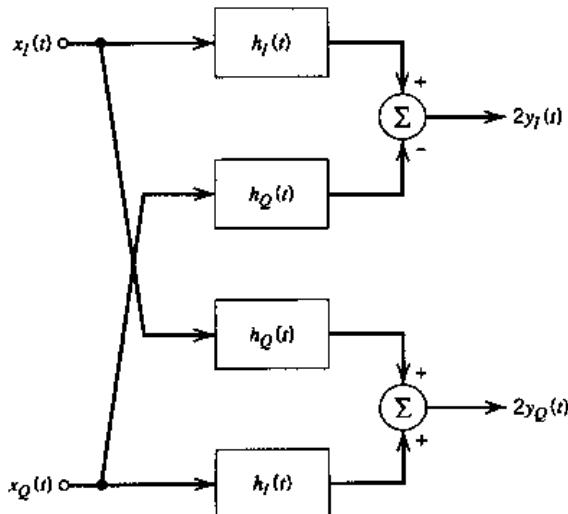


图 A2.8 带通滤波器输出响应的同相和正交分量与输入信号的同相和正交分量间的关系框图

综上所述,当输入信号为带通信号(载频为 f_c)时,带通系统(中心频率为 f_c)输出响应的估算可按如下过程进行:

1. 输入带通信号 $x(t)$ 用其复包络 $\hat{x}(t)$ 代替, $\hat{x}(t)$ 与 $x(t)$ 的关系为

$$x(t) = \operatorname{Re}[\hat{x}(t) \exp(j2\pi f_c t)]$$

2. 冲激响应为 $b(t)$ 的带通系统用低通模型代替,复冲激响应 $\tilde{b}(t)$ 与 $b(t)$ 的关系:

$$b(t) = \operatorname{Re}[\tilde{b}(t) \exp(j2\pi f_c t)]$$

3. 输出带通信号 $y(t)$ 的复包络 $\tilde{y}(t)$ 由 $\tilde{b}(t)$ 和 $\hat{x}(t)$ 的卷积得到

$$2\tilde{y}(t) = \tilde{b}(t) \star \hat{x}(t)$$

4. 最后,由复包络 $\tilde{y}(t)$,通过下面的关系式得到所需要的输出 $y(t)$:

$$y(t) = \operatorname{Re}[\tilde{y}(t) \exp(j2\pi f_c t)]$$

附录 3 贝塞尔函数

A3.1 贝塞尔等式的级数解

n 阶贝塞尔等式的最基本形式是

$$x^2 \frac{d^2y}{dx^2} + x \frac{dy}{dx} + (x^2 - n^2)y = 0 \quad (\text{A3.1})$$

这是最重要的变系数差分方程之一。^[1]对于每个 n ,都有一个用幕级数定义的解为

$$J_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{1}{2}x\right)^{n+2m}}{m!(n+m)!} \quad (\text{A3.2})$$

函数 $J_n(x)$ 称为第一类 n 阶贝塞尔函数。在式(A3.1)中有两个系数函数,分别是 $1/x$ 和 $(1 - n^2/x^2)$ 。因此,除了原点外,函数 $J_n(x)$ 没有有限奇异点。由此可知,式(A3.2)对所有的 $x > 0$ 级数展开式都是收敛的。这样,式(A3.2)就可被用于对 $J_n(x)$ 进行数值计算,其中, $n = 0, 1, 2, \dots$ 表 A6.5 给出了不同阶数 n 和不同 x 时的 $J_n(x)$ 。不难注意到, $J_0(x)$ 和 $J_1(x)$ 的图形分别与 $\cos x$ 和 $\sin x$ 的图形相似(参见第 2 章图 2.23)。

$J_n(x)$ 也可以表示成积分的形式为

$$J_n(x) = \frac{1}{\pi} \int_0^\pi \cos(x \sin \theta - n\theta) d\theta \quad (\text{A3.3})$$

或者等效地表示为

$$J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(jx \sin \theta - jn\theta) d\theta \quad (\text{A3.4})$$

A3.2 贝塞尔函数的性质

贝塞尔函数 $J_n(x)$ 具有如下性质:

1. $J_n(x) = (-1)^n J_{-n}(x) \quad (\text{A3.5})$

为了证明这个关系式,在式(A3.3)中,用 $(\pi - \theta)$ 代替 θ 。由于 $\sin(\pi - \theta) = \sin \theta$, 所以有

$$\begin{aligned} J_n(x) &= \frac{1}{\pi} \int_0^\pi \cos(x \sin \theta + n\theta - n\pi) d\theta \\ &= \frac{1}{\pi} \int_0^\pi [\cos(n\pi) \cos(x \sin \theta + n\theta) + \sin(n\pi) \sin(x \sin \theta + n\theta)] d\theta \end{aligned}$$

由于 n 为整数,因此有

$$\begin{aligned}\cos(n\pi) &= (-1)^n \\ \sin(n\pi) &= 0\end{aligned}$$

所以

$$J_n(x) = \frac{(-1)^n}{\pi} \int_0^\pi \cos(x \sin \theta + n\theta) d\theta \quad (\text{A3.6})$$

由式(A3.3)可知,用 $-n$ 代替 n ,可得

$$J_{-n}(x) = \frac{1}{\pi} \int_0^\pi \cos(x \sin \theta - n\theta) d\theta \quad (\text{A3.7})$$

由式(A3.6)和式(A3.7)可以立即得到所需的结果。

$$2. \quad J_n(x) = (-1)^n J_{-n}(-x) \quad (\text{A3.8})$$

在式(A3.3)中用 $-x$ 代替 x ,然后利用式(A3.6)即可得到上式。

$$3. \quad J_{n-1}(x) + J_{n+1}(x) = \frac{2n}{x} J_n(x) \quad (\text{A3.9})$$

这个递推公式在构造贝塞尔系数表时特别有用,它可由式(A3.2)的幂级数推出。

4. 对于较小的 x 值,有

$$J_n(x) \approx \frac{x^n}{2^n n!} \quad (\text{A3.10})$$

这个关系式只需在式(A3.2)中保留幂级数的第一项,并忽略高次项,即可得到。于是,当 x 较小时,有

$$J_0(x) \approx 1$$

$$J_1(x) \approx \frac{x}{2} \quad (\text{A3.11})$$

$$J_n(x) \approx 0 \quad \text{for } n > 1$$

5. 对于较大的 x 值,有

$$J_n(x) \approx \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{\pi}{4} - \frac{n\pi}{2}\right) \quad (\text{A3.12})$$

上式说明,对于较大的 x 值,贝塞尔函数 $J_n(x)$ 类似于一个幅度递减的正弦信号。

6. 当 x 为一个固定的实数时,随着阶数 n 趋向于无穷, $J_n(x)$ 趋近于 0。

$$7. \quad \sum_{n=-\infty}^{\infty} J_n(x) \exp(jn\phi) = \exp(jx \sin \phi) \quad (\text{A3.13})$$

下面证明这个性质。由和式 $\sum_{n=-\infty}^{\infty} J_n(x) \exp(jn\phi)$,并利用式(A3.4),可得

$$\sum_{n=-\infty}^{\infty} J_n(x) \exp(jn\phi) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \exp(jn\phi) \int_{-\pi}^{\pi} \exp(jx \sin \theta - jn\theta) d\theta$$

交换积分次序然后求和,得

$$\sum_{n=-\infty}^{\infty} J_n(x) \exp(jn\phi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta \exp(jx \sin \theta) \sum_{n=-\infty}^{\infty} \exp[jn(\phi - \theta)] \quad (\text{A3.14})$$

这里需用到傅里叶变换理论中的关系式

$$\delta(\phi) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \exp[jn(\phi)], \quad -\pi \leq \phi \leq \pi \quad (\text{A3.15})$$

其中, $\delta(\phi)$ 是一个 δ 函数。将(A3.15)代入式(A3.14), 并应用 δ 函数的移位性质, 可得

$$\begin{aligned} \sum_{n=-\infty}^{\infty} J_n(x) \exp(jn\phi) &= \int_{-\pi}^{\pi} \exp(jx \sin \theta) \delta(\phi - \theta) d\theta \\ &= \exp(jx \sin \phi) \end{aligned}$$

这正是所需的结果。

$$8. \quad \sum_{n=-\infty}^{\infty} J_n^2(x) = 1 \quad \text{对于所有 } x \quad (\text{A3.16})$$

下面证明这个性质。可认为 $J_n(x)$ 均为实数。因此, 将式(A3.4)乘上其自身的复共轭, 并对所有可能的 n 求和, 得到

$$\sum_{n=-\infty}^{\infty} J_n^2(x) = \frac{1}{(2\pi)^2} \sum_{n=-\infty}^{\infty} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \exp(jx \sin \theta - jn\theta - jx \sin \phi + jn\phi) d\theta d\phi$$

交换两个积分号与求和的次序, 得

$$\begin{aligned} \sum_{n=-\infty}^{\infty} J_n^2(x) &= \\ \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} d\theta d\phi \exp[jx(\sin \theta - \sin \phi)] &\sum_{n=-\infty}^{\infty} \exp[jn(\phi - \theta)] \end{aligned} \quad (\text{A3.17})$$

将式(A3.15)代入式(A3.17), 并应用 δ 函数的移位特性, 可得

$$\sum_{n=-\infty}^{\infty} J_n^2(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta = 1$$

这正是所需的结果。

贝塞尔函数 $J_n(x)$ 的这些性质中的一部分也可通过参考表 A6.5, 用数值项来进行说明。

A3.3 修正贝塞尔函数

n 阶修正贝塞尔方程为

$$x^2 \frac{d^2y}{dx^2} + x \frac{dy}{dx} - (x^2 + n^2)y = 0 \quad (\text{A3.18})$$

由 $j^2 = -1$ (其中 j 是 -1 的平方根), 可将式(A3.18)写为

$$x^2 \frac{d^2y}{dx^2} + x \frac{dy}{dx} + (j^2 x^2 - n^2)y = 0$$

由上式可知, 式(A3.18)显然就是贝塞尔函数, 即式(A3.1), 其中, x 被 jx 代替。同样, 在

式(A3.2)中也用 jx 代替 x , 得到

$$\begin{aligned} J_n(jx) &= \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{jx}{2}\right)^{n+2m}}{m!(n+m)!} \\ &= j^n \sum_{m=0}^{\infty} \frac{\left(\frac{x}{2}\right)^{n+2m}}{m!(n+m)!} \end{aligned}$$

不难注意到 $J_n(jx)$ 乘上一个常数仍然是贝塞尔方程的解。因此, 将 $J_n(jx)$ 乘以一个常数 j^{-n} , 得到

$$j^{-n} J_n(jx) = \sum_{m=0}^{\infty} \frac{\left(\frac{1}{2}x\right)^{n+2m}}{m!(n+m)!}$$

这个新的函数称为第一类 n 阶修正贝塞尔函数, 用 $I_n(x)$ 来表示。这样, 我们就可以正式地给出式(A3.18)修正贝塞尔方程的解

$$\begin{aligned} I_n(x) &= j^{-n} J_n(jx) \\ &= \sum_{m=0}^{\infty} \frac{\left(\frac{1}{2}x\right)^{n+2m}}{m!(n+m)!} \end{aligned} \quad (\text{A3.19})$$

对于所有 n , 修正贝塞尔函数 $I_n(x)$ 是自变量 x 的单调递增实函数, 图 A3.1 给出了 $n = 0, 1, 2$ 时的情形。

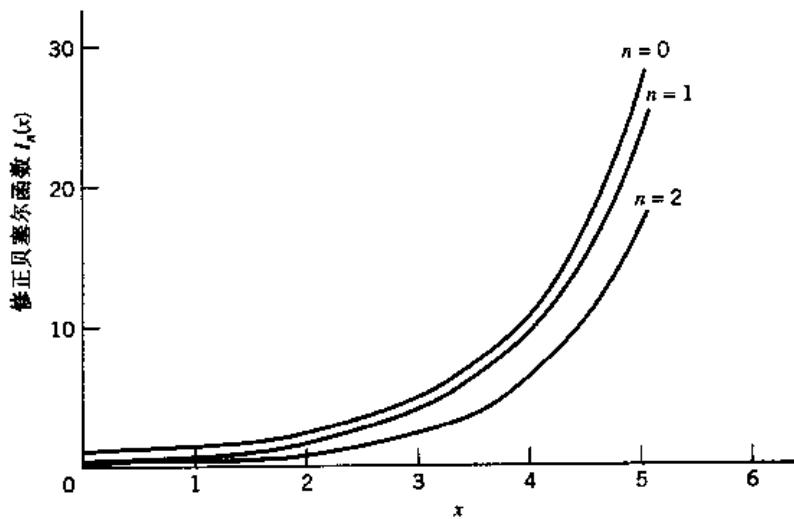


图 A3.1 不同阶数 n 对应的修正贝塞尔函数 $I_n(x)$

除了一个重要差别之外, 修正贝塞尔函数 $I_n(x)$ 与原始贝塞尔函数 $J_n(x)$ 是相同的。这个重要差别就是, 式(A3.19)的级数展开中各项都是正的, 而在式(A3.2)中是正负号交替的。 $J_n(x)$ 和 $I_n(x)$ 的关系与三角函数 $\cos x$ 和 $\sin x$, 以及双曲函数 $\cosh x$ 和 $\sinh x$ 的关系类似。

修正贝塞尔函数 $I_n(x)$ 的一个有趣的性质来自于式(A3.13)。特别地,在式(A3.13)中用 jx 代替 x ,用 $\theta - \pi/2$ 代替 ϕ ,并在式(A3.19)的第一行引用 $I_n(x)$ 的定义,可得

$$\sum_{n=-\infty}^{\infty} I_n(x) \exp(jn\theta) = \exp(x \cos \theta) \quad (\text{A3.20})$$

从这个关系式中可得到

$$I_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(x \cos \theta) \cos(n\theta) d\theta \quad (\text{A3.21})$$

当然,这个积分公式也可以由式(A3.4)经过适当的变化得到。

当自变量 x 较小时,可以由式(A3.19)的级数表示法直接得到以下的渐近估计值:

$$I_0(x) \rightarrow 1 \quad \text{对于 } x \rightarrow 0 \quad (\text{A3.22})$$

和

$$I_n(x) \rightarrow 0 \quad \text{对于 } n \geq 1 \text{ 且 } x \rightarrow 0 \quad (\text{A3.23})$$

对于较大的 x ,可以得到以下对所有 $n \geq 0$ 都成立的 $I_n(x)$ 的渐近估计值

$$I_n(x) \approx \frac{\exp(x)}{\sqrt{2\pi x}} \quad \text{对于 } x \rightarrow \infty \quad (\text{A3.24})$$

注意,对于较大的 x , $I_n(x)$ 的这种渐近特性与阶数 n 是独立的。

注释与参考

- [1] 式(A3.1)是以德国数学家和天文学家 Friedrich Wilhelm Bessel(1784 ~ 1846)来命名的。关于该方程解的详细分析和相关论述,可参看 Wylie 和 Barrett(1982)以及 Watson(1966)的论文。

附录 4 汇合型超几何函数

A4.1 库默尔方程

汇合型超几何函数^[1]是库默尔(Kummer)差分方程的一个解：

$$x \frac{d^2y}{dx^2} + (b - x) \frac{dy}{dx} - ay = 0 \quad (\text{A4.1})$$

通常，这里的参数 a 和 b 是复数。在 $b \neq 0, -1, -2, \dots$ 的情况下，库默尔方程的解可以用级数的形式来定义为

$${}_1F_1(a; b; x) = 1 + \frac{a}{b} \frac{x}{1!} + \frac{a(a+1)}{b(b+1)} \frac{x^2}{2!} + \dots \quad (\text{A4.2})$$

${}_1F_1(a; b; x)$ 代表参数为 a 和 b 的汇合型超几何函数。在这个表示法中，第一个下标代表式(A4.2)中通项的分子中的因子数，第二个下标代表分母中除 $n!$ 之外的因子数。显然，式(A4.2)中两个下标都是 1。

A4.2 汇合型超几何函数的性质

性质 1

对于较小的 x 值，汇合型超几何函数近似为

$${}_1F_1(a; b; x) \approx 1 + \frac{a}{b} x \quad \text{对于 } x \rightarrow 0 \quad (\text{A4.3})$$

这个性质可以由式(A4.2)的级数展开式直接得到。

性质 2

对于 $a = -1, b = 1$ 的情况，有精确的恒等式

$${}_1F_1(-1; 1; x) = 1 - x \quad \text{对于所有 } x \quad (\text{A4.4})$$

这个性质也可以由式(A4.2)的级数展开式直接得到。

性质 3

当 $a = -1/2, b = 1$ 时，对于所有的 x ，汇合型超几何函数与修正贝塞尔函数有如下关系：

$${}_1F_1\left(-\frac{1}{2}; 1; -x\right) = \exp\left(-\frac{x}{2}\right) \left((1+x)I_0\left(\frac{x}{2}\right) + xI_2\left(\frac{x}{2}\right) \right) \quad (\text{A4.5})$$

其中， $I_n(x)$ 是 n 阶变形贝塞尔函数。

当 x 较大时，式(A4.5)会出现特殊的情况。从附录 3 中给出的修正贝尔塞函数的定义，当

x 较大时,有如下的渐近公式:

$$I_n(x) \approx \frac{\exp(x)}{\sqrt{2\pi x}} \quad \text{对于 } x \rightarrow \infty \quad (\text{A4.6})$$

因此,将式(A4.5)与式(A4.6)结合,可得到简单的结果

$${}_1F_1\left(-\frac{1}{2}; 1; -x\right) \approx 2\sqrt{\frac{x}{\pi}} \quad \text{对于 } x \rightarrow \infty \quad (\text{A4.7})$$

注释与参考

- [1] 对汇合型超几何函数的讨论,可参阅 Jeffreys 和 Jeffreys (1956)。Abramowitz 和 Stegun (1965) 给出了这些函数值的列表。

附录 5 密 码 学

对信息传输的安全性和完整性来说,保密无疑是重要的。的确,由于如今的商业、贸易和个人事务的开展都需要通过计算机这个中间媒介,计算机通信已基本取代了过去以纸为媒介的传统通信方式。因此,与过去相比,安全通信变得更加意义深远。

密码学是保密通信科学的统称。这个名称起源于希腊语中的 *kryptos* 和 *logos*,其含义分别是“隐藏的”和“词”。^[1]密码学的主题可分为加密技术和密码分析。加密技术就是通过加密将消息转变成编码形式,并通过解密恢复原始消息的技术。被加密的原始消息称为明文,加密后的消息称为密段或密文,密段和密文两个词可以互换。用于加密的数据变换的集合称为密码。通常这种变换是以一个或多个密钥为参数。另一方面,加密分析研究如何通过截断密码或伪造可被当成原始信号来接收的编码信号来破坏密码通信。

加密系统提供三个重要的服务:

1. 保密性,是指未授权用户不能获得信息。
2. 可靠性,是指消息源的有效性。
3. 完整性,是指要保证消息在传送过程中不会被偶然的或故意的破坏所改变。

传统的密码系统依靠一条专用且必须保密的信息来工作,这条信息被称为密钥。因此,传统的加密技术是指单密钥加密技术或秘密密钥加密技术。^[2]运用这种形式的加密技术时有一个前提,就是加密者与解密者要都知道密钥,而其他的人都不知道。这个假设就是说一旦消息被加密,如果不知道密钥,就(几乎)不可能将它解密。

公开密钥加密技术^[3]又称为双密钥加密技术。这种形式的加密技术不同于传统的加密技术,两个用户不再共享一个密钥,而是每个用户提供他自己的密钥资料。这些密钥资料被分为两个部分:公开部分和专用部分。公开部分由公开变换生成,而专用部分由专用变换生成。当然,为保证通信安全,在两用户之间,专用变换必须保密。

A5.1 密钥的加密技术

密钥加密系统中的基本信息流程如图 A5.1 所示。消息源生成一个明文消息,明文消息在系统发送的末端被加密成一个密码。这个密码经过一个“不安全的”信道被送给一个授权用户。如果信道的安全性不能满足用户的需要,这个信道就可认为是不安全的。也就是说,可认为在传输过程中密码可能被敌方密码分析者窃听^[4](即加密系统可能会被入侵)。因此就需要对消息进行加密,使敌人不能了解到明文的内容。

抽象地说,加密系统或者密码(较短的)就是明文空间(即可能的明文消息的集合)到密码空间(即所有可能的密段的集合)的一组可逆变换。每个特定的变换对应于一个特定密钥的明文加密。变换的可逆性是指,在知道密钥的情况下,只有一种可能的解密结果。令 **X** 代表明文,**Y** 代表密段,**Z** 代表密钥,**F** 代表(可逆的)产生密段 **Y** 的变换,则

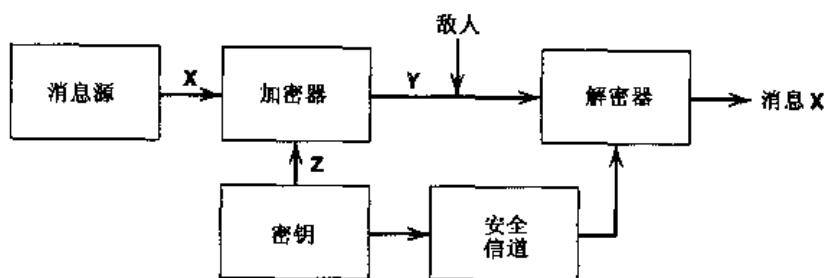


图 A5.1 秘密密钥加密系统框图

$$Y = F(X, Z) = F_z(X) \quad (A5.1)$$

这个变换就是要产生对敌人无用的密段 Y 。在系统的接收端,密段 Y 通过逆变换 F^{-1} 被解密并恢复为原始的明文 X :

$$F^{-1}(Y, Z) = F_z^{-1}(Y) = F_z^{-1}(F_z(X)) = X \quad (A5.2)$$

从物理实现的角度来看,加密系统由指令集、物理硬件或计算机程序组成。在任何情况下,系统都要设计成可以采用不同的方式对明文进行加密(对密段进行解密)。在实际应用中,具体采用什么方式进行加密,由指定密钥决定。

系统的安全性取决于密钥的保密特性,如图 A5.1 所示。密钥必须由安全信道(如注册邮件、信使服务等)发往接收机。此图描述的加密系统为保密性问题提供了一种解决方案,它能防止敌人从不安全信道内传送的消息中提取信息。加密同样也为鉴别问题提供了一种解决方案,它可防止敌方密码分析者模仿消息的发送者。在后一种情况中,敌方密码分析者是指那些产生虚假密段 Y' 的人,如图 A5.2 所示。这些虚假的密段 Y 被送往了接收者(解密者)。真正的密码(如虚线所示)送给了敌方密码分析者。这说明敌人产生了虚假密段 Y' ,但他们并没有看到真实的密码。而接收者此时就可以根据正确的密钥 Z 进行解密,从而辨别出 Y' 是伪造的密段。因此,图中从接收者输出到信宿之间用虚线表示,说明接收用户拒绝了虚假的密段 Y' 。

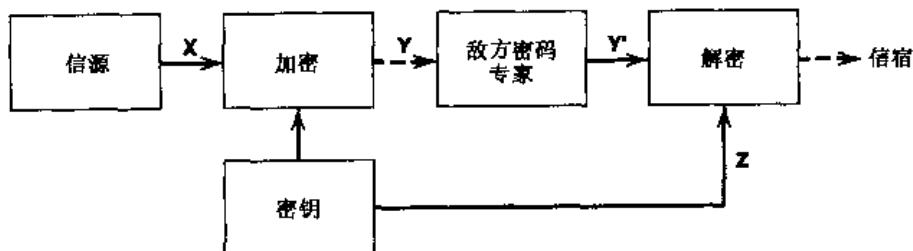


图 A5.2 敌方密码分析者入侵示意图

A5.2 块密码和流密码

和纠错编码可分为分组码和卷积码一样,加密系统(密码)也可以分成两大类:块密码和流密码。块密码完全是以组合方式来处理大量的明文块,而流密码是按小段(字符或比特)来处理明文的。

图 A5.3 给出了块密码的一般形式。明文(由串行数据组成)被分成了很多的大块,其中的每一块都由固定数目的比特构成。连续的明文块采用相同的密钥进行加密,否则,就是相互独立的。这样得到的加密块最后被转换为串行形式。所以,和前面的分块相同,一个特定的明文块可生成一个相应的密文块。特别地,一个特定密文块的每一个比特都是组成相应明文块和密钥的所有比特的函数。块密码的目的就是使明文中没有一个特定比特直接在密文中出现。

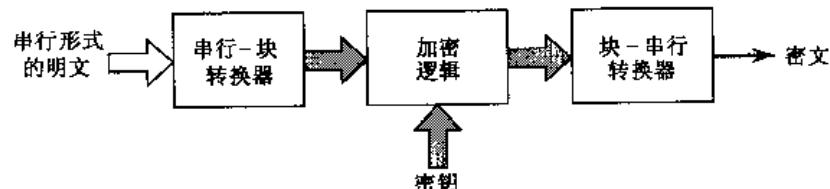


图 A5.3 块密码框图

块密码的操作是以块到块为基础的,并且对大量的明文数据块采用一个固定的转换。流密码则相反,它的操作是以应用到明文的不同比特的时变转换为基础的。最常用的流密码称为二进制加性流密码。其一般形式如图 A5.4 所示。在这种加密方式中,密钥被用来控制密钥流产生器,密钥流产生器输出的二进制序列称为密钥流,其长度远大于密钥。令 x_n , y_n 和 z_n 分别代表时刻 n 的明文比特、密文比特和密钥比特。密文比特此时仅由明文比特和密钥比特通过简单的模 2 加法来确定,即

$$y_n = x_n \oplus z_n, \quad n = 1, 2, \dots, N \quad (\text{A5.3})$$

其中, N 是密钥流的长度。因为模 2 加法和减法是完全相同的,因此有

$$x_n = y_n \oplus z_n, \quad n = 1, 2, \dots, N \quad (\text{A5.4})$$

于是,在二进制加法流密码中,相同的设备既可以用来加密,也可以用来解密,如图 A5.4 所示。密钥的选择可以根据某些概率分布进行。为了保证加密的安全性,密钥流应尽可能与掷硬币序列(完全随机)相似。

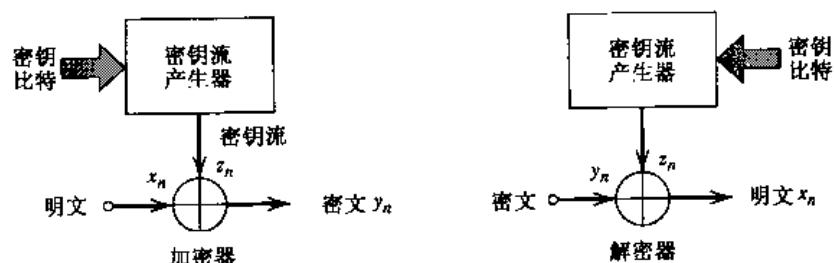


图 A5.4 二进制加法流密码

块密码一般采用这样的方式来设计: 输入的明文之中的一个小变化会引起输出结果的较大变化。块密码的这种误码传播特性在身份鉴别中是很有价值的,因为这种特性使得敌方的密码分析者不可能去改变加密数据,除非他们已经知道密钥。而二进制加法流密码没有误码传播特性,密文中失真比特的解密只会影响输出结果的相应比特。

流密码一般更适于在较易产生误码的信道上进行数据的安全传输,如应用于高数据速率或小传输时延的场合。^[5]

A5.2.1 保密要求

在密码学中,一个基本的假设是敌方密码分析者除了不知道密钥外,了解用于加密的所有机制。依靠附加知识的可用性,我们可以识别下面几种敌方密码分析者攻击时采用的方式:

1. **惟密文攻击**,这是一种密码分析型攻击。在这种攻击中,敌方密码分析者可以访问部分或全部密文。
2. **已知明文攻击**,这也是一种密码分析型攻击。在这种攻击中,敌方密码分析者已了解了部分由同一个实际密钥产生的密文和明文对。
3. **选择明文攻击**,这也是一种密码分析型攻击。在这种攻击中,敌方密码分析者可以提交任何选择好的明文,并在回复中接收到对应于实际密钥的密文。
4. **选择密文攻击**,也是一种密码分析型攻击。在这种攻击中,敌方密码分析者能选择任意一种密文,并且能找到其解密结果。

在实际应用中,惟密文攻击经常出现。在这种攻击形式中,敌方密码分析者只能了解所使用语言的统计结构(例如在英语中字母 e 的使用概率为 13%,而字母 q 后面总是紧跟着字母 u)和某些可能出现的词(例如一封信很可能是以“Dear Sir/Madam:”来开头)。已知明文攻击可以依靠编程语言和数据产生的标准计算机格式来进行。在任何情况下,惟密文攻击对加密系统的威胁都可认为是最小的,任何在这种攻击下不能正常工作的系统都可以认为是不可靠的。所以,对于一个可提供保密的加密系统来说,它至少应能抵抗惟密文攻击。实际上,它也应能抵抗已知明文攻击。

A5.3 信息论近似

在香农 1949 年的(关于保密系统的信息论近似的)划时代论文中公布的密码学香农模型中,敌方密码分析者被假设为具有无限的时间和计算能力。但实际上,敌人可能因为种种限制而只能采用惟密文攻击。在给定了密文以及不同明文和密钥的先验概率的情况下,香农模型的密码分析被定义为寻找密钥的过程。当敌方密码分析者能成功地进行解密,并获得惟一的密文解时,我们就认为这个保密系统已经被破坏了。^[6]

令 $\mathbf{X} = (X_1, X_2, \dots, X_N)$ 代表 N 比特的明文消息, $\mathbf{Y} = (Y_1, Y_2, \dots, Y_N)$ 代表相应的 N 比特的密文。明文和密文具有相同的比特数。设用来构造密文的密钥 \mathbf{Z} 是根据某些概率分布提出的。 \mathbf{X} 的不确定性用熵 $H(\mathbf{X})$ 来表示,已知 \mathbf{Y} 后, \mathbf{X} 的不确定性用条件熵 $H(\mathbf{X}|\mathbf{Y})$ 来表示。 \mathbf{X} 和 \mathbf{Y} 之间的互信息定义为

$$I(\mathbf{X}; \mathbf{Y}) = H(\mathbf{X}) - H(\mathbf{X}|\mathbf{Y}) \quad (\text{A5.5})$$

互信息 $I(\mathbf{X}; \mathbf{Y})$ 为香农模型中安全性的一个基本度量。

A5.3.1 完善的保密性

如果敌方密码分析者只能得到密文 \mathbf{Y} ,则可以定义这个加密系统具有完善的保密性,也就是说,明文 \mathbf{X} 和密文 \mathbf{Y} 之间是统计独立的。换言之,有

$$I(\mathbf{X}; \mathbf{Y}) = 0 \quad (\text{A5.6})$$

由式(A5.5), 可将完善的保密性条件写为

$$H(\mathbf{X} | \mathbf{Y}) = H(\mathbf{X}) \quad (\text{A5.7})$$

式(A5.7)说明, 敌方密码分析者最多只能在得到密文 \mathbf{Y} 时, 根据所有可能消息的概率分布来猜测明文消息 \mathbf{X} 。

假设密钥为 \mathbf{Z} , 则

$$\begin{aligned} H(\mathbf{X} | \mathbf{Y}) &\leq H(\mathbf{X}, \mathbf{Z} | \mathbf{Y}) \\ &= H(\mathbf{Z} | \mathbf{Y}) + H(\mathbf{X} | \mathbf{Y}, \mathbf{Z}) \end{aligned} \quad (\text{A5.8})$$

当且仅当 \mathbf{Y} 和 \mathbf{Z} 一起惟一确定 \mathbf{X} 时, 条件熵 $H(\mathbf{X} | \mathbf{Y}, \mathbf{Z})$ 为零。如果解密过程是在已知 \mathbf{Z} 的情况下进行, 这个假设实际上是有效的。因此, 式(A5.8)可以简化为

$$\begin{aligned} H(\mathbf{X} | \mathbf{Y}) &\leq H(\mathbf{Z} | \mathbf{Y}) \\ &\leq H(\mathbf{Z}) \end{aligned} \quad (\text{A5.9})$$

将式(A5.9)代入到式(A5.7)可得, 一个加密系统要提供完善的保密性, 就必须满足如下条件

$$H(\mathbf{Z}) \geq H(\mathbf{X}) \quad (\text{A5.10})$$

不等式(A5.10)是完善保密性的香农基本边界。这说明, 要得到完善的保密性, 密钥 \mathbf{Z} 的不确定性必须至少等于被密钥 \mathbf{X} 隐藏的明文的不确定性。

当明文表和密钥表大小相同时, 用完善保密性的香农边界可以得到如下结果: 密钥必须至少与明文一样长。从这个结果中可以得出的结论是, 构建一个完善保密性的加密系统所需要的密钥长度, 对于大多数应用来说, 都是大到了不可能实现的程度。不过, 完善的保密性还是有一定的应用场合的, 它可以用于(可能的)消息数较少的情况或(最重要的)需要有完善保密性的场合。

一种著名的完善保密密码是一次填充密码^[7], 有时也称为佛纳姆(Vernam)密码, 这种密码一般用在非通用场合, 如有高度机密性要求的两用户的热线通信。一次填充密码是一种流密码, 密钥同样也是一个密钥流, 如图 A5.5 所示。输入端由两个部分组成, 一个由消息序列比特 $\{x_n | n = 1, 2, \dots\}$ 代表的消息, 以及一个由统计独立的均匀分布比特 $\{z_n | n = 1, 2, \dots\}$ 构成的序列代表的密钥。合成的密码 $\{y_n | n = 1, 2, \dots\}$ 是由两个输入序列经模 2 加法器得到的, 即

$$y_n = x_n \oplus z_n, \quad n = 1, 2, \dots$$

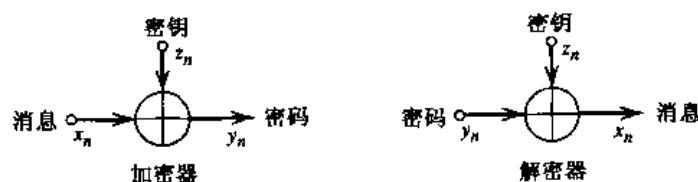


图 A5.5 一次性密码(佛纳姆密码)

例如二进制消息序列为 00011010, 二进制密钥序列为 01101001。这两个序列的模 2 加可以写为

消息: 00011010
 密钥: 01101001
 密码: 01110011

这里描述的加密规则中,密钥比特 1 使消息序列中的 0 和 1 互换,而密钥比特 0 保持消息比特不变。只需将密码序列和密钥序列进行模 2 加,就能够恢复出消息序列,即

密码: 01110011
 密钥: 01101001
 消息: 00011010

一次填充密码具有完善的保密性,因为消息与密码之间的相互信息为 0,因此,这种密码是完全不可破译的。

A5.3.2 单一距离

假设实际上采用了一种不完善的密码,那么,敌方密码分析者何时能够破译这个密码? 随着窃听内容的增加,可直观地认为,拥有无限时间和计算能力的敌方密码分析者总会找到密钥并破译密码。在香农模型中,存在一个临界点,称为单一距离。单一距离的正式定义为,使条件熵 $H(\mathbf{Z} | Y_1, Y_2, \dots, Y_N)$ 近似等于 0 的最小 N 值。对于“随机密码”的特殊情况,单一距离可以由下式近似给出:^[8]

$$N_0 \approx \frac{H(\mathbf{Z})}{r \log L_y} \quad (\text{A5.11})$$

其中, $H(\mathbf{Z})$ 是密钥 \mathbf{Z} 的熵, L_y 是密文表的大小。参数 r 是 N 比特密文中消息信息的冗余百分比,其定义为

$$r = 1 - \frac{H(\mathbf{X})}{N \log L_x} \quad (\text{A5.12})$$

其中, $H(\mathbf{X})$ 是明文 \mathbf{X} 的熵。在大多数加密系统中,密文表的大小 L_y 与各明文表的大小 L_x 是相等的。在这种情况下, r 正好就等于明文自身的冗余百分比。虽然式(A5.11)是在一个特定的、假设“随机密码”的基础上得到的,但也可用来估计普通类型密码的单一距离,实际上也都是这样做的。

假设 K 是在大小为 L_z 的码表中选定的密钥 \mathbf{Z} 的数位数,则密钥 \mathbf{Z} 的熵可以表示成

$$H(\mathbf{Z}) \leq \log(L_z^K) = K \log L_z \quad (\text{A5.13})$$

当且仅当密钥的选择是完全随机时,上式取等号。令密钥表的大小 L_z 和密码表的大小 L_y 相等,且密钥的选择是完全随机的,则此时单一距离最大。将式(A5.13)取等号并代入式(A5.11),可得到简单的结果

$$N_0 \approx \frac{K}{r} \quad (\text{A5.14})$$

下面举例说明式(A5.14)。考虑一个加密系统,其中 $L_x = L_y = L_z$,该系统被用来对英文文本进行加密。典型的英文文本的冗余百分比 r 为 75%。因此,根据式(A5.14),敌方密码分析者只要窃听到密码数据的 1.333 K 比特就可以破译这个密码了。

但是,一个潜在可破译的不完善密码仍然有其实用价值,这一点也是很重要的。当窃听的密文包含了足以满足式(A5.11)的信息时,也不能保证敌方密码分析者就真的能破译这个密码,因为他们只有有限的计算资源。特别地,对于以这种方式设计的密码,很可能会由于密码分析的任务太重,以至于所有的计算资源被一点点耗尽而不能破译,尽管大家都知道通过有限次的计算就可以得到结果。在这种情况下,这种不完善的密码可以认为是计算安全的。

A5.3.3 密码学中数据压缩的作用

无损的数据压缩是密码学中的有用工具。这么说是因为数据压缩去除了数据中的冗余,从而增加了单一距离 N_0 ,见式(A5.11)。利用这个思想,在发射机对信息进行加密之前,先进行数据压缩,而在接收机解密之后再将冗余重新插入。最后的结果是,接收机输出端的授权用户不会发现差错,同时,传输的信息变得更加安全。采用完善的压缩技术能够去除所有的冗余信息,从而将消息源转变成完全随机的,并使得对于任何大小的密钥都有 $N_0 = \infty$,这的確是很吸引人的。遗憾的是,现在还不存在这样的对现实消息源进行完善压缩的设备,而且将来也不可能出现。因此,想仅靠数据压缩来实现数据安全是不可能的。不过,有限的数据压缩的确可以使安全性增加,这就是加密者将数据压缩视为有效手段的原因。

A5.3.4 扩散与混淆

在密码学的香农模型中,有两个方法可以作为指导实际加密者的设计通用原则。这两个方法称为扩散与混淆。这两种方法(单独使用或同时使用)的目的就是要阻挠敌人对密文进行统计分析,从而使密码的破译变得极其困难。

扩散方法就是通过将明文中单个比特的影响扩展到密文中的大量比特中,从而隐藏明文中的统计结构。这种扩展可以迫使敌人为确定明文的统计结构而不得不窃听大量的材料,因为这种结构只有在许多信息块中才会显现,而每个这样的信息块出现的概率又很小。在混淆方法中,数据转换被设计成使密文统计对明文统计的依赖方式变得复杂化。因此,一个好的加密者一般是将扩散与混淆结合起来使用。

然而,一个有实用价值的密码,必须使得敌方密码分析者破译起来非常困难,而在给定密钥的情况下又能够很容易地进行加密解密。采用基于“分治”概念的乘积密码,就可以满足这两个设计要求。特别地,一个强密码的实现可以由简单的部分密码连接起来得到,其中每个部分都进行适当的扩散和混淆来组成整个密码。乘积密码通常以代换密码和转置密码作为其基本成分,接下来就介绍这些简单密码。

1. 代换密码

在代换密码中,明文中的每个字母都有一个固定的替代,这个替代通常是来自同一字母表中的字母,具体的替代规则由密钥决定。于是,明文为

$$\mathbf{X} = (x_1, x_2, x_3, x_4, \dots)$$

其中, x_1, x_2, x_3, \dots 是连续的字母,变换得到的密文为

$$\begin{aligned}\mathbf{Y} &= (y_1, y_2, y_3, y_4, \dots) \\ &= (f(x_1), f(x_2), f(x_3), f(x_4), \dots)\end{aligned}\tag{A5.15}$$

其中, $f(\cdot)$ 是一个可逆的函数。当替代是字母时, 密钥是一个字母表的置换。例如图 A5.6 所示的密文表, 第一个字母 Y 代替 A , 第二个字母 D 代替 B , 以此类推。使用代换密码的结果是产生了混淆。

明文字母 ABCDEFGHIJKLMNOPQRSTUVWXYZ

密文字母 YDUBHNACSVXELPFMKQJRWGOZIT

图 A5.6 代换密码

2. 转置密码

在转置密码中, 明文被分成具有固定周期 d 的分组, 并对每个分组采用同样的置换, 具体的置换规则由密钥决定。例如图 A5.7 所示的置换规则, 其中周期 $d = 4$ 。根据这种密码, 字母 x_1 由明文中的位置 1 移到了密文中的位置 4。于是明文

$$\mathbf{X} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, \dots)$$

变换成密文

$$\mathbf{Y} = (x_3, x_4, x_2, x_1, x_7, x_8, x_6, x_5, \dots)$$

虽然密文 \mathbf{Y} 中单个字母的统计等于明文 \mathbf{X} 中单个字母的统计, 但是, 更高阶的统计发生了变化。使用转置密码的结果是产生了扩展。

将简单的代换和转置交织在一起, 并多次进行这种交织过程, 构造一个具有好的扩展和混淆的强码是有可能的。

明文字母 $x_1 \ x_2 \ x_3 \ x_4$

密文字母 $x_3 \ x_4 \ x_2 \ x_1$

图 A5.7 转置密码

例 A5.1

设明文消息为

THE KING IS DEAD LONG LIVE THE KING

使用图 A5.6 中描述的改序的字母表来进行代换加密, 则明文变成密文

RCHXSPASJBHYBEFPAESGHRCRXSPA

接下来采用图 A5.7 所示的置换规则来进行转置加密, 于是, 由代换加密得到的密文再次转变为

HXRASPSSHYBJFBEBSGEACHRHPASX

这与原来的明文已经没有相似之处。

A5.4 数据加密标准

数据加密标准(DES)^[9]无疑是最知名且使用最广泛的密钥加密算法。“算法”一词是用来描述一系列计算的。基本的 DES 算法可以用于数据加密和数据认证。DES 算法是数据存储和邮

件系统、电子转账(批发和零售)和电子商务数据交换的标准加密算法。

DES 算法是一种采用 64 比特明文数据块和 56 比特密钥的有效块密码。这种算法是依据扩散和混淆的香农模型来设计的, 加密和解密必须采用相同的算法。DES 算法采用的所有转换可以记为 $P^{-1}\{F[P(X)]\}$, 其中, X 是指明文, P 是某种置换, 函数 F 是代换和转置的结合。函数 F 可以由某个函数 f 级联而成, 级联的每一级可看成是一轮。

图 A5.8 给出了 DES 算法加密的详细情况。经过一次初始置换后, 64 比特的明文被分成了左边的一半 L_0 和右边的一半 R_0 , 每一半都是 32 比特。然后, 算法执行 16 轮与密钥相关的计算, 其中, 第 i 轮计算为

$$L_i = R_{i-1} \quad i = 1, 2, \dots, 16 \quad (\text{A5.16})$$

$$R_i = L_{i-1} \oplus f(R_{i-1}, Z_i) \quad i = 1, 2, \dots, 16 \quad (\text{A5.17})$$

在式(A5.17)中等号的右边, 加法是模 2 运算。 Z_i 用于第 i 轮的 48 比特密钥块。函数 $f(\cdot, \cdot)$ 是一个输出为 32 比特的函数。第 16 轮的结果被保存, 得到序列 $R_{16} L_{16}$ 。将这 32 比特序列送入最后的置换运算 P^{-1} , 生成 64 比特的密文, 目的就是经过 16 轮与密钥相关的计算后, 在密文中不可能检测出原始明文的图样。由式(A5.16)和式(A5.17)可知, 解密函数 $f(\cdot, \cdot)$ 并非必须可逆, 因为 (L_{i-1}, R_{i-1}) 可以由 (L_i, R_i) 按照下式恢复出来:

$$R_{i-1} = L_i \quad i = 1, 2, \dots, 16 \quad (\text{A5.18})$$

$$L_{i-1} = R_i \oplus f(L_i, Z_i) \quad i = 1, 2, \dots, 16 \quad (\text{A5.19})$$

即使 $f(\cdot, \cdot)$ 是多对一的函数(其逆不惟一), 式(A5.19)照样成立。

图 A5.9 给出了计算函数 $f(\cdot, \cdot)$ 的流程图。首先, 将连续的 4 比特码字的边缘比特(比特编号为 1, 4, 5, 8, 9, 12, 13, 16, ..., 28, 29, 32)进行重复, 从而将 32 比特块 R 扩展为一个新的 48 比特块 R' 。于是, 如果设 32 比特块 R 为

$$R = \underbrace{r_1 r_2 r_3 r_4}_{\substack{\text{第1块} \\ 4 \text{ 比特码字}}} \quad \underbrace{r_5 r_6 r_7 r_8}_{\substack{\text{第2块} \\ 4 \text{ 比特码字}}} \quad \cdots \quad \underbrace{r_{29} r_{30} r_{31} r_{32}}_{\substack{\text{第8块} \\ 4 \text{ 比特码字}}}$$

则可将扩展得到的 48 比特块 R' 构造成

$$R' = \underbrace{r_{32} r_1 r_2 r_3 r_4 r_5}_{\substack{\text{第1块} \\ 6 \text{ 比特码字}}} \quad \underbrace{r_4 r_5 r_6 r_7 r_8 r_9}_{\substack{\text{第2块} \\ 6 \text{ 比特码字}}} \quad \cdots \quad \underbrace{r_{28} r_{29} r_{30} r_{31} r_{32} r_1}_{\substack{\text{第8块} \\ 6 \text{ 比特码字}}}$$

将 48 比特块 R' 和 Z_i 进行模 2 加, 结果为 8 个 6 比特码字。令这些码字分别由 B_1, B_2, \dots, B_8 代表, 可记为

$$B_1 B_2 \cdots B_8 = R' \oplus Z_i \quad (\text{A5.20})$$

每个 6 比特的码字 B_i 被送入查表形式的替代盒, 生成 4 比特的输出 $S_i(B_i)$ 。替代盒 $S_i(B_i)$ 的每个输出比特都是 6 比特码字 B_i 的布尔函数。8 个输出 $S_1(B_1), S_2(B_2), \dots, S_8(B_8)$ 被排列成一个 32 比特块, 并输入到置换盒 $P[\cdot]$ 中。于是, 置换盒的输出就产生了所需的 32 比特函数 $f(R, Z_i)$, 记为

$$f(R, Z_i) = P[S_1(B_1)S_2(B_2) \cdots S_8(B_8)] \quad (\text{A5.21})$$

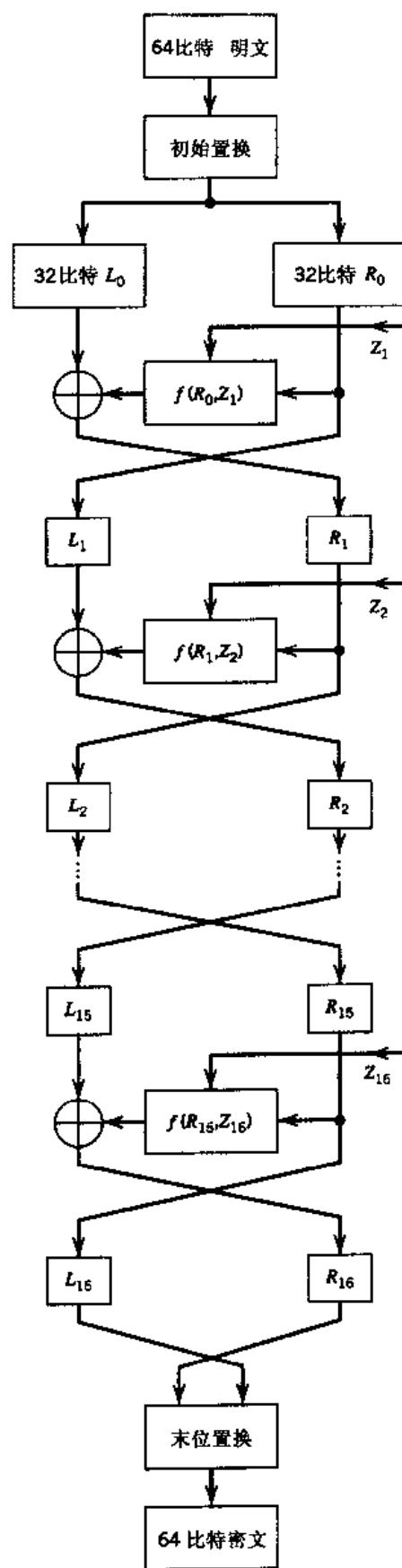


图 A5.8 DES(经 IEEE 允许摘自 Diffie 和 Hellman, 1979)

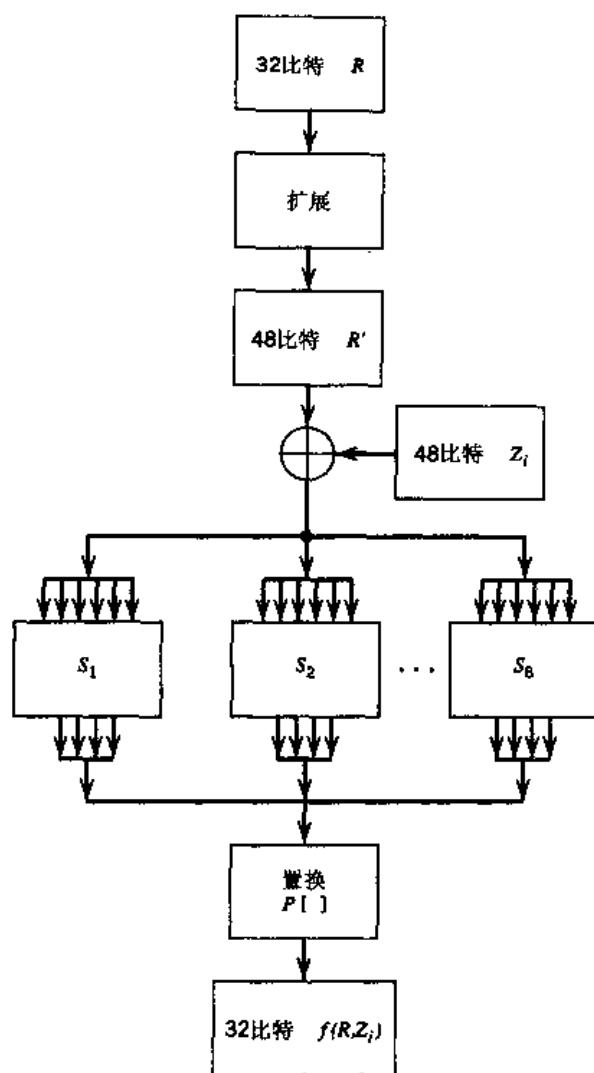


图 A5.9 $f(R, K)$ 流程图(经 IEEE 允许摘自 Diffie 和 Hellman, 1979)

第 i 次迭代中, 48 比特块 Z_i 使用的是 64 比特密钥 Z_0 的不同子集。确定每个 Z_i 的过程称为密钥表计算, 其流程图如图 A5.10 所示。密钥 Z_0 有 8 个奇偶校验比特, 分布在位置 8, 16, ..., 64 上, 分别用于各自 8 比特字节的误码检测。相关的误码可在密钥 Z_0 的生成、分配和存储中产生。置换选择 1 丢弃掉 Z_0 中的奇偶校验比特, 并对保留的 56 比特进行重新排序, 然后, 将它们分别送入两个 28 比特的移位寄存器, 每个移位寄存器有 24 个时延。两个移位寄存器的 48 个时延用于 16 次计算的迭代, 每次迭代都包含一次或两次循环左移, 然后进行置换排序。这个置换排序称为置换选择 2。这些 16 次迭代产生的输出, 提供了不同密钥的 48 比特块 Z_1, Z_2, \dots, Z_{16} , 分别用于第 1, 2, ..., 16 次迭代中。

尽管所有问题都是一分为二的, 但是, 目前尚没有人可以证明 DES 算法有什么缺陷。虽然围绕着 DES 算法的使用有很多争论, 但是, DES 算法的最大贡献也许就在于它有助于提高人们对加密技术的兴趣, 并以此作为计算机网络安全的一种机制。

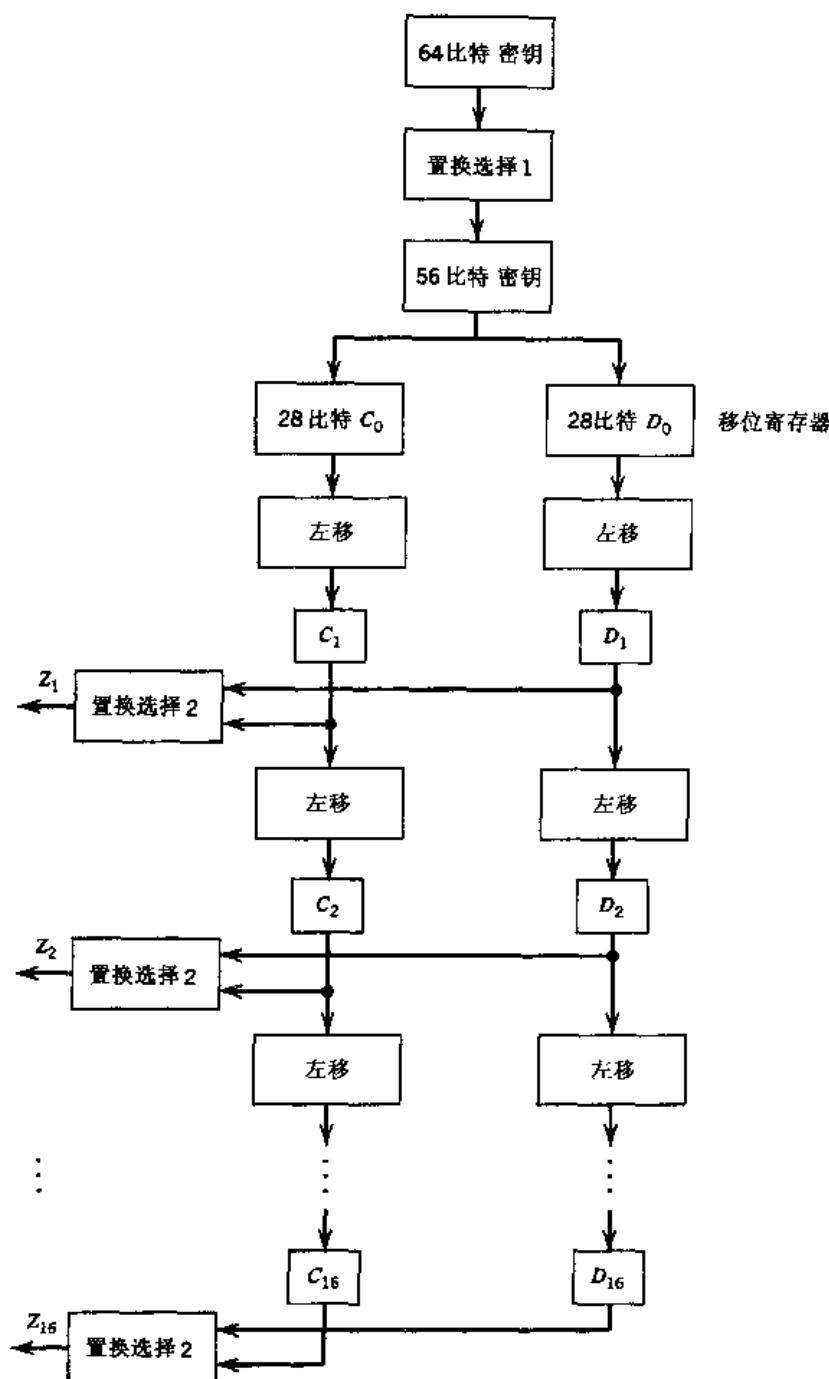


图 A5.10 密钥表计算流程图(经 IEEE 允许摘自 Diffie 和 Hellman, 1979)

A5.5 公钥密码^[10]

对于试图通过不安全信道进行加密通信的两个用户来说，在通信之前进行密钥交换是必要的。所有的加密系统都要求在授权用户之间进行安全的密钥分配。在传统的加密技术中，用户采用物理的安全信道(注册邮件或信使服务)进行密钥分配。然而，这种辅助信道的使用引出了传统加密的一个主要局限。显然，用信使服务和注册邮件来分配密钥，既昂贵又不方便。

便,还会降低带宽和速率。此外,这种方法也不总是安全的。

在大规模的通信网络中,特别强调密钥的分配问题,因为 n 个用户会使可能的连接数上升为 $(n^2 - n)/2$ 个。当 n 较大时,密钥分配的成本变得非常高。因此,随着大规模安全通信网络的发展,我们不得不依靠不安全信道来进行相互密钥信息的交换,然后再进行安全通信。这个限制就提出了一个基本问题:在不安全信道上怎样安全地进行密钥信息的交换?在公钥密码中,这个看起来困难的问题通过将部分密钥信息公开而得到了解决,并由此大大地简化了密钥管理的任务。这与传统的密码技术是完全不同的,在传统的加密技术中,密钥是完全对敌方密码分析者保密的。

公钥密码系统可以用两个计算可逆函数(变换)的算法集来表示。这两个算法集分别用 $\{E_z\}$ 和 $\{D_z\}$ 表示,即都用 z 来标记。由这个算法计算得到的可逆变换可以记为

$$E_z: f_z(x) = y \quad (A5.22)$$

$$D_z: f_z^{-1}(y) = x \quad (A5.23)$$

其中, x 为用 z 标记的函数 f_z 定义域中的某个输入信息, y 为 f_z 的值域内的对应密码。系统的基本要求是,函数 f_z 必须是一个陷门单向函数。“单向”一词是指,对于 f_z 的定义域中的 x ,由已知的算法 E_z 必须能够简单地计算出 $f_z(x)$,但对于值域中的某个密码 y ,敌方密码分析者要计算逆函数 $f_z^{-1}(y)$ 却极其困难。而对于拥有相关算法 D_z 的授权用户来说,计算逆函数 $f_z^{-1}(y)$ 则比较容易。所以,秘密密钥(算法) D_z 提供了一个“陷门”,使得从密码分析者的角度来看,求函数 f_z 的逆显得极其困难,但对 D_z 的拥有者(惟一授权)来说又相当容易。由于仅依靠密钥(算法) E_z 本身并不能计算出函数 f_z 的逆,所以可以将其公开,因此称为“公钥密码”。

公钥密码系统的概念是密钥变成了互逆对(公开密钥和秘密密钥),每个密钥对都具有两个基本性质:

1. 无论消息用哪个密钥加密,都可以用另一个密钥解密。
2. 已知公开密钥,要找到密钥,从计算角度看仍是不可实现的。

使用这里描述的公钥密码可以解决保密问题。例如在某一安全通信系统中,用户都将其公开密钥按姓名和地址列入一个“电话目录”中。当一个用户要向另一个用户发送秘密消息时,只需查找那个用户的地址对应的公开密钥,并用这个密钥对消息进行加密即可。加密的消息(密文)只能由指定的公开密钥的持有者阅读。实际上,原始消息(明文)已经丢失了,即使是消息的发送者,要从密文中恢复出原来的消息也是非常困难的。

公钥密码的密钥管理使得它特别适用于大规模安全通信网络。实际上,公钥密码已由一个简单的概念发展成为加密技术的一个主流。

A5.5.1 Diffie-Hellman 公开密钥分配

Diffie-Hellman 公钥分配系统是一种简单而又性能良好的系统。这个系统的设计思想来源于这样一个事实:计算离散指数比较容易,而计算离散对数则十分困难。具体地说,假设有一个离散指数函数

$$Y = \alpha^X \bmod p \quad \text{对于 } 1 \leq X \leq p - 1 \quad (A5.24)$$

其中,计算是按模 p 进行的。 α 是一个本原整数(α 的所有次幂生成与 $p - 1$ 互质的模 p 的

所有元素)。对应地, X 就是指模 p 的以 α 为底的 Y 的离散对数:

$$X = \log_{\alpha} Y \bmod p \quad \text{对于 } 1 \leq Y \leq p - 1 \quad (\text{A5.25})$$

由 X 计算 Y 是比较容易的, 可以使用平方与乘积的方法。例如, 当 $X = 16$ 时, 有

$$Y = \alpha^{16} = [(\alpha^2)^2]^2$$

而由 Y 来计算 X 却是非常困难的。

在 Diffie-Hellman 公钥分配系统中, 假设所有用户都已知 α 和 p 。如果用户 i 从整数集 $\{1, 2, \dots, p\}$ 中机会均等地随机选择了独立数 X_i 作为秘密密钥。但离散指数

$$Y_i = \alpha^{X_i} \bmod p \quad (\text{A5.26})$$

与用户的姓名和地址一起存入了公共目录中。系统的所有其他用户也如此操作。现在, 假设用户 i 和 j 需要进行秘密通信。为此, 用户 i 从公共目录中取出 Y_j , 并用秘密密钥 X_i 计算

$$\begin{aligned} K_{ji} &= (Y_j)^{X_i} \bmod p \\ &= (\alpha^{X_j})^{X_i} \bmod p \\ &= \alpha^{X_j X_i} \bmod p \end{aligned} \quad (\text{A5.27})$$

与此类似, 用户 j 也计算 K_{ij} 。于是可得到

$$K_{ji} = K_{ij} \quad (\text{A5.28})$$

从而, 用户 i 和 j 一起得到了 K_{ji} , 并以此作为传统加密系统中的密钥。而其他的用户则必须在公共目录中查到 Y_i 和 Y_j , 并利用替换公式

$$K_{ji} = (Y_j)^{\log_{\alpha} Y_i} \bmod p \quad (\text{A5.29})$$

来计算 K_{ji} 。显然, 敌方没有其他的方法可以用来计算 K_{ji} (但这一点我们还不能证明)。如前所述, 式(A5.29)计算起来很困难, 这是因为其中包含了离散的对数运算, 而式(A5.27)计算起来就比较容易, 因为其中只包含了离散的指数运算。因此, 系统的安全性是依靠离散对数计算的困难性来得到的。

Diffie-Hellman 公钥分配系统是这一类系统中最老的一种, 但仍然被认为是最安全且实用的公钥分配系统之一。

A5.6 RSA 系统

要开发一种公钥密码系统, 并不是件容易的事。实际上, 各种文献中已提出了大量这样的系统, 但遗憾的是, 它们中的大多数都已证明是不安全的。到目前为止, 最成功地实现了公钥密码的是 RSA(Rivest-Shamir-Adleman)系统^[11], 该系统采用了经典数论的思想。RSA 系统之所以被认为是最安全的密码系统之一, 是因为它已经经受住了许多此领域的专家对它的攻击。

RSA 算法是一个块密码, 它是基于这样一个事实, 即寻找一个随机的大素数(例如 100 位)比较容易计算, 而要将两个这样的数的乘积进行因子分解, 目前从计算的角度看是不可行的。特别地, 对于 RSA 算法, 指定参数的计算按如下步骤进行:

1. 随机地选择两个非常大的素数 p 和 q 。素数必须均等且仔细地选择, 因为有些素数会

使系统非常脆弱。

2. 将 p 和 q 相乘, 得到乘积

$$pq = n \quad (\text{A5.30})$$

可用下式来寻找 n 的欧拉 ϕ 函数

$$\phi(n) = (p - 1)(q - 1) \quad (\text{A5.31})$$

式(A5.31)是根据欧拉 ϕ 函数的定义得到的。 $\phi(n)$ 是小于 n 的正整数 i 的个数, 其中, i 和 n 的最大公因子为 1。

3. 令 e 为小于 $\phi(n)$ 的正整数, e 和 $\phi(n)$ 的最大公因子为 1。寻找一个小于 $\phi(n)$ 的正整数 d , 使得

$$de = 1 \pmod{\phi(n)} \quad (\text{A5.32})$$

于是, RSA 的陷门单向函数可简单地由离散指数运算定义为

$$f_z(x) = x^e = y \pmod{n} \quad (\text{A5.33})$$

n 和 e 的值构成公开密钥, 因此, 公开易于寻找的计算函数 f_z 的算法 E_z , 相当于分开了 n 和 e 。

素数 p 和 q 组成了密钥。由于 d 是与 p 和 q 相关的, 因此, 拥有了易于寻找的(当一个用户已知陷门 z 时)用来计算逆函数 f_z^{-1} 的算法 D_z , 就相当于知道了 p 和 q 。特别地, 逆函数定义为

$$f_z^{-1}(y) = y^d \pmod{n} \quad (\text{A5.34})$$

解密指数 d 可由式(A5.32)求得, 在通常的整数运算中, 对于一些整数 Q , 式(A5.32)等价于

$$de = \phi(n)Q + 1 \quad (\text{A5.35})$$

值得注意的是, 由式(A5.31)可知, $\phi(n)$ 也与 p 和 q 有关。由于 $y = x^e$, 则可由式(A5.32)和式(A5.33)得到

$$\begin{aligned} y^d &= x^{de} \\ &= x^{\phi(n)Q+1} \\ &= ((x^{\phi(n)})^Q)x \end{aligned} \quad (\text{A5.36})$$

下面应用著名的欧拉(Euler)定理, 即对于任何的正整数 x 和 n , 其中 $x < n$, 有

$$x^{\phi(n)} = 1 \pmod{n} \quad (\text{A5.37})$$

因此, 将式(A5.37)代入式(A5.36), 就可得到所需的解密结果

$$y^d = x \quad (\text{A5.38})$$

这样, 在给定素数 p 和 q 后, 寻找逆函数 f_z^{-1} 就变得容易了。

RSA 加密算法的安全性依赖于一个前提, 即函数 f_z 的求逆等价于分解 $n = pq$ 。这个等价引出了一个问题: 通过分解 n 来攻击, 在计算上可能实现吗? 如果素数 p 和 q 都在 100 位十进制数的数量级上, 并且因子分解没有革命性的突破, 则答案显然是否定的。

A5.6.1 数字签名^[12]

为了在商务中用电子邮件系统代替普通的纸张邮件, 系统的用户必须能够签署电子信息。

使用数字签名可以为某个消息确实来源于某个发送者提供证据。为了满足这个条件,数字签名必须具备以下性质:

- 电子信息的接收者能验证发送者的签名。
- 签名是不可伪造的。
- 电子签名的发送者不能否认这个签名。

用 RSA 算法实现数字签名的过程如下。拥有密钥 d 的用户可以对一个已知的消息块 m 用下面的信号来签署电子信息:

$$s = m^d \bmod n \quad (\text{A5.39})$$

如果不知道秘密密钥 d ,要计算出 s 是很困难的。因此,由式(A5.39)定义的数字签名是很难伪造的。此外,消息 m 的发送者不能否认是他发送了这个信息,因为别人无法产生信号 s 。接收者用公开密钥 e 来计算

$$\begin{aligned} s^e &= (m^d)^e \bmod n \\ &= m^{de} \bmod n \\ &= m \bmod n \end{aligned} \quad (\text{A5.40})$$

上式最后一行是由式(A5.32)得到的。因此,通过 $s^e \pmod{n}$ 的计算,能够产生与解密消息 m 相同的结果,接收者可以验证发送者的签名。这样,RSA 算法就满足了数字签名所有的必须性质。

A5.7 总结与讨论

密码学是一个很“热门”的研究领域。这样说不是为了哗众取宠。在这样一个信息时代,作为一种安全机制,密码学的重要性将越来越明显。在本附录中,我们已经对这个高度重要的课题进行了介绍性的论述。

根据信息加密和解密中所用密钥是完全秘密还是部分公开,可将密码学分为秘密密钥密码学和公开密钥密码学。另外,根据实现的方式,又可将密码系统分为块密码和流密码。块密码有误码传播特性,这一点在认证过程具有很高的应用价值。

密码学发展到今天,在许多密码系统中,DES 和 RSA 算法仍然是最成功的算法。这两个加密算法都是块密码。它们的不同之处在于,DES 算法采用的是秘密密钥,而 RSA 算法采用的是公开密钥。在秘密密钥系统中,相同密钥被发送者与接收者双方共享。而在公开密钥系统中,密钥被分成了两部分其中公开密钥位于发送方,秘密密钥位于接收方。对于公开密钥系统,在没有秘密密钥的情况下,要从加密信息中恢复出明文信息,从计算的角度讲是不可能的。

虽然像 RSA 这样的公开密钥加密系统提供了有效的密钥管理方式,但大量的加密信息会由于低带宽而使效率降低。相反,像 DES 这样的传统加密系统却能够提供更好的吞吐量,但其中需要密钥管理。这就暗示着可以采用混合的方法,充分利用两个加密系统的优点,并以此作为实际应用中设计安全通信系统的基础。例如,将 RSA 算法用于验证,而将 DES 算法用于加密。

注释与参考

- [1] 对密码学的介绍性论述可参阅 Adámek(1991)的第 15 章。对密码学的多侧面全面论述可参阅 Simmons(1992), 该书是 IEEE(1949)在密码学方面的论文集特刊的扩充版。书中相关章节的稿件由研究密码学课题的前沿权威人士提供。van Tilborg(1988)中也给出了密码学的较好论述。
- [2] Shannon(1949)发表的具有里程碑性质的论文, 开创了科学的秘密密钥密码学的新纪元, 也在密码学与信息论之间建立起了联系。
- [3] Diffie 和 Hellman(1976)发表的另一篇具有里程碑性质的论文, 开创了公开密钥密码学的新纪元, 第一次证明了在发送者和接收者之间不进行任何密钥传递就可以进行保密通信是可能的。正是这篇论文引发了人们研究密码学的兴趣, 这个兴趣一直延续至今。
- [4] 在密码学中常用的“敌方密码分析者”一词是指密码的破译者, 这个应用是起源于军事领域。
- [5] 关于流密码的全面论述参阅 Simmons(1992)编写的“Contemporary Cryptology”一书中由 R. A. Rueppel 所著的第 2 章。
- [6] 关于密码学香农模型的更通俗的论述, 可参阅 Simmons(1992)编撰的书中由 Massey 所著的第 1 章。
- [7] 一次填充密码这个名称起源于(不久之前, 即二战期间及以后)几个国家的间谍对它的使用, 他们常收到一个写有随机选择的密钥的纸垫, 并被告知这个密钥只能用于一次加密。一次填充密码又称为佛纳姆密码, 这是以其发明者 G. S. Vernam 的名字来命名的。
- [8] 式(A5.11)可参阅香农(1949)的原文。
- [9] 在Simmons(1992)的第1章中, M. E. Smid 和 D. K. Branstad 对 DES 算法的历史进行了重新叙述。DES 算法的详细描述可参阅 Diffie 和 Hellman(1979)的论文。也可参阅 Meyer, Matyas(1982)和 Torrieri(1992 Chapter 6)。
- [10] 对公开密钥密码学的全面论述, 可参阅 Simmons(1992)中由 J. Nechvatal 所著的第 4 章。书中还包括由 W. Diffie 执笔的一章, 其中描述了几种关于公开密钥密码算法的设想和相应协议的逐渐演变。
- [11] RSA 算法已获得专利, 并以其发明者 R. L. Rivest, A. Shamir 和 L. Adleman 的名字命名。关于 RSA 算法的最早文献是 Rivest, Shamir 和 Adleman(1978)的论文。
- [12] 数字签名的思想是由Diffie和Hellman(1976)首先提出的。该思想的 RSA 算法实现在 Rivset, Shamir 和 Adleman(1978)的论文中曾有描述。关于数字签名的详细论述, 可参阅 Simmons(1992)中由 C. J. Michell, F. Piper 和 R. Wild 编写的第 6 章。

附录 6 表 格

汇编在附录中的 12 个表格涵盖下列内容：

- ASCII 码表
- 傅里叶和希尔伯特变换
- 贝塞尔函数
- 误差函数
- 调制解调器标准选编
- 三角恒等式、级数展开式和积分
- 常用的常数和推荐的单位前缀

表 A6.1 ASCII 码

比特位置				7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0	7	6	5	4	3	2	1	0
				0	0	0	0	0	0	1	1	0	1	0	0	1	0	0	1	1	0	1	0	1	1	0	1
0	0	0	0	NUL	DLE	SP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	SOH	DC1	!	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	1	0	STX	DC2	"	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
0	0	1	1	ETX	DC3	#	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
0	1	0	0	EOT	DC4	\$	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
0	1	0	1	ENQ	NAK	%	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
0	1	1	0	ACK	SYN	&	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
0	1	1	1	BEL	ETB	'	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
1	0	0	0	BS	CAN	(8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
1	0	0	1	HT	EM)	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
1	0	1	0	LF	SUB	*	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
1	0	1	1	VT	ESC	+	;	;	;	;	;	;	;	;	;	;	;	;	;	;	;	;	;	;	;	;	;
1	1	0	0	FF	FS	,	<	<	<	<	<	<	<	<	<	<	<	<	<	<	<	<	<	<	<	<	<
1	1	0	1	CR	GS	-	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	
1	1	1	0	SO	RS	.	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>	>
1	1	1	1	SI	US	/	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?

ACK	应答	ENQ	询问	NUL	空或全零
BEL	振铃或警报	FOT	传输结束	RS	记录分隔符
BS	退格	ESC	转义字符	SI	移入
CAN	取消	ETB	块传输结束	SO	移出
CR	回车	ETX	文本结束	SOH	报文开始
DC1	设备控制 1	FF	换页、走纸	SP	空格
DC2	设备控制 2	FS	文件分隔符	STX	文本开始
DC3	设备控制 3	GS	分组符	SUB	替换
DC4	设备控制 4	H1	横表符	SYN	同步空载
DEL	删除	LF	进行符	US	单元分隔符
DLE	数字通信换码	NAK	否认应答	VT	竖表符
EM	媒体结束				

(经 Macmillan 允许摘自 Couch, 1990)

表 A6.2 傅里叶变换的性质

性质	数学描述
1. 线性	$ag_1(t) + bg_2(t) \Leftrightarrow aG_1(f) + bG_2(f)$ 其中 a, b 为常数
2. 时域缩放	$g(at) \Leftrightarrow \frac{1}{ a } G\left(\frac{f}{a}\right)$ 其中 a 为常数
3. 对偶性	如果 $g(t) \Leftrightarrow G(f)$, 那么 $G(t) \Leftrightarrow g(-f)$
4. 时移特性	$g(t - t_0) \Leftrightarrow G(f) \exp(-j2\pi f t_0)$
5. 频移特性	$\exp(j2\pi f_c t) g(t) \Leftrightarrow G(f - f_c)$
6. $g(t)$ 曲线下的面积	$\int_{-\infty}^{\infty} g(t) dt = G(0)$
7. $G(f)$ 曲线下的面积	$g(0) = \int_{-\infty}^{\infty} G(f) df$
8. 时域求导	$\frac{d}{dt} g(t) \Leftrightarrow j2\pi f G(f)$
9. 时域积分	$\int_{-\infty}^t g(\tau) d\tau \Leftrightarrow \frac{1}{j2\pi f} G(f) + \frac{G(0)}{2} \delta(f)$
10. 结合公式	如果 $g(t) \Leftrightarrow G(f)$, 那么 $g^*(t) \Leftrightarrow G^*(-f)$
11. 时域乘法	$g_1(t) g_2(t) \Leftrightarrow \int_{-\infty}^{\infty} G_1(\lambda) G_2(f - \lambda) d\lambda$
12. 时域卷积	$\int_{-\infty}^{\infty} g_1(\tau) g_2(t - \tau) d\tau \Leftrightarrow G_1(f) G_2(f)$

表 A6.3 傅里叶变换对

时域函数	傅里叶变换
$\text{rect}\left(\frac{t}{T}\right)$	$T \text{sinc}(fT)$
$\text{sinc}(2Wt)$	$\frac{1}{2W} \text{rect}\left(\frac{f}{2W}\right)$
$\exp(-at) u(t), \quad a > 0$	$\frac{1}{a + j2\pi f}$
$\exp(-a t), \quad a > 0$	$\frac{2a}{a^2 + (2\pi f)^2}$
$\exp(-\pi t^2)$	$\exp(-\pi f^2)$
$\begin{cases} 1 - \frac{ t }{T}, & t < T \\ 0, & t \geq T \end{cases}$	$T \text{sinc}^2(fT)$
$\delta(t)$	1
1	$\delta(f)$
$\delta(t - t_0)$	$\exp(-j2\pi f t_0)$
$\exp(j2\pi f_c t)$	$\delta(f - f_c)$
$\cos(2\pi f_c t)$	$\frac{1}{2} [\delta(f - f_c) + \delta(f + f_c)]$

(续表)

时域函数	傅里叶变换
$\sin(2\pi f_c t)$	$\frac{1}{2j} [\delta(f - f_c) - \delta(f + f_c)]$
$\text{sgn}(t)$	$\frac{1}{j\pi f}$
$\frac{1}{\pi t}$	$-j \text{sgn}(f)$
$u(t)$	$\frac{1}{2} \delta(f) + \frac{1}{j2\pi f}$
$\sum_{i=-\infty}^{\infty} \delta(t - iT_0)$	$\frac{1}{T_0} \sum_{n=-\infty}^{\infty} \delta(f - \frac{n}{T_0})$

注: $u(t)$, 单位阶跃函数 $\delta(t)$, δ 函数或单位冲激 $\text{rect}(t)$, 以原点为中心的单位幅度、单位时延的矩形函数 $\text{sgn}(t)$, 符号函数 $\text{sinc}(t)$, sinc 函数

表 A6.4 希尔伯特变换对

(在最前面两个变换对中, 假设 $m(t)$ 是在 $-W \leq f \leq W$ 内的带限信号, 其中, $W < f_c$)

时域函数	希尔伯特变换
$m(t) \cos(2\pi f_c t)$	$m(t) \sin(2\pi f_c t)$
$m(t) \sin(2\pi f_c t)$	$-m(t) \cos(2\pi f_c t)$
$\cos(2\pi f_c t)$	$\sin(2\pi f_c t)$
$\sin(2\pi f_c t)$	$-\cos(2\pi f_c t)$
$\frac{\sin t}{t}$	$\frac{1 - \cos t}{t}$
$\text{rect}(t)$	$-\frac{1}{\pi} \log \left \frac{t - \frac{1}{2}}{t + \frac{1}{2}} \right $
$\delta(t)$	$\frac{1}{\pi t}$
$\frac{1}{1+t^2}$	$\frac{t}{1+t^2}$
$\frac{1}{t}$	$-\pi \delta(t)$

注: $\delta(t)$, δ 函数 $\text{rect}(t)$, 以原点为中心的单位幅度、单位时延的矩形函数

log, 自然对数

表 A6.5 贝塞尔函数表^{*}

$n \setminus x$	0.5	1	2	3	4	6	8	10	12
0	0.9385	0.7652	0.2239	-0.2601	-0.3971	0.1506	0.1717	-0.2459	0.0477
1	0.2423	0.4401	0.5767	0.3391	-0.0660	-0.2767	0.2346	0.0435	-0.2234
2	0.0306	0.1149	0.3528	0.4861	0.3641	-0.2429	-0.1130	0.2546	-0.0849
3	0.0026	0.0196	0.1289	0.3091	0.4302	0.1148	-0.2911	0.0584	0.1951
4	0.0002	0.0025	0.0340	0.1320	0.2811	0.3576	-0.1054	-0.2196	0.1825
5	—	0.0002	0.0070	0.0430	0.1321	0.3621	0.1858	-0.2341	-0.0735
6	—	0.0012	0.0114	0.0491	0.2458	0.3376	-0.0145	-0.2437	—
7	—	0.0002	0.0025	0.0152	0.1296	0.3206	0.2167	-0.1703	—
8	—	0.0005	0.0040	0.0565	0.2235	0.3179	0.0451	—	—
9	—	0.0001	0.0009	0.0212	0.1263	0.2919	0.2304	—	—

(续表)

$n \setminus x$	0.5	1	2	3	4	6	8	10	12
10			—	0.0002	0.0070	0.0608	0.2075	0.3005	
11			—		0.0020	0.0256	0.1231	0.2704	
12					0.0005	0.0096	0.0634	0.1953	
13					0.0001	0.0033	0.0290	0.1201	
14					—	0.0010	0.0120	0.0650	

* 修正贝塞尔函数表可参阅 Watson (1966, pp.666 - 697) 和 Abramowitz 和 Stegun (1965, pp.358 - 406)。

表 A6.6 误差函数*

u	$\text{erf}(u)$	u	$\text{erf}(u)$
0.00	0.000 00	1.10	0.880 21
0.05	0.056 37	1.15	0.896 12
0.10	0.112 46	1.20	0.910 31
0.15	0.168 00	1.25	0.922 90
0.20	0.222 70	1.30	0.934 01
0.25	0.276 33	1.35	0.943 76
0.30	0.328 63	1.40	0.952 29
0.35	0.379 38	1.45	0.959 70
0.40	0.428 39	1.50	0.966 11
0.45	0.475 48	1.55	0.971 62
0.50	0.520 50	1.60	0.976 35
0.55	0.563 32	1.65	0.980 38
0.60	0.603 86	1.70	0.983 79
0.65	0.642 03	1.75	0.986 67
0.70	0.677 80	1.80	0.989 09
0.75	0.711 16	1.85	0.991 11
0.80	0.742 10	1.90	0.992 79
0.85	0.770 67	1.95	0.994 18
0.90	0.796 91	2.00	0.995 32
0.95	0.820 89	2.50	0.999 59
1.00	0.842 70	3.00	0.999 98
1.05	0.862 44	3.30	0.999 998

* 在一些参考文献中给出了误差函数扩展列表,例如 Abramowitz 和 Stegun (1965, pp.297 - 316)。

表 A6.7 部分 ITU 音频(电话线)调制解调器标准

	ITU 标准*	调制类型	比特率, bps	符号率, 波特
(a) 对称调制解调器	V.21	二进制 FSK	300	300
	V.22 bis	QPSK	1200	600
	V.26	QPSK	2400	1200
	V.27	8-PSK	4800	2400
	V.32	16-QAM	9600	2400
	V.34	1024-QAM	28 800	3429
	V.34(高速)	4 个 960-QAM 星座的嵌套星座	33 600	
(b) 非对称调制解调器	V.90: 下行	数字	56 000	
	上行	V.34(高速)	33 600	

* 后缀 Bis 代表指定标准的第二版。

表 A6.8 三角几何恒等式

$\exp(\pm j\theta) = \cos \theta \pm j \sin \theta$
$\cos \theta = \frac{1}{2} [\exp(j\theta) + \exp(-j\theta)]$
$\sin \theta = \frac{1}{2j} [\exp(j\theta) - \exp(-j\theta)]$
$\sin^2 \theta + \cos^2 \theta = 1$
$\cos^2 \theta - \sin^2 \theta = \cos(2\theta)$
$\cos^2 \theta = \frac{1}{2} [1 + \cos(2\theta)]$
$\sin^2 \theta = \frac{1}{2} [1 - \cos(2\theta)]$
$2 \sin \theta \cos \theta = \sin(2\theta)$
$\sin(\alpha \pm \beta) = \sin \alpha \cos \beta \pm \cos \alpha \sin \beta$
$\cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta$
$\tan(\alpha \pm \beta) = \frac{\tan \alpha \pm \tan \beta}{1 \mp \tan \alpha \tan \beta}$
$\sin \alpha \sin \beta = \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)]$
$\cos \alpha \cos \beta = \frac{1}{2} [\cos(\alpha - \beta) + \cos(\alpha + \beta)]$
$\sin \alpha \cos \beta = \frac{1}{2} [\sin(\alpha - \beta) + \sin(\alpha + \beta)]$

表 A6.9 级数展开

泰勒级数	$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots$
	其中 $f^{(n)}(a) = \left. \frac{d^n f(x)}{dx^n} \right _{x=a}$
麦克劳林级数	$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \cdots$
	其中 $f^{(n)}(0) = \left. \frac{d^n f(x)}{dx^n} \right _{x=0}$
二项式级数	$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \cdots, \quad nx < 1$
指数级数	$\exp x = 1 + x + \frac{1}{2!}x^2 + \cdots$
对数级数	$\log(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots$
三角级数	$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots$
	$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots$
	$\tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \cdots$
	$\sin^{-1} x = x + \frac{1}{6}x^3 + \frac{3}{40}x^5 + \cdots$
	$\tan^{-1} x = x - \frac{1}{3}x^3 + \frac{1}{5}x^5 - \cdots, \quad x < 1$
	$\operatorname{sinc} x = 1 - \frac{1}{3!}(\pi x)^2 + \frac{1}{5!}(\pi x)^4 - \cdots$

表 A6.10 积分

不定积分

$$\begin{aligned}\int x \sin(ax) dx &= \frac{1}{a^2} [\sin(ax) - ax \cos(ax)] \\ \int x \cos(ax) dx &= \frac{1}{a^2} [\cos(ax) + ax \sin(ax)] \\ \int x \exp(ax) dx &= \frac{1}{a^2} \exp(ax)(ax - 1) \\ \int x \exp(ax^2) dx &= \frac{1}{2a} \exp(ax^2) \\ \int \exp(ax) \sin(bx) dx &= \frac{1}{a^2 + b^2} \exp(ax)[a \sin(bx) - b \cos(bx)] \\ \int \exp(ax) \cos(bx) dx &= \frac{1}{a^2 + b^2} \exp(ax)[a \cos(bx) + b \sin(bx)] \\ \int \frac{dx}{a^2 + b^2 x^2} &= \frac{1}{ab} \tan^{-1}\left(\frac{bx}{a}\right) \\ \int \frac{x^2 dx}{a^2 + b^2 x^2} &= \frac{x}{b^2} - \frac{a}{b^3} \tan^{-1}\left(\frac{bx}{a}\right)\end{aligned}$$

定积分

$$\begin{aligned}\int_0^\infty \frac{x \sin(ax)}{b^2 + x^2} dx &= \frac{\pi}{2} \exp(-ab), \quad a > 0, b > 0 \\ \int_0^\infty \frac{\cos(ax)}{b^2 + x^2} dx &= \frac{\pi}{2b} \exp(-ab), \quad a > 0, b > 0 \\ \int_0^\infty \frac{\cos(ax)}{(b^2 - x^2)^2} dx &= \frac{\pi}{4b^3} [\sin(ab) - ab \cos(ab)], \quad a > 0, b > 0 \\ \int_0^\infty \sin x dx &= \int_0^\infty \sin^2 x dx = \frac{1}{2} \\ \int_0^\infty \exp(-ax^2) dx &= \frac{1}{2} \sqrt{\frac{\pi}{a}}, \quad a > 0 \\ \int_0^\infty x^2 \exp(-ax^2) dx &= \frac{1}{4a} \sqrt{\frac{\pi}{a}}, \quad a > 0\end{aligned}$$

表 A6.11 常用的常数

物理常数

波尔兹曼常数	$k = 1.38 \times 10^{-23} \text{ J/K}$
普朗克常数	$b = 6.626 \times 10^{-34} \text{ J·S}$
电子(基本)电荷	$q = 1.602 \times 10^{-19} \text{ C}$
真空中的光速	$c = 2.998 \times 10^8 \text{ m/s}$
标准(绝对)温度	$T_0 = 273 \text{ K}$
热电压	$V_T = 0.026 \text{ V(室温)}$
标准温度下的热能	$kT_0 = 3.77 \times 10^{-21} \text{ J}$

1 Hz = 1 r/s(转每秒) 1 rad = 2π rad(弧度)

1 W = 1 J/S

数学常数

自然对数的底	$e = 2.718 281 8$
以 2 为底 e 的对数	$\log_2 e = 1.442 695$
以 e 的底 2 的对数	$\log_e 2 = 0.693 147$
以 10 为底 2 的对数	$\log_{10} 2 = 0.301 03$
Pi	$\pi = 3.141 592 7$

表 A6.12 推荐的单位前缀

倍数和约数	前 缀	符 号
10^{12}	tera	T
10^9	giga	G
10^6	mega	M
10^3	kilo	K(k)
10^{-3}	milli	m
10^{-6}	micro	μ
10^{-9}	nano	n
10^{-12}	pico	p

术语表

习惯用法和符号

1. 符号 $||$ 是指复数量的绝对值或模。
2. 符号 $\arg(\cdot)$ 是指复数量的相角。
3. 符号 $\text{Re}[\cdot]$ 是指实部, 符号 $\text{Im}[\cdot]$ 是指虚部。
4. 如果不特别说明, 自然对数用 \log 表示。以 2 和 10 为底的对数分别用 \log_2 和 \log_{10} 表示。
5. 用上标 * 来表示复数的共轭, 例如 x^* 是指 x 的复共轭。
6. \rightleftharpoons 是指傅里叶变换对, 例如 $g(t) \rightleftharpoons G(f)$ 中, 小写字母代表时间的函数, 相应的大写字母代表频率的函数。
7. 符号 $F[\cdot]$ 是指傅里叶变换运算, 如 $F[g(t)] = G(f)$, 而符号 $F^{-1}[\cdot]$ 代表傅里叶反变换, 如, $F^{-1}[G(f)] = g(t)$ 。
8. \star 代表卷积, 如

$$x(t) \star b(t) = \int_{-\infty}^{+\infty} x(\tau) b(t - \tau) d\tau$$

9. 除了第 10 章中的模 2 加用普通加号来表示以外, 其他各章都用符号 \oplus 代表模 2 加。
10. 用下标 T_0 说明相关的函数 $g_{T_0}(t)$ 是一个周期为 T_0 的时间 t 的周期函数。
11. 函数上方的小帽子有两个含义:
 - (a) 代表函数的希尔伯特变换, 如函数 $\hat{g}(t)$ 是函数 $g(t)$ 的希尔伯特变换。
 - (b) 代表未知函数的估计值, 如值 $\hat{\alpha}(\mathbf{x})$ 代表基于观察矢量 \mathbf{x} 所得的未知函数 α 的估计值。
12. 函数上方的波浪线是指窄带信号的复包络, 如函数 $\tilde{g}(t)$ 是窄带信号 $g(t)$ 的复包络。在 10.8 节关于 Turbo 译码的描述中, 波浪线表示外部信息, 从而将其与对数似然比区分开。这是个例外。
13. 下标“+”表示信号的预包络, 如函数 $g_+(t)$ 是指信号 $g(t)$ 的预包络。于是有 $g_+(t) = g(t) + j\hat{g}(t)$, 其中 $\hat{g}(t)$ 是 $g(t)$ 的希尔伯特变换。用下标“-”表示 $g_-(t) = g(t) - j\hat{g}(t) = g_-'(t)$ 。
14. 下标 I 和 Q 分别表示载波为 $\cos(2\pi f_c t)$ 的窄带信号、窄带随机过程或窄带滤波器的冲激响应的同相分量和正交分量。
15. 对于低通信息信号, 最高频率分量或信息带宽由 W 表示。低通信号的频谱占用了 $-W \leq f \leq W$ 的频谱间隔, 而在其他地方为零。载波为 f_c 的带通信号频谱占用的是 $f_c - W \leq f \leq f_c + W$ 和 $-f_c - W \leq f \leq -f_c + W$ 的频谱间隔, 所以信号的带宽用 $2W$ 表示。带通信号的复包络(低通)频谱占用的是 $-W \leq f \leq W$ 的频谱间隔。

带通滤波器的带宽用 B 表示。常用的滤波器带宽定义是该滤波器的幅度响应比零频率点低 3 dB 时对应的频率点。对于中心频率为 f_c 的带通滤波器来说，带宽用 $2B$ 来表示，其中心在 f_c 。该带通滤波器的复低通等效带宽为 B 。

传输调制信号的通信信道的传输带宽用 B_1 表示。

16. 随机变量或随机矢量用大写字母表示(如 X 或 \mathbf{X})，而它们的样本值用小写字母表示(如 x 或 \mathbf{x})。
17. 表达式中的竖线表示“在…条件下”，如 $f_x(x|H_0)$ 是假设 H_0 成立的条件下，随机变量 X 的概率密度函数。
18. 符号 $E[\cdot]$ 是指括号中随机变量的平均值。 E 可看成是一个运算符。
19. 符号 $\text{var}[\cdot]$ 是指括号中随机变量的方差。
20. 符号 $\text{cov}[\cdot]$ 是指括号中两个随机变量的协方差。
21. 平均符号差错概率用 P_e 表示。
在二进制信号的情况下， p_{10} 代表发送符号 0 时的条件差错概率， p_{01} 代表发送符号 1 时的条件差错概率。符号 0 和 1 的先验概率分别用 p_0 和 p_1 来表示。
22. 符号 $\langle \cdot \rangle$ 表示括号内样本函数的时间平均值。
23. 粗体字代表矢量或矩阵。方阵 \mathbf{R} 的逆矩阵用 \mathbf{R}^{-1} 来表示。矢量 \mathbf{w} 的转置用 \mathbf{w}^T 来表示。复数矢量 \mathbf{x} 的埃尔米特转置用 \mathbf{x}^H 表示。埃尔米特转置包括转置和复共轭。
24. 矢量 \mathbf{x} 的长度用 $\|\mathbf{x}\|$ 表示。两个矢量 \mathbf{x}_i 和 \mathbf{x}_j 间的欧氏距离用 $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$ 表示。
25. 两个实数矢量 \mathbf{x} 和 \mathbf{y} 的内积用 $\mathbf{x}^T \mathbf{y}$ 来表示，外积用 $\mathbf{x} \mathbf{y}^T$ 来表示。如果矢量 \mathbf{x} 和 \mathbf{y} 是复数，则它们的内积是 $\mathbf{x}^H \mathbf{y}$ ，外积是 $\mathbf{x} \mathbf{y}^H$ 。
26. $M \times 1$ 的矢量 α 和 β 的矢量积是一个 $M \times 1$ 的矢量，其定义为

$$\alpha \cdot \beta = \begin{bmatrix} \alpha_1 \beta_1 \\ \alpha_2 \beta_2 \\ \vdots \\ \alpha_M \beta_M \end{bmatrix}$$

其中， α_k 和 β_k 分别是 α 和 β 的第 k 个元素。 $\alpha \cdot \beta$ 的矢量积 L_1 的模为

$$\|\alpha \cdot \beta\|_1 = \sum_{m=1}^M |\alpha_m \beta_m|$$

函数

1. 矩形函数：

$$\text{rect}(t) = \begin{cases} 1, & -\frac{1}{2} < t < \frac{1}{2} \\ 0 & |t| > \frac{1}{2} \end{cases}$$
2. 单位阶跃函数：

$$u(t) = \begin{cases} 1, & t > 0 \\ 0, & t < 0 \end{cases}$$
3. 符号函数：

$$\text{sgn}(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1, & t < 0 \end{cases}$$

4. (狄拉克) δ 函数: $\delta(t) = 0, t \neq 0$
 $\int_{-\infty}^{\infty} \delta(t) dt = 1$
 或等效为: $\int_{-\infty}^{\infty} g(t) \delta(t - t_0) dt = g(t_0)$
5. Sinc 函数: $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$
6. 正弦积分: $\text{Si}(u) = \int_0^u \frac{\sin x}{x} dx$
7. 误差函数: $\text{erf}(u) = \frac{2}{\sqrt{\pi}} \int_0^u \exp(-z^2) dz$
 互补误差函数: $\text{erfc}(u) = 1 - \text{erf}(u).$
8. 贝努利系数: $\binom{n}{k} = \frac{n!}{(n-k)! k!}$
9. n 阶第一类贝塞尔函数: $J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(jx \sin \theta - jn\theta) d\theta$
10. 零阶第一类修正贝塞尔函数: $I_0(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(x \cos \theta) d\theta$
11. 汇合型超几何函数: ${}_1F_1(a; b; x) = 1 + \frac{a}{b} \frac{x}{1!} + \frac{a(a+1)}{b(b+1)} \frac{x^2}{2!} + \dots$

缩写

A:	安(培)
AC:	交流
ADC:	模-数转换
ADM:	自适应增量调制
ADPCM:	自适应差分脉冲编码调制
ADSL:	非对称数字用户线
AM:	幅度调制
ANSI:	美国国家标准协会
APB:	后向估计自适应预测
APF:	前向估计自适应预测
AQB:	后向估计自适应量化
AQF:	前向估计自适应量化
ARQ:	自动请求重发
ASCII:	美国国家信息交换标准码
ASK:	幅移键控
ATM:	异步传输模式
AWGN:	加性高斯白噪声
bps:	比特/秒
BER:	误比特率

BISDN:	宽带 ISDN
BPF:	带通滤波器
BSC:	二进制对称信道
CAP:	无载波幅度/相位调制
CCITT:	国际电话与电报咨询委员会(现在改名为 ITU)
CDM:	码分复用
CDMA:	码分多址
CELP:	码激励线性预测(模型)
CO:	中心局
codec:	编/译码
CPFSK:	连续相位频移键控
CRC:	循环冗余校验
CW:	连续波
DAC:	数-模转换
dB:	分贝
dBW:	对应于 1 W 的 dB 值
dBmW:	对应于 1 mW 的 dB 值
DC:	直流
DEM:	解调器
DES:	数据加密标准
DFT:	离散傅里叶变换
DM:	增量调制
DMT:	离散多音
DPCM:	差分脉冲编码调制
DPSK:	差分相移键控
DSB -SC:	抑制载波的双边带
DS/BPSK:	直接序列/二进制相移键控
DSL:	数字用户线
exp:	指数
FDM:	频分复用
FDMA:	频分多址
FEXT:	远端串话
FFT:	快速傅里叶变换
FH:	跳频
FH/MFSK:	跳频/ M 进制频移键控
FMFB:	反馈频率调制器
FSK	频移键控
GMSK:	高斯滤波的最小频移键控
GSM:	全球移动通信系统

HDTV:	高清晰度电视
Hz:	赫(兹)
IDFT:	离散傅里叶反变换
IF:	中频
I/O:	输入/输出
IP:	因特网协议
IS-95:	暂行标准 95
ISDN:	综合业务数字网
ISI:	码间干扰
ISO:	国际标准化组织
ITU:	国际电信联盟
JPEG:	联合图像专家组
LAN:	局域网
LDM:	线性增量调制
LMS:	最小均方
log:	对数
\log_2 :	以 2 为底的对数
\log_{10} :	以 10 为底的对数
LPC:	线性预测编码(模型)
LPF:	低通滤波器
MAP:	最大后验概率
ML:	最大似然
mmse:	最小均方误差
modem:	调制解调器
MPEG:	活动图像专家组
ms:	毫秒
μs :	微秒
MSK:	最小频移键控
NCO:	数控振荡器
NEXT:	近端串话
nm:	纳秒
NRZ:	不归零
NTSC:	国家电视系统委员会
OC:	光载波
OFDM:	正交频分复用
OOK:	开 - 关键控
OSI:	开放系统互联
PAM:	脉冲幅度调制
PCM:	脉冲编码调制

PDM:	脉冲持续时间调制
PG:	处理增益
PLL:	锁相环
PN:	伪噪声(伪随机)
POTS:	传统电话服务
PPM:	脉冲位置调制
PSK:	相移键控
PSTN:	公众电话交换网
PWM:	脉冲宽度调制
QAM:	正交幅度调制
QoS:	服务质量
QPSK:	正交相移键控
RF:	射频
rms:	均方根
RS:	里德-索罗蒙(码)
RS-232:	推荐标准 232(端口)
RSA:	RSA(算法)
RSC:	回归系统卷积(码)
s:	秒
SDH:	同步数字系列
SDMA:	空分多址
SDR:	信号失真比
SNR:	信噪比
SONET:	同步光网络
STFF:	短时傅里叶变换
STM:	同步转移模式
TC:	时间压缩
TCM:	网格编码调制
TDM:	时分复用
TDMA:	时分多址
TV:	电视
UHF:	超高频
V:	伏(特)
VCO:	压控振荡器
VHF:	甚高频
VLSI:	超大规模集成电路
W:	瓦(特)
WDM:	波分复用

参 考 文 献

- M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (New York: Dover Publications, 1965).
- N. Abramson, *Information Theory and Coding* (New York: McGraw-Hill, 1963).
- J. Adamék, *Foundations of Coding* (New York: Wiley, 1991).
- Y. Akaiwa, *Introduction to Digital Mobile Communication* (New York: Wiley, 1997).
- J.B. Anderson, T. Aulin, and C.E. Sundberg, *Digital Phase Modulation* (New York: Plenum Publishers, 1986).
- J.B. Anderson and S. Mohan, *Source and Channel Coding: An Algorithmic Approach* (Boston, Mass.: Kluwer Academic, 1991).
- J.B. Anderson, *Digital Transmission Engineering* (Piscataway, N.J.: IEEE Press, 1999).
- R.B. Ash, *Information Theory* (New York: Wiley, 1965).
- Bell Laboratories Technical Staff, *A History of Engineering Science in the Bell System: The Early Years (1875–1925)*, (Books on Demand, Ann Arbor, Michigan: 1975).
- J.C. Bellamy, *Digital Telephony*, Second Edition (New York: Wiley, 1991).
- S. Benedetto and E. Biglieri, *Principles of Digital Transmission with Wireless Applications* (New York: Kluwer Academic/Plenum Publishers, 1999).
- S. Benedetto, E. Biglieri, and V. Castellani, *Digital Transmission Theory* (Englewood Cliffs, N.J.: Prentice-Hall, 1987).
- W.R. Bennett, *Introduction to Signal Transmission* (New York: McGraw-Hill, 1970).
- K.B. Benson and J.C. Whitaker, *Television Engineering Handbook*, rev. ed. (New York: McGraw-Hill, 1992).
- T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression* (Englewood Cliffs, N.J.: Prentice-Hall, 1971).
- E.R. Berlekamp, *Algebraic Coding Theory* (New York: McGraw-Hill, 1968).
- E.R. Berlekamp (editor), *Key Papers in the Development of Coding Theory* (Piscataway, N.J.: IEEE Press, 1974).
- V.K. Bhargava, D. Haccoun, R. Matyas, and P. Nuspl, *Digital Communications by Satellite: Modulation, Multiple Access, and Coding* (New York: Wiley, 1981).
- E. Biglieri, D. Divsalar, P.J. McLane, and M.K. Simon, *Introduction to Trellis-Coded Modulation with Applications* (New York: Macmillan, 1991).
- J.A.C. Bingham, *The Theory and Practice of Modem Design* (New York: Wiley, 1988).
- R.B. Blackman and J.W. Tukey, *The Measurement of Power Spectra, from the Point of View of Communication Engineering* (New York: Dover, 1958).
- H.S. Black, *Modulation Theory* (Princeton, N.J.: Van Nostrand, 1953).
- R.E. Blahut, *Principles and Practice of Information Theory* (Reading, Mass.: Addison-Wesley, 1987).
- R.E. Blahut, *Digital Transmission of Information* (Reading, Mass.: Addison-Wesley, 1990).
- G.E.P. Box and G.M. Jenkins, *Time Series Analysis: Forecasting and Control* (San Francisco: Holden Day, 1976).
- R.N. Bracewell, *The Fourier Transform and Its Applications*, 2nd ed., rev. (New York: McGraw-Hill, 1986).
- L. Brillouin, *Science and Information Theory*, 2nd ed. (New York: Academic Press, 1962).
- K.W. Cattermole, *Principles of Pulse-code Modulation* (New York: American Elsevier, 1969).
- W.Y. Chen, *DSL: Simulation Techniques and Standards Development for Digital Subscriber Line Systems* (Indianapolis, Ind.: Macmillan Technical Publishing, 1998).
- J.M. Cioffi, *Digital Data Transmission*, EE379C Course Textbook, Stanford University, 1998.

- G.C. Clark, Jr., and J.B. Cain, *Error-correction Coding for Digital Communications* (New York: Plenum Publishers, 1981).
- L.W. Couch, *Digital and Analog Communication Systems*, 5th ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1997).
- T.M. Cover and J.A. Thomas, *Elements of Information Theory* (New York: Wiley, 1991).
- H. Cramér and M.R. Leadbetter, *Stationary and Related Stochastic Processes: Sample Function Properties and Their Applications* (New York: Wiley, 1967).
- W.B. Davenport, Jr., and W.I. Root, *An Introduction to the Theory of Random Signals and Noise* (New York: McGraw-Hill, 1958).
- R.C. Dixon, *Spread Spectrum Systems*, 2nd ed. (New York: Wiley, 1984).
- R.C. Dixon (editor), *Spread Spectrum Techniques* (New York: IEEE Press, 1976).
- L.J. Doob, *Stochastic Processes* (New York: Wiley, 1953).
- J.J. Downing, *Modulation Systems and Noise* (Englewood Cliffs, N.J.: Prentice-Hall, 1964).
- W. Feller, *An Introduction to Probability Theory and its Application*, vol. 1, 3rd ed. (New York: Wiley, 1968).
- T.L. Fine, *Theories of Probability: An Examination of Foundations* (New York: Academic Press, 1973).
- L.E. Franks (editor), *Data Communication: Fundamentals of Baseband Transmission* (Dowden, Hutchison, and Ross, 1974).
- L.E. Franks, *Signal Theory* (Englewood Cliffs, N.J.: Prentice-Hall, 1969).
- R.L. Freeman, *Telecommunications Transmission Handbook*, 4th ed. (New York: Wiley, 1998).
- R.M. Gagliardi, *Introduction to Communications Engineering*, 2nd ed. (New York: Wiley, 1988).
- R.G. Gallager, *Information Theory and Reliable Communication* (New York: Wiley, 1968).
- R.G. Gallager, *Low-Density Parity-Check Codes* (Cambridge, Mass.: MIT Press, 1963).
- F.M. Gardner, *Phaselock Techniques*, 2nd ed. (New York: Wiley, 1979).
- V.K. Gary and J.E. Wilkes, *Principles & Applications of GSM* (Englewood Cliffs, N.J.: Prentice-Hall, 1999).
- A. Gersho and R.M. Gray, *Vector Quantization and Signal Compression* (Boston, Mass.: Kluwer Academic, 1992).
- J.D. Gibson (editor), *The Mobile Communications Handbook* (Piscataway, N.J.: IEEE Press, 1996).
- R.D. Gitlin, J.F. Hayes, and S.B. Weinstein, *Data Communications Principles* (New York: Plenum, 1992).
- B. Goldberg and H.S. Bennett (editors), *Communication Channels: Characterization and Behavior* (New York: IEEE Press, 1976).
- S.W. Golomb (editor), *Digital Communications with Space Applications* (Englewood Cliffs, N.J.: Prentice-Hall, 1964).
- S.W. Golomb, *Shift Register Sequences* (San Francisco: Holden-Day, 1967).
- R.M. Gray and L.D. Davisson, *Random Processes: A Mathematical Approach for Engineers* (Englewood Cliffs, N.J.: Prentice-Hall, 1986).
- P.E. Green, Jr., *Computer Network Architectures and Protocols* (New York: Plenum, 1982).
- P.E. Green, Jr., *Fiber Optic Networks* (Englewood Cliffs, N.J.: Prentice-Hall, 1993).
- M.S. Gupta (editor), *Electrical Noise: Fundamentals and Sources* (New York: IEEE Press, 1977).
- R.W. Hamming, *Coding and Information Theory* (Englewood Cliffs, N.J.: Prentice-Hall, 1980).
- S. Haykin, *Communication Systems*, 3rd ed. (New York: Wiley, 1994).
- S. Haykin, *Adaptive Filter Theory*, 3rd ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1996).
- S. Haykin and B. Van Veen, *Signals and Systems* (New York: Wiley, 1999).
- C. Heegard and S.B. Wicker, *Turbo Coding* (Boston, Mass.: Kluwer Academic Publishers, 1999).
- G. Held, *The Complete Modem Reference*, 3rd ed. (New York: Wiley, 1997).
- C.W. Helstrom, *Statistical Theory of Signal Detection* (Elmsford, N.Y.: Pergamon Press, 1968).
- C.W. Helstrom, *Probability and Stochastic Processes for Engineers*, 2nd ed. (New York: Macmillan, 1990).
- K. Henney (editor), *Radio Engineering Handbook* (New York: McGraw-Hill, 1959).
- J.K. Holmes, *Coherent Spread Spectrum Systems* (New York: Wiley, 1982).
- W.C. Jakes, Jr. (editor), *Microwave Mobile Communications* (New York: Wiley, 1974).

- N.S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video* (Englewood Cliffs, N.J.: Prentice-Hall, 1984).
- N.S. Jayant (editor), *Waveform Quantization and Coding* (New York: IEEE Press, 1976).
- H. Jeffreys, Sir, *Theory of Probability*, 3rd ed. (Oxford: Clarendon Press, 1967).
- H. Jeffreys, Sir, and B.S. Jeffreys, *Methods of Mathematical Physics*, 3rd ed. (Cambridge University Press, 1956).
- M.C. Jeruchim, B. Balaban, and J.S. Shanmugan, *Simulation of Communication Systems* (New York: Plenum, 1992).
- E.C. Jordan and K.G. Balmain, *Electromagnetic Waves and Radiating Systems*, 2nd ed. (Englewood Cliffs, N.J.: Prentice Hall, 1968).
- A. Khintchin, *Mathematical Foundations of Information Theory* (New York: Dover, 1957).
- A.N. Kolmogorov, *Foundations of the Theory of Probability* (New York: Chelsea Publishing, 1956).
- V.A. Kotelnikov, *The Theory of Optimum Noise Immunity* (New York: McGraw-Hill, 1960).
- J.D. Kraus, *Antennas* (New York: McGraw-Hill, 1950).
- S. Kullback, *Information Theory and Statistics* (New York: Dover, 1968).
- P. Lafrance, *Fundamental Concepts in Communication* (Englewood Cliffs, N.J.: Prentice-Hall, 1990).
- B.P. Lathi, *Modern Digital and Analog Communication Systems*, 2nd ed. (Oxford University Press, 1995).
- I. Lebow, *Information Highways and Byways* (Piscataway, N.J.: IEEE Press, 1995).
- E.A. Lee and D.G. Messerschmitt, *Digital Communication*, 2nd ed. (Boston, Mass.: Kluwer Academic, 1994).
- J.S. Lee and L.E. Miller, *CDMA Systems Engineering Handbook* (Boston, Mass.: Artech House Publishers, 1998).
- Y.W. Lee, *Statistical Theory of Communication* (New York: Wiley, 1960).
- W.C.(Y.) Lee, *Mobile Communications Engineering* (New York: McGraw-Hill, 1982).
- A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, 2nd ed. (Reading, Mass.: Addison-Wesley, 1994).
- C.R. Lewart, *The Ultimate Modem Handbook* (Englewood Cliffs, N.J.: Prentice-Hall, 1998).
- S. Lin and D.J. Costello, Jr., *Error Control Coding: Fundamentals and Applications* (Englewood Cliffs, N.J.: Prentice-Hall, 1983).
- W.C. Lindsey, *Synchronization Systems in Communication and Control* (Englewood Cliffs, N.J.: Prentice-Hall, 1972).
- W.C. Lindsey and M.K. Simon (editors), *Phase-locked Loops and Their Applications* (New York: IEEE Press, 1978).
- W.C. Lindsey and M.K. Simon, *Telecommunication Systems Engineering* (Englewood Cliffs, N.J.: Prentice-Hall, 1973).
- M. Loève, *Probability Theory* (Princeton, N.J.: Van Nostrand, 1963).
- R.W. Lucky, *Silicon Dreams: Information, Man, and Machine* (New York: St. Martin's Press, 1989).
- R.W. Lucky, J. Salz, and E.J. Weldon, Jr., *Principles of Data Communication* (New York: McGraw-Hill, 1968).
- F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-correcting Codes* (Amsterdam: North-Holland, 1977).
- V.K. Madisetti and D.B. Williams (editors), *The Digital Signal Processing Handbook* (Piscataway, N.J.: IEEE Press, 1998).
- R.J. Marks II, *Introduction to Shannon Sampling and Interpolation Theory* (New York/Berlin: Springer-Verlag, 1991).
- J.C. McDonald (editor), *Fundamentals of Digital Switching*, 2nd ed. (New York: Plenum, 1990).
- R. McDonough and A.D. Whalen, *Detection of Signals in Noise*, 2nd ed. (New York: Academic Press, 1995).
- R.J. McEliece, *The Theory of Information and Coding: A Mathematical Framework for Communication* (Reading, Mass.: Addison-Wesley, 1977).
- A. Mengali and N. D'Andrea, *Synchronization Techniques for Digital Receivers* (New York: Plenum, 1997).

- D.J.G. Mestdagh, *Fundamentals of Multiaccess Optical Fiber Networks* (Boston, Mass.: Artech House Publishers, 1995).
- C.H. Meyer and S.M. Matyas, *Cryptography: A New Dimension in Computer Data Security* (New York: Wiley, 1982).
- H. Meyr and G. Ascheid, *Synchronization in Digital Communications*, vol. 1 (New York: Wiley, 1990).
- H. Meyr, M. Moeneclaey, and S.A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing* (New York: Wiley, 1998).
- A.M. Michelson and A.H. Levesque, *Error-control Techniques for Digital Communication* (New York: Wiley, 1985).
- D. Middleton, *An Introduction to Statistical Communication Theory* (New York: McGraw-Hill, 1960).
- J.G. Nellist, *Understanding Telecommunications and Lightwave Systems: An Entry Level Guide* (Piscataway, N.J.: IEEE Press, 1992).
- C.F.J. Overhage (editor), *The Age of Electronics* (New York: McGraw-Hill, 1962).
- P.F. Panter, *Modulation, Noise and Spectral Analysis, Applied to Information Transmission* (New York: McGraw-Hill, 1965).
- A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. (New York: McGraw-Hill, 1984).
- J.D. Parsons, *The Mobile Radio Propagation Channel* (New York: Wiley, 1992).
- K. Pahlavan and A.H. Levesque, *Wireless Information Networks* (New York: Wiley, 1996).
- J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (San Mateo, Calif: Morgan Kaufman Publishers, 1988).
- W.W. Peterson and E.J. Weldon, Jr., *Error Correcting Codes*, 2nd ed. (Cambridge, Mass.: MIT Press, 1972).
- J.R. Pierce and A.M. Noll, *Signals: The Science of Telecommunications* (New York: Scientific American Library, 1990).
- J.R. Pierce, *Symbols, Signals and Noise: The Nature and Process of Communication* (New York: Harper, 1961).
- H.V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. (New York/Berlin: Springer-Verlag, 1994).
- T. Pratt and C.W. Bostian, *Satellite Communications* (New York: Wiley, 1986).
- J.G. Proakis, *Digital Communications*, 3rd ed. (New York: McGraw-Hill, 1995).
- L.R. Rabiner and R.W. Schafer, *Digital Processing of Speech Signals* (Englewood Cliffs, N.J.: Prentice-Hall, 1978).
- K.R. Rao and P. Yip, *Discrete Cosine Transform: Algorithms, Advantages, Applications* (New York: Academic Press, 1990).
- T.S. Rappaport, *Smart Antennas* (Piscataway, N.J.: IEEE Press, 1998).
- T.S. Rappaport, *Wireless Communications: Principles and Practice* (Piscataway, N.J.: IEEE Press, 1996).
- S.O. Rice, "Noise in FM receivers," in M. Rosenblatt (editor), *Proceedings of the Symposium on Time Series Analysis* (New York: Wiley, 1963), pp. 395-411.
- J.H. Roberts, *Angle Modulation: The Theory of Systems Assessment*, IEE Communication Series 5 (London: Institution of Electrical Engineers, 1977).
- H.E. Rowe, *Signals and Noise in Communication Systems* (Princeton, N.J.: Van Nostrand, 1965).
- D.J. Sakrison, *Communication Theory: Transmission of Waveforms and Digital Information* (New York: Wiley, 1968).
- C. Schlegel, *Trellis Coding* (Piscataway, N.J.: IEEE Press, 1997).
- M. Schwartz, W.R. Bennett, and S. Stein, *Communication Systems and Techniques* (New York: McGraw-Hill, 1966).
- M. Schwartz, *Information Transmission, Modulation and Noise: A Unified Approach*, 3rd ed. (New York: McGraw-Hill, 1980).
- M. Schwartz, *Telecommunication Networks: Protocols, Modeling, and Analysis* (Reading, Mass.: Addison Wesley, 1987).

- K.S. Shanmugan, *Digital and Analog Communication Systems* (New York: Wiley, 1979).
- C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (Urbana: University of Illinois Press, 1949).
- G.J. Simmons (editor), *Contemporary Cryptology: The Science of Information Integrity* (Piscataway, N.J.: IEEE Press, 1992).
- M.K. Simon, J.K. Omura, R.A. Scholtz, and B.K. Levitt, *Spread Spectrum Communications*, vols. I, II, and III (New York: Computer Science Press, 1985).
- B. Sklar, *Digital Communications: Fundamentals and Applications* (Englewood Cliffs, N.J.: Prentice-Hall, 1988).
- D. Slepian (editor), *Key Papers in the Development of Information Theory* (New York: IEEE Press, 1974).
- N.J.A. Sloane and A.D. Wyner, *Claude Shannon: Collected Papers* (Piscataway, N.J.: IEEE Press, 1993).
- D.R. Smith, *Digital Transmission Systems* (Princeton, N.J.: Van Nostrand Reinhold, 1985).
- I.S. Sokolnikoff and R.M. Redheffer, *Mathematics of Physics and Modern Engineering* (New York: McGraw-Hill, 1966).
- J.J. Spilker, Jr., *Digital Communications by Satellite* (Englewood Cliffs, N.J.: Prentice-Hall, 1977).
- W. Stallings, *ISDN and Broadband ISDN*, 2nd ed. (New York: Macmillan, 1992).
- T. Starr, J.M. Cioffi, and P.J. Silverman, *Understanding Digital Subscriber Line Technology* (Englewood Cliffs, N.J.: Prentice-Hall, 1999).
- R. Steele, *Delta Modulation Systems* (New York: Wiley, 1975).
- R. Steele and L. Hanzo (editors), *Mobile Radio Communications*, 2nd ed. (New York: Wiley, 1999).
- J.J. Stiffler, *Theory of Synchronous Communications* (Englewood Cliffs, N.J.: Prentice-Hall, 1971).
- E.D. Sunde, *Communications Systems Engineering Theory* (New York: Wiley, 1969).
- A.S. Tanenbaum, *Computer Networks*, 2nd ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1995).
- S. Tantaratana and K.M. Ahmed, *Wireless Applications of Spread Spectrum Systems: Selected Readings* (Piscataway, N.J.: IEEE Press, 1998).
- T.M. Thompson, *From Error-correcting Codes Through Sphere Packings to Simple Groups* (The Mathematical Association of America, Washington D.C.: 1983).
- D.J. Torrieri, *Principles of Military Communication Systems*, 2nd ed. (Boston, Mass.: Artech House Publishers, 1992).
- G.L. Turin, *Notes on Digital Communications* (Princeton, N.J.: Van Nostrand-Reinhold, 1969).
- A. Van der Ziel, *Noise: Source, Characterization, Measurement* (Englewood Cliffs, N.J.: Prentice-Hall, 1970).
- H.F. Vanlandingham, *Introduction to Digital Control Systems* (New York: Macmillan, 1985).
- H.C.A. van Tilborg, *An Introduction to Cryptology* (Boston, Mass.: Kluwer, 1988).
- H.L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part I (New York: Wiley, 1968).
- A.J. Viterbi, *Principles of Coherent Communication* (New York: McGraw-Hill, 1966).
- A.J. Viterbi and J.K. Omura, *Principles of Digital Communication and Coding* (New York: McGraw-Hill, 1979).
- G.N. Watson, *A Treatise in the Theory of Bessel Functions*, 2nd ed. (New York: Cambridge University Press, 1966).
- N. Wax (editor), *Selected Papers on Noise and Stochastic Processes* (New York: Dover Publications, 1954).
- E.T. Whittaker and G.N. Watson, *A Course in Modern Analysis*, 4th ed. (New York: Cambridge University Press, 1952).
- S.B. Wicker and V.K. Bhargava (editors), *Reed-Solomon Codes* (Piscataway, N.J.: IEEE Press, 1994).
- B. Widrow and S.D. Stearns, *Adaptive Signal Processing* (Englewood Cliffs, N.J.: Prentice-Hall, 1985).
- N. Wiener, *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series, with Engineering Applications* (New York: Wiley, 1949).
- S.G. Wilson, *Digital Modulation and Coding* (Englewood Cliffs, N.J.: Prentice-Hall, 1996).
- E. Wong, *Stochastic Processes in Information and Dynamical Systems* (New York: McGraw-Hill, 1971).

- P.M. Woodward, *Probability and Information Theory, with Applications to Radar*, 2nd ed. (Elmsford, N.Y.: Pergamon Press, 1964).
- J.M. Wozencraft and I.M. Jacobs, *Principles of Communication Engineering* (New York: Wiley, 1965).
- W.W. Wu, *Elements of Digital Satellite Communication*, vol. I (New York: Computer Science Press, 1984).
- C.R. Wylie and L.C. Barrett, *Advanced Engineering Mathematics*, 5th ed. (New York: McGraw-Hill, 1982).
- R.D. Yates and D.J. Goodman, *Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers* (New York: Wiley, 1999).
- J.H. Yuen (editor), *Deep Space Telecommunications Systems Engineering* (New York: Plenum, 1983).
- R.E. Ziemer and R.L. Peterson, *Digital Communications and Spread Spectrum Systems* (New York: Macmillan, 1985).
- R.E. Ziemer and W.H. Tranter, *Principles of Communications*, 3rd ed. (Boston, Mass.: Houghton Miflin, 1990).

■ PAPERS, REPORTS, PATENTS¹

- M.R. Aaron and D.W. Tufts, "Intersymbol interference and error probability," *IEEE Trans. on Information Theory*, vol. IT-12, pp. 26–34, 1966.
- J.E. Abate, "Linear and adaptive delta modulation," *Proceedings of the IEEE*, vol. 55, pp. 298–308, 1967.
- A.N. Akansu, P. Duhamel, X. Lin, and M. de Courville, "Orthogonal transmultiplexers in communications: A review," *IEEE Transactions on Signal Processing*, vol. 46, pp. 979–995, 1998.
- Y. Akaiwa and Y. Nagata, "Highly efficient digital mobile communications with a linear modulation method," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, pp. 890–895, 1987.
- O. Al-Shaykh, R. Neff, D. Taubman, and A. Zakhour, "Video sequence compression." In V.K. Madisetti and D.B. Williams (editors), *The Digital Signal Processing Handbook*, CRC Press, pp. 55-1–55-19, 1998.
- F. Amoroso, "The bandwidth of digital data signals," *IEEE Communications Magazine*, vol. 18, no. 6, pp. 13–24, 1980.
- J.B. Anderson and D.P. Taylor, "A bandwidth-efficient class of signal space codes," *IEEE Transactions on Information Theory*, vol. IT-24, pp. 703–712, 1978.
- R.R. Anderson and J. Salz, "Spectra of digital FM," *Bell System Tech. J.*, vol. 44, pp. 1165–1189, 1965.
- R. Arens, "Complex processes for envelopes of normal noise," *IRE Trans. on Information Theory*, vol. IT-3, pp. 204–207, 1957.
- E.H. Armstrong, "A method of reducing disturbances in radio signaling by a system of frequency modulation," *Proceedings of the IRE*, vol. 24, pp. 689–740, 1936.
- E. Arthurs and H. Dym, "On the optimum detection of digital signals in the presence of white Gaussian noise—A geometric interpretation and a study of three basic data transmission systems," *IRE Trans. on Communication Systems*, vol. CS-10, pp. 336–372, 1962.
- B.S. Atal and J.R. Remde, "A new model of LPC excitation for producing natural-sounding speech at low bit rates," *Proc. ICASSP '82*, pp. 614–17, 1982.
- B.S. Atal and M.R. Schroeder, "Stochastic coding of speech signals at very low bit rates," *IEEE International Conference on Communications*, May 1984.
- M. Austin, "Decision-feedback equalization for digital communication over dispersive channels," *MIT Research Laboratory of Electronics Technical Report 461*, 1967.
- E. Ayanoglu, N.R. Dagdeviren, J.E. Mazo, and R. Saltzberg, "High-speed modem synchronized to a remote codec," United States Patent 5,394,437, February 28, 1995.
- E. Ayanoglu, N.R. Dagdeviren, G.D. Golden, and J.E. Mazo, "An equalizer design technique for the PCM modem: a new modem for the digital public switched network," *IEEE Transactions on Communications*, vol. 46, pp. 763–774, 1998.

- L.R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. IT-20, pp. 284-287, 1974.
- G. Battail, "Coding for the Gaussian channel: the promise of weighted output decoding," *International J. Satellite Communications*, vol. 7, pp. 183-192, 1989.
- G. Battail, "Pondération des symbols décodés par l'algorithme de Viterbi," *Ann. Télécommunication*, vol. 42, pp. 31-38, 1987.
- E. Bedrosian, "The analytic signal representation of modulated waveforms," *Proceedings of the IRE*, vol. 50, pp. 2071-2076, 1962.
- P.A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Transactions on Communication Systems*, vol. CS-11, pp. 360-393, 1963.
- S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Transactions on Information Theory*, vol. 42, pp. 409-428, 1996.
- W.R. Bennett, "Spectra of quantized signals," *Bell System Tech. J.*, vol. 27, pp. 446-472, 1948.
- N. Benvenuto, et al., "The 32 kb/s ADPCM coding standard," *AT&T Technical Journal*, vol. 65, pp. 12-22, Sept./Oct. 1986.
- C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo codes," *IEEE Transactions on Communications*, vol. 44, pp. 1261-1271, 1996.
- C. Berrou and A. Glavieux, "Reflections on the Prize Paper: Near optimum error-correcting coding and decoding turbo codes," *IEEE Information Theory Society Newsletter*, vol. 48, no. 2, p. 1 and pp. 24-31, June 1998.
- C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correction coding and decoding: turbo codes," *International Conference on Communications*, pp. 1064-1090, Geneva, Switzerland, May 1993.
- V.K. Bhargava, "Forward error correction schemes for digital communications," *IEEE Communications Magazine*, vol. 21, no. 1, pp. 11-19, 1983.
- R.C. Bose and D.K. Ray-Chaudhuri, "On a class of error correcting binary group codes," *Information and Control*, vol. 3, pp. 68-79, 1960.
- K. Brandenburg and G. Stoll, "ISO-MPEG-1 Audio: A generic standard for coding of high-quality digital audio," *Journal of the Audio Engineering Society*, vol. 42, pp. 780-792, 1994.
- D.G. Brennan, "Linear diversity combining techniques," *Proceedings of the IRE*, vol. 47, pp. 1075-1102, 1959.
- A. Buzo, A.H. Gray, Jr., R.M. Gray, and J.D. Markel, "Speech coding based upon vector quantization," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-28, pp. 562-574, 1980.
- C.R. Cahn, "Combined digital phase and amplitude modulation communication systems," *IRE Transactions on Communication Systems*, vol. CS-8, pp. 150-155, 1960.
- J.R. Carson, "Notes on the theory of modulation," *Proceedings of the IRE*, vol. 10, pp. 57-64, 1922.
- J.R. Carson and T.C. Fry, "Variable frequency electric circuit theory with application to the theory of frequency modulation," *Bell System Tech. J.*, vol. 16, pp. 513-540, 1937.
- E.F. Casas and C. Leung, "OFDM for data communication over mobile radio FM channels," *IEEE Transactions on Communications*, vol. 39, pp. 783-793, 1991.
- J.G. Chaffee, "The application of negative feedback to frequency-modulation systems," *Bell System Tech. J.*, vol. 18, pp. 404-437, 1939.
- R.W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell System Tech. J.*, vol. 45, pp. 1775-1796, 1996.
- W.Y. Chen, G.H. Im, and J.J. Werner, "Design of digital carrierless AM/PM transceivers," *Standard Project, T1E1.4/92-149, AT&T and Bellcore*, August 19, 1992.
- S. Chennakeshu and G.J. Sauliner, "Differential detection of $\pi/4$ -shifted-DQPSK for digital cellular radio," *IEEE Transactions on Vehicular Technology*, vol. 42, pp. 46-57, 1993.
- J.M. Cioffi, V. Oksman, J.-J. Werner, T. Pollet, P.M.P. Spruyt, J.S. Chow, and K.S. Jacobsen, "Very-high-speed digital subscriber lines," *IEEE Communications Magazine*, vol. 37, pp. 72-79, April, 1999.
- L.J. Cimini, Jr., and Y. Li, "Orthogonal frequency division multiplexing for wireless communica-

- tions," tutorial notes, TU18, International Conference on Communications '99, Vancouver, British Columbia, Canada, June, 1999.
- A.C. Clarke, "Extraterrestrial relays," *Wireless World*, vol. 51, pp. 305-308, October 1945.
- C.E. Cook and H.S. Marsh, "An introduction to spread spectrum," *IEEE Communications Magazine*, vol. 21, no. 2, pp. 8-16, 1983.
- J.P. Costas, "Synchronous communications," *Proceedings of the IRE*, vol. 44, pp. 1713-1718, 1956.
- J.P. Costas, "Poisson, Shannon, and the radio amateur," *Proceedings of the IRE*, vol. 47, pp. 2058-2068, 1959.
- M.G. Crosby, "Frequency modulation noise characteristics," *Proceedings of the IRE*, vol. 25, pp. 472-514, April 1937.
- C.C. Cutler, "Differential quantization of communication signals," United States Patent 2-505-361, 1952.
- C.L. Dammann, L.D. McDaniel and C.L. Maddox, "D2 channel bank—Multiplexing and coding," *Bell System Tech. J.*, vol. 51, pp. 1675-1699, 1972.
- F. Daneshgaran and M. Mondin, "Design of interleavers for turbo codes: Iterative interleaver growth algorithms of polynomial complexity," *IEEE Transactions on Information Theory*, vol. 45, pp. 1845-1859, 1999.
- R. deBuda, "Coherent demodulation of frequency-shift keying with low deviation ratio," *IEEE Trans. on Communications*, vol. COM-20, pp. 429-435, 1972.
- F.E. DeJager, "Deltamodulation, a method of PCM transmission using the 1-unit code," *Phillips Research Reports*, vol. 7, pp. 442-46, 1952.
- F.E. DeJager and C.B. Dekker, "Tamed frequency modulation: A novel method to achieve spectrum economy in digital transmission," *IEEE Transactions on Communications*, vol. COM-26, pp. 534-542, 1978.
- J.A. Develet, "A threshold criterion for phase-lock demodulation," *Proceedings of the IEEE*, vol. 51, pp. 349-356, 1963.
- W. Diffie and M.E. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, vol. IT-22, pp. 644-654, 1976.
- W. Diffie and M.E. Hellman, "Privacy and authentication: An introduction to cryptography," *Proceedings of the IEEE*, vol. 67, pp. 397-427, 1979.
- D. Divsalar, "Turbo codes," MILCOM 96 tutorial, San Diego, November 1996.
- M.I. Doelz and E.H. Heald, "Minimum shift data communication system," U.S. Patent 2977417, March 1961.
- R.M. Dolby, "An audio reduction system," *Journal of the Audio Engineering Society*, vol. 15, p. 383, 1967.
- J. Dungundji, "Envelopes and pre-envelopes of real wave-forms," *IRE Transactions on Information Theory*, vol. IT-4, pp. 53-57, 1958.
- P. Elias, "Coding for noisy channels," *IRE Convention Record*, Part 4, pp. 37-46, March 1955.
- L.H. Enloe, "Decreasing the threshold in FM by frequency feedback," *Proceedings of the IRE*, vol. 50, pp. 18-30, 1962.
- V.M. Eyuboglu, "Detection of coded modulation signals on linear, severely distorted channels using decision-feedback noise prediction with interleaving," *IEEE Transactions on Communications*, vol. COM-36, pp. 401-409, 1988.
- D.D. Falconer, "Carrierless AM/PM," *Bell Laboratories, Internal Memorandum*, July 3, 1975.
- K. Feher, "MODEMS for emerging digital cellular-mobile radio system," *IEEE Transactions on Vehicular Technology*, vol. 40, pp. 355-365, 1991.
- J.L. Flanagan, M.R. Schroeder, B.S. Atal, R.E. Crochiere, N.S. Jayant, and J.M. Tribollet, "Speech coding," *IEEE Transactions on Communications*, vol. COM-27, pp. 710-737, 1979.
- B. LeFloch, R. Halbert-Lassalle, and D. Castelain, "Digital sound broadcasting to mobile receivers," *IEEE Transactions of Broadcasting*, vol. 35, pp. 493-503, 1989.
- G.D. Forney, Jr., "Maximum likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transactions on Information Theory*, vol. IT-18, pp. 363-378, 1972.
- G.D. Forney, Jr. "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, pp. 268-278, 1973.

- G.D. Forney, Jr., and M.V. Eyuboglu, "Combined equalization and coding using precoding," *IEEE Communications Magazine*, vol. 29, no. 12, pp. 25-34, 1991.
- G.D. Forney, Jr., L. Brown, M.V. Eyuboglu, and J.L. Moran III, "The V.34 high-speed modem standard," *IEEE Communications Magazine*, pp. 28-93, December 1996.
- L.E. Franks, "Carrier and bit synchronization in data communications—A tutorial review," *IEEE Transactions on Communications*, vol. COM-28, pp. 1107-1121, 1980.
- B.J. Frey and D.J.C. MacKay, "Irregular turbocodes," *Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing*, Allerton House, Illinois, September 1999.
- H.T. Friis, "Noise figures in radio receivers," *Proceedings of the IRE*, vol. 32, pp. 419-422, 1944.
- K.E. Fulz and D.B. Penick, "T1 carrier system," *Bell System Tech. J.*, vol. 44, pp. 1405-1451, 1965.
- D. Gabor, "Theory of communications," *Journal of IEE (London)*, vol. 93, Part III, pp. 429-457, 1946.
- D.L. Gall, "MPEG: a video compression standard for multimedia applications," *Communications of the ACM*, vol. 34, pp. 47-58, 1991.
- R.G. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, pp. 21-28, 1962.
- W.A. Gardner, "Introduction to Einstein's contribution to time-series analysis," *IEEE ASSP Magazine*, vol. 4, pp. 4-5, October 1987.
- W.A. Gardner and L.E. Franks, "Characterization of cyclostationary random signal processes," *IEEE Transactions on Information Theory*, vol. IT-21, pp. 4-14, 1975.
- D.A. George, "Matched filters for interfering signals," *IEEE Transactions on Information Theory*, vol. IT-11, pp. 153-154, 1965.
- A. Gershho, "Adaptive equalization of highly dispersive channels for data transmission," *Bell System Tech. J.*, vol. 48, pp. 55-70, 1969.
- R.A. Gibby and J.W. Smith, "Some extensions of Nyquist's telegraph transmission theory," *Bell Systems Tech. J.*, vol. 44, pp. 1487-1510, 1965.
- R.D. Gitlin and E.Y. Ho, "The performance of staggered quadrature amplitude modulation in the presence of phase jitter," *IEEE Transactions on Communications*, vol. COM-23, pp. 348-352, 1975.
- M.J.E. Golay, "Note on digital coding," *Proceedings of the IRE*, vol. 37, p. 657, 1949.
- M.J.E. Golay, "Binary coding," *IRE Transactions on Information Theory*, vol. PGIT-4, pp. 23-28, 1954.
- R. Gold, "Optimal binary sequences for spread spectrum multiplexing," *IEEE Transactions on Information Theory*, vol. IT-13, pp. 619-621, 1967.
- R. Gold, "Maximal recursive sequences with 3-valued recursive cross correlation functions," *IEEE Transactions on Information Theory*, vol. IT-14, pp. 154-156, 1968.
- B. Goode, "Scanning the issue: Special issue on global information infrastructure," *Proceedings of the IEEE*, vol. 85, pp. 1883-1886, 1997.
- R.M. Gray, "Vector quantization," *IEEE ASSP Magazine*, vol. 1, no. 2, pp. 4-29, 1984.
- W.J. Gruen, "Theory of AFC synchronization," *Proceedings of the IRE*, vol. 41, pp. 1043-1048, 1953.
- P. Guinand and J. Lodge, "Trellis termination for turbo encoders," *Proceedings of 18th Biennial Symposium on Communications, Queen's University*, Kingston, Canada, June 1996.
- D.W. Hagelbarger, "Recurrent codes: Easily mechanized, burst-correcting binary codes," *Bell System Tech. J.*, vol. 38, pp. 969-984, 1959.
- J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory*, vol. 42, pp. 429-445, 1996.
- J. Hagenauer and P. Hoeher, "A Viterbi algorithm with soft-decision outputs and its applications," *IEEE Globecom 89*, pp. 47.11-47.17, November 1989, Dallas, Texas.
- R.W. Hamming, "Error detecting and error correcting codes," *Bell System Tech. J.*, vol. 29, pp. 147-160, 1950.
- J.C. Hancock and R.W. Lucky, "Performance of combined amplitude and phase-modulated communication systems," *IRE Transactions on Communication Systems*, vol. CS-8, pp. 232-237, 1960.

- H.H. Hanning and J.W. Pan, "D2 channel bank system aspects," *Bell System Tech. J.*, vol. 51, pp. 1641–1657, 1972.
- H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Transactions on Communications*, vol. COM-20, pp. 774–779, 1972.
- T.V.L. Hartley, "Transmission of information," *Bell System Tech. J.*, vol. 7, pp. 535–563, 1928.
- F.S. Hill, Jr., "On time-domain representations for vestigial sideband signals," *Proceedings of the IEEE*, vol. 62, pp. 1032–1033, 1974.
- D.A. Huffman, "A method for the construction of minimum redundancy codes," *Proceedings of the IRE*, vol. 40, pp. 1098–1101, 1952.
- P.A. Humblet and M.G. Toulis, "The information driveway," *IEEE Communications Magazine*, pp. 64–68, December, 1996.
- G.-H. Im and J.-J. Werner, "Bandwidth-efficient digital transmission over unshielded twisted-pair wiring," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1643–1655, 1995.
- H. Insoe, Y. Yasuda, and J. Murakami, "A telemetering system by code modulation: Δ - Σ modulation," *IRE Transactions on Space Electronics and Telemetry*, vol. SET-8, pp. 204–209, 1962.
- M. Ishizuka and K. Hirade, "Optimum Gaussian filter and deviated-frequency locking scheme for coherent detection of MSK," *IEEE Transactions on Communications*, vol. COM-28, pp. 850–857, 1980.
- I.M. Jacobs, "Practical applications of coding," *IEEE Transactions on Information Theory*, vol. IT-20, pp. 305–310, 1974.
- N.S. Jayant, "Adaptive delta modulation with a one-bit memory," *Bell System Tech. J.*, vol. 49, pp. 321–342, 1970.
- N.S. Jayant, "Digital coding of speech waveforms, PCM, DPCM and DM quantizers," *Proceedings of the IEEE*, vol. 62, pp. 611–632, 1974.
- N.S. Jayant, "Coding speech at low bit rates," *IEEE Spectrum*, vol. 23, no. 8, pp. 58–63, 1986.
- A.J. Jerri, "The Shannon sampling theorem—its various extensions and applications: A tutorial review," *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1565–1596, 1977.
- J.B. Johnson, "Thermal agitation of electricity in conductors," *Physical Review*, second series, vol. 32, pp. 97–109, 1928.
- P. Kabal and S. Pasupathy, "Partial-response signaling," *IEEE Transactions on Communications*, vol. COM-23, pp. 921–934, 1975.
- I. Kalet, "The multitone channel," *IEEE Transactions on Communications*, vol. 37, pp. 119–124, 1989.
- I. Kalet, J.E. Mazo, and B.R. Saltzberg, "The capacity of PCM voiceband channels," *IEEE International Conference on Communications*, pp. 507–511, Geneva, Switzerland, 1993.
- H. Kaneko, "A unified formulation of segment companding laws and synthesis of codes and digital companders," *Bell System Tech. J.*, vol. 49, pp. 1555–1588, 1970.
- A.I. Khintchine, "Korrelationstheorie der stationären stochastischen Prozesse," *Mathematische Annalen*, vol. 1, 109, pp. 415–458, 1934.
- H. Kobayashi, "Correlative level coding and maximum-likelihood decoding," *IEEE Transactions on Information Theory*, vol. IT-17, pp. 586–594, 1971.
- R. Kohn, "Spatial and temporal communication theory using adaptive antenna array," *IEEE Personal Communications*, pp. 28–35, February, 1998.
- E.T. Kretzmer, "Generalization of a technique for binary data communications," *IEEE Transactions on Communication Technology*, vol. COM-14, pp. 67–68, Feb. 1966.
- F.R. Kschischang and B.J. Frey, "Interactive decoding of compound codes by probability propagation in graphical models," *IEEE Journal on Selected Areas in Communication*, vol. 16, pp. 219–230, 1998.
- J.W. Lechleider, "Line codes for digital subscriber lines," *IEEE Communications Magazine*, vol. 27, pp. 25–32, September 1989.
- B.M. Leiner, V.G. Cerf, D.D. Clark, R.E. Kohn, L. Kleinrock, D.C. Lynch, J. Postel, L.G. Roberts, and S. Wolff, "A brief history of the Internet," *Commun. ACM*, vol. 40, pp. 102–108, February 1997.
- A. Lender, "The duobinary technique for high-speed data transmission," *IEEE Transactions on Communications and Electronics*, vol. 82, pp. 214–218, May 1963.

- A. Lender, "Correlative digital communication techniques," *IEEE Transactions on Communication Technology*, vol. COM-12, pp. 128-135, 1964.
- A. Lender, "Correlative level coding for binary-data transmission," *IEEE Spectrum*, vol. 3, no. 2, pp. 104-115, 1966.
- N.-S. Lin and C.-P.J. Tzeng, "Full-duplex data over local loops," *IEEE Communications Magazine*, vol. 26, pp. 31-42, February 1988.
- S. Lin, D.J. Costello, and M.J. Miller, "Automatic-repeat-request error control schemes," *IEEE Communications Magazine*, vol. 22, no. 12, pp. 5-16, 1984.
- Y. Linde, A. Buzo and R.M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. on Communications*, vol. COM-28, pp. 84-95, 1980.
- D. Linden, "A discussion of sampling theorems," *Proceedings of the IRE*, vol. 47, pp. 1219-1226, 1959.
- C.L. Liu and K. Feher, "Noncoherent detection of $\pi/4$ -shifted systems in a CCI-AWGN combined interference environment," *Proceedings of the IEEE 40th Vehicular Technology Conference*, San Francisco, 1989.
- S.P. Lloyd, "Least squares quantization in PCM," unpublished Bell Laboratories Technical Note, 1957. This report was reprinted in *IEEE Transactions on Information Theory*, vol. IT-28, pp. 129-137, 1982.
- J. Lodge, R. Young, P. Hoher, and J. Hagenauer, "Separable MAP 'filters' for the decoding of product and concatenated codes," *Proceedings of the IEEE International Conference on Communications*, pp. 1740-1745, Geneva, Switzerland, May 1993.
- R. W. Lucky, "Automatic equalization for digital communication," *Bell System Tech. J.*, vol. 44, pp. 547-588, 1965.
- R. W. Lucky, "Techniques for adaptive equalization of digital communication systems," *Bell System Tech. J.*, vol. 45, pp. 255-286, 1966.
- R. Lugannani, "Intersymbol interference and probability of error in digital systems," *IEEE Transactions on Information Theory*, vol. IT-15, pp. 682-688, 1969.
- V.H. MacDonald, "Advanced mobile phone service: the cellular concept," *Bell System Tech. J.*, vol. 58, pp. 15-41, 1979.
- D.J.C. MacKay, "Good error-correcting codes based on very sparse matrices," *IEEE Transactions on Information Theory*, vol. 45, pp. 399-431, 1999.
- D.J.C. MacKay and R.M. Neal, "Near Shannon limit performance of low density parity check codes," *Electronics Letters*, vol. 33, No. 6, pp. 457-458, 1997; and vol. 32, no. 18, pp. 1645-1646, 1996.
- D.J.C. MacKay, S.T. Wilson, and M.C. Davey, "Comparison of constructions of irregular Gallager codes," *IEEE Transactions on Communications*, vol. 47, pp. 1449-1454, 1999.
- J. Max, "Quantizing for minimum distortion," *IRE Transactions on Information Theory*, vol. IT-6, pp. 7-12, 1960.
- K. Maxwell, "Asymmetric digital subscriber line: Interim technology for the next forty years," *IEEE Communications Magazine*, vol. 34, pp. 100-106, October 1996.
- R.J. McEliece, D.J.C. MacKay, and J.-F. Cheng, "Turbo coding as an instance of Pearl's belief propagation algorithm," *IEEE Journal on Selected Areas of Communication*, vol. 16, pp. 140-152, 1998.
- D. Mennie, "AM stereo: Five competing options," *IEEE Spectrum*, vol. 15, no. 6, pp. 24-31, 1978.
- M.L. Moher, "Cross-entropy and iterative detection," Ph.D. thesis, Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada, May 1997.
- M.L. Moher and T.A. Gulliver, "Cross-entropy and iterative decoding," *IEEE Transactions on Information Theory*, vol. 44, pp. 3097-3104, 1998.
- P. Monsen, "Feedback equalization for fading dispersive channels," *IEEE Transactions on Information Theory*, vol. IT-17, pp. 56-64, 1971.
- K.H. Mueller and J.J. Werner, "A hardware efficient passband equalizer structure for data transmissions," *IEEE Transactions on Communications*, vol. COM-30, pp. 538-541, 1982.
- K. Murota and K. Hirade, "GMSK modulation for digital mobile radio telephone," *IEEE Transactions on Communications*, vol. COM-29, pp. 1044-1050, 1981.
- E. Murphy, "Whatever happened to AM stereo?" *IEEE Spectrum*, vol. 25, p. 17, 1988.

- P. Noll, "MPEG digital audio coding standards." In V.K. Madisetti and D.B. Williams (editors), *The Digital Signal Processing Handbook*, Piscataway, N.J.: IEEE Press, pp. 40-1-40-28, 1998.
- D.O. North, "An analysis of the factors which determine signal/noise discrimination in pulsed carrier systems," *Proceedings of the IEEE*, vol. 51, pp. 1016-1027, 1963; this paper is a reprint of a classified RCA Report published in 1943.
- H. Nyquist, "Certain factors affecting telegraph speed," *Bell System Tech. J.*, vol. 3, pp. 324-346, 1924.
- H. Nyquist, "Thermal agitation of electric charge in conductors," *Physical Review*, second series, vol. 32, pp. 110-113, 1928.
- H. Nyquist, "Certain topics in telegraph transmission theory," *Transactions of the AIEE*, vol. 47, pp. 617-644, Feb. 1928.
- M.W. Oliphant, "The mobile phone meets the Internet," *IEEE Spectrum*, vol. 36, pp. 20-28, August, 1999.
- B.M. Oliver, J.R. Pierce, and C.E. Shannon, "The philosophy of PCM," *Proceedings of the IRE*, vol. 36, pp. 1324-1331, 1948.
- D.Y. Pan, "Digital audio compression," *Digital Technical Journal*, vol. 5, pp. 1-14, 1993.
- S. Pasupathy, "Nyquist's third criterion," *Proceedings of the IEEE*, vol. 62, pp. 860-861, 1974.
- S. Pasupathy, "Correlative coding—A bandwidth-efficient signaling scheme," *IEEE Communications Magazine*, vol. 15, no. 4, pp. 4-11, 1977.
- S. Pasupathy, "Minimum shift keying—A spectrally efficient modulation," *IEEE Communications Magazine*, vol. 17, no. 4, pp. 14-22, 1979.
- A.J. Paulraj and B.C. Ng, "Space-time modems for wireless personal communications," *IEEE Personal Communications*, pp. 36-48, February, 1998.
- A.J. Paulraj and C.B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, pp. 49-83, November, 1997.
- R.L. Pickholtz, D.L. Schilling, and L.B. Milstein, "Theory of spread-spectrum communications—A tutorial," *IEEE Transactions on Communications*, vol. COM-30, pp. 855-884, 1982.
- R. Price, "Nonlinearly feedback-equalized PAM vs. capacity for noisy filter channels," International Conference on Communications, ICC '72, pp. 22.12-22.17, June 1972, Philadelphia.
- R. Price and P.E. Green, Jr., "A communication technique for multipath channels," *Proceedings of the IRE*, vol. 46, pp. 555-570, 1958.
- J.G. Proakis, "Advances in equalization for intersymbol interference," *Advances in Communications Systems*, edited by A.J. Viterbi, vol. 4, pp. 123-198, Academic Press, 1975.
- S. Qureshi, "Adaptive equalization," *IEEE Communications Magazine*, vol. 20, no. 2, pp. 9-16, March 1982.
- S. Qureshi, "Adaptive equalization," *Proceedings of the IEEE*, vol. 73, pp. 1349-1387, 1985.
- T.A. Ramstad, "Still image compression." In V.K. Madisetti and D.B. Williams (editors), *The Digital Signal Processing Handbook*, Piscataway, N.J.: IEEE Press, pp. 52-1-52-27, 1998.
- I.S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *Journal of SIAM*, vol. 8, pp. 300-304, 1960.
- A.H. Reeves, "The past, present and future of PCM," *IEEE Spectrum*, vol. 12, no. 5, pp. 58-63, 1975.
- S.A. Rhodes, "Effect of noisy phase reference on coherent detection of offset-QPSK signals," *IEEE Transactions on Communications*, vol. COM-22, pp. 1046-1055, 1974.
- S.O. Rice, "Mathematical analysis of random noise," *Bell System Tech. J.*, vol. 23, pp. 282-332, 1944; vol. 24, pp. 46-156, 1945.
- S.O. Rice, "Statistical properties of a sine-wave plus random noise," *Bell System Tech. J.*, vol. 27, pp. 109-157, 1948.
- S.O. Rice, "Envelopes of narrow-band signals," *Proceedings of the IEEE*, vol. 70, pp. 692-699, 1982.
- T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of provably good low-density parity check codes," submitted in 1999 to *IEEE Transactions on Information Theory*.
- R.L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public key cryptosystems," *Communications of the ACM*, vol. 21, pp. 120-126, 1978.
- W.L. Root, "Remarks, mostly historical, on signal detection and signal parameter estimation," *Proceedings of the IEEE*, vol. 75, pp. 1446-1457, 1987.

- A. Ruiz, J.M. Cioffi, and S. Kasturia, "Discrete multiple tone modulation with coset coding for the spectrally shaped channel," *IEEE Transactions on Communications*, vol. 40, pp. 1012–1029, 1992.
- W.D. Rummier, "A new selective fading model—Application to propagation data," *Bell System Tech. J.*, vol. 58, pp. 1037–1071, 1979.
- B.R. Saltzberg, "Comparison of single-carrier and multitone digital modulation for ADSL applications," *IEEE Communications Magazine*, vol. 36, pp. 114–121, November, 1998.
- B.R. Saltzberg, "Performance of an efficient parallel data transmission system," *IEEE Transactions on Communication Technology*, vol. COM-15, pp. 805–811, 1967.
- S.D. Sandberg and M.A. Tzannes, "Overlapped discrete multitone modulation for high speed copper wire communications," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1571–1585, 1995.
- D.V. Sarwate and M.B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proceedings of the IEEE*, vol. 68, pp. 593–619, 1980.
- B. Sayar and S. Pasupathy, "Nyquist 3 pulse shaping in continuous phase modulation," *IEEE Transactions on Communications*, vol. COM-35, pp. 57–67, 1987.
- H.R. Schindler, "Delta modulation," *IEEE Spectrum*, vol. 7, no. 10, pp. 69–78, 1970.
- R.A. Scholz, "The origins of spread-spectrum communications," *IEEE Transactions on Communications*, vol. COM-30, pp. 822–854, May 1982.
- R.A. Scholz, "Notes on spread-spectrum history," *IEEE Transactions on Communications*, vol. COM-31, pp. 82–84, 1983.
- J.S. Schouten, F. DeJager, and J.A. Greefkes, "Delta modulation, a new modulation system for telecommunication," *Phillips Technical Review*, vol. 13, pp. 237–245, 1952.
- C.E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.
- C.E. Shannon, "Communication theory of secrecy systems," *Bell System Tech. J.*, vol. 28, pp. 656–715, 1949.
- C.E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, pp. 10–21, 1949.
- M.K. Simon and D. Divsalar, "On the implementation and performance of single and double differential detection schemes," *IEEE Trans. on Communications*, vol. 40, pp. 278–291, 1992.
- B. Sklar, "A primer on turbo code concepts," *IEEE Communications Magazine*, vol. 35, pp. 94–102, December 1997.
- B. Sklar, "A structural overview of digital communications—A tutorial review," Part I, *IEEE Communications Magazine*, vol. 21, no. 5, pp. 4–17, 1983; Part II, vol. 21, no. 7, pp. 6–21, 1983.
- D. Slepian, "On bandwidth," *Proceedings of the IEEE*, vol. 64, pp. 292–300, 1976.
- B. Smith, "Instantaneous companding of quantized signals," *Bell System Tech. J.*, vol. 36, pp. 653–709, 1957.
- E.S. Sousa and S. Pasupathy, "Pulse shape design for teletext data transmission," *IEEE Trans. on Communications*, vol. COM-31, pp. 871–878, 1983.
- S. Stein, "Unified analysis of certain coherent and noncoherent binary communication systems," *IEEE Transactions on Information Theory*, vol. IT-10, pp. 43–51, 1964.
- C.E. Sundberg, "Continuous phase modulation," *IEEE Communications Magazine*, vol. 24, no. 4, pp. 25–38, 1986.
- M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronics Letters*, vol. 7, pp. 138–139, March 1971.
- D.W. Tufts, "Nyquist's problem—The joint optimization of transmitter and receiver in pulse amplitude modulation," *Proceedings of the IEEE*, vol. 53, pp. 248–259, 1965.
- G.L. Turin, "An introduction to matched filters," *IRE Transactions on Information Theory*, vol. IT-6, pp. 311–329, 1960.
- G.L. Turin, "An introduction to digital matched filters," *Proceedings of the IEEE*, vol. 64, pp. 1092–1112, 1976.
- G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, IT-28, pp. 55–67, 1982.

- G. Ungerboeck, "Trellis-coded modulation with redundant signal sets," Parts 1 and 2, *IEEE Communications Magazine*, vol. 25, no. 2, pp. 5-21, 1987.
- M.C. Valenti, "An introduction to turbo codes," EE Department, Virginia Polytechnic Institute & State University, Blacksburg, Virginia, unpublished, 1998.
- B. van der Pol, "The fundamental principles of frequency modulation," *Journal of IEE (London)*, vol. 93, part III, pp. 253-258, 1946.
- J.H. Van Vleck and D. Middleton, "A theoretical comparison of visual, aural, and meter reception of pulsed signals in the presence of noise," *Journal of Applied Physics*, vol. 17, pp. 940-971, 1946.
- A.J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. on Information Theory*, vol. IT-13, pp. 260-269, 1967.
- A.J. Viterbi, "Spread-spectrum communications—Myths and realities," *IEEE Communications Magazine*, vol. 17, no. 3, pp. 11-18, May 1979.
- A.J. Viterbi, "When not to spread spectrum—A sequel," *IEEE Communications Magazine*, vol. 23, no. 4, pp. 12-17, 1985.
- A.J. Viterbi, "Wireless digital communication: A view based on three lessons learned," *IEEE Communications Magazine*, vol. 29, no. 9, pp. 33-36, 1991.
- G.K. Wallace, "The JPEG still picture compression standard," *Communications of the ACM*, vol. 34, pp. 31-44, 1991.
- D.K. Weaver, Jr., "A third method of generation and detection of single-sideband signals," *Proceedings of the IRE*, vol. 44, pp. 1703-1705, 1956.
- J. Weiss and D. Schremp, "Putting data on a diet," *IEEE Spectrum*, vol. 30, pp. 36-39, August 1993.
- T.A. Welch, "A technique for high performance data compression," *Computer*, vol. 17, no. 6, pp. 8-19, 1984.
- L.-F. Wei, "Rotationally invariant convolutional channel coding with expanded signal space—part I: 180°," *IEEE Journal on Selected Areas in Communications*, vol. SAC-2, pp. 659-671, 1984.
- L.-F. Wei, "Rotationally invariant convolutional channel coding with expanded signal space—part II: nonlinear codes," *IEEE Journal on Selected Areas in Communications*, vol. SAC-2, pp. 672-686, 1984.
- L.-F. Wei, "Trellis-coded modulation with multidimensional constellations," *IEEE Transactions on Information Theory*, vol. IT-33, pp. 483-501, 1987.
- S.B. Weinstein, "Echo cancellation in the telephone network," *IEEE Communications Magazine*, vol. 15, no. 1, pp. 8-15, 1977.
- S.B. Weinstein and P.M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Transactions on Communications*, vol. COM-19, pp. 628-634, 1971.
- J.J. Werner, "Tutorial on carrierless AM/PM—Part I: Fundamentals and digital CAP transmitter," *AT&T Bell Laboratories Report*, Minneapolis, June 23, 1992.
- J.J. Werner, "Tutorial on carrierless AM/PM—Part II: Performance of bandwidth-efficient line codes," *AT&T Bell Laboratories Report*, Middletown, February 6, 1993.
- B. Widrow and M.E. Hoff, Jr., "Adaptive switching circuits," *WESCON Convention Record*, Pt. 4, pp. 96-104, 1960.
- J.H. Winters, "Smart antennas for wireless systems," *IEEE Personal Communications*, pp. 23-27, February, 1998.
- J.H. Winters, "Adaptive antennas for wireless communications," International Conference on Communications '99, Tutorial Notes TUS, Vancouver, June 6, 1999.
- A.D. Wyner, "Fundamental limits in information theory," *Proceedings of the IEEE*, vol. 69, pp. 239-251, 1981.
- J.L. Yen, "On the non-uniform sampling of bandwidth-limited signals," *IRE Transactions on Circuit Theory*, vol. CT-3, pp. 251-257, 1956.
- O.C. Yue, R. Luganani, and S.O. Rice, "Series approximations for the amplitude distribution and density of shot processes," *IEEE Transactions on Communications*, vol. COM-26, pp. 45-54, 1978.

-
- N. Zervos and I. Kalet, "Optimized decision feedback equalization versus optimized orthogonal frequency division multiplexing for high-speed data transmission over the local cable network," International Conference on Communications, ICC '89, pp. 35.2.1-35.2.6, June 1989.
- J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Transactions on Information Theory*, vol. IT-23, pp. 337-343, 1977.
- J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE Transactions on Information Theory*, vol. IT-24, pp. 530-536, 1978.
- W.Y. Zou and Y. Wu, "COFDM: An overview," *IEEE Transactions on Broadcasting*, vol. 41, pp. 1-5, 1995.

索引

- ADM; ;199 ~ 202
ADM 算法;199
ADM 系统;200 ~ 202
ADSL;见非对称数字用户线
AM(幅度调制);16;76 ~ 79;640
~ 的定义;76
~ 的缺陷;78 ~ 79
~ 的优点;138 ~ 139
~ 的类型;138 ~ 139
~ 的优点;78
AM-PM 转换;108
AM 接收机
~ 的比较;115 ~ 116
包络检波;114 ~ 116
~ 的性能;115 ~ 116
AM 系统
~ 的噪声分析;114
~ 的过程;76 ~ 77
AM 信号
~ 的解调;138 ~ 139
与 FM 信号对比;139
~ ;75 ~ 76
ARPANET;(美国国防部)高级研究计划局计算机网络
~ 和计算机通信中的碰撞;23
~ 和开创工作;23
ARQ;见自动重发请求
ASCII;见美国国家信息交换标准码
ATM;见异步传输模式
AWGN;见加性高斯白噪声
AWGN 模型;452
AWGN 信道;489
容量;376 ~ 377
~ 和平均符号差错概率;288
~ 的接收信号;351
~ 和信号检测;285
~ 和信号传输;268
- ~ 中的技术;382
A 律
~ 的性能;174
~ 的定义;173;174
BCH 码;572 ~ 573
BCJR 算法
~ 的表示式;594
~ 的基础;597
~ 的基本假设;594
~ 与维特比算法相比;594
B-ISDN;见宽带综合业务数字网
CAP;321;324;331
~ 的带宽;327
~ 的基本思想;321
CAP 接收机
~ 的数字实现;330
~ 的性能改进;330
未知环境下的 ~ ;330
CAP 系统
~ 的应用;330
~ 的基本结构;329
~ 调制;321;324;326
~ 接收机;330
~ 结构;329
~ 发射机;329
CAP 发射机;329
CDMA;见码分多址
CPFSK;见连续相位频移键控
CRC;见循环冗余校验码
DES;见数据加密标准检测
~ 与误码纠正;550
脉冲信号的 ~ ;213
DFT;见离散傅里叶变换
Diffie-Hellman 公开密钥分配系统;
665
Dirichlet 条件;627
DMI;见直接矩阵求逆
- DMT;见离散多音
DonaldDuck 语音效应;84
DPSK;见差分相移键控
DS/BPSK 波形;430;431
DS/MPSK 系统;445
DSB-SC 已调信号;80
DSB-SC 接收机
~ 与 AM 接收机对比;115 ~ 116
~ 的模型;112 ~ 113
DSB-SC 调制;113;114
~ 的定义;79
~ 的产生;80
FDM;见频分复用
FDM 系统
~ 的框图;89
~ 的调制步骤;89 ~ 90
FDMA;见频分多址
FH/MFSK 系统
快慢对比;441
~ 对接收机的干扰影响;439
~ 的符号差错;439
FM(频率调制);16;21;640
~ 的性能;140
~ 的定义;92
间接 ~ ;101 ~ 102
~ 是一个非线性过程;92
~ 的原理;107
FM 门限的降低;129 ~ 131
FM 的门限效应;126 ~ 129
FMFB 接收机;131
FMFB 解调器;129 ~ 131
FM 接收机
~ 的失效点;126 ~ 127
~ 中的干扰抑制;126
~ 的模型;121
~ 的噪声分析;121
~ 中的噪声;121

- ~ 中的门限效应;129
FM 解调器
 带负反馈的 ~ ;129
 ~ 与振荡器类型;129
FM 立体声
 ~ 复用;105 ~ 106
 ~ 技术标准;105
 ~ 的传输;105
FM 系统
 ~ 中的非线性影响;107 ~ 108
 ~ 的噪声分析;121 ~ 125
 ~ 类似于 PPM 系统;165
FM 信号
 ~ ;360
 ~ 的带宽范围;90
 降低调制指数的 ~ ;129 ~ 130
 ~ 的平均功率;97
 ~ 的复包络;96
 ~ 的解调;102 ~ 105
 ~ 的所期望的性质;346
 ~ 的检测;346
 ~ 与 AM 信号的区别;92
 ~ 的传输带宽;99 ~ 101
 ~ 的基本特点;93
 ~ 的产生;101 ~ 102
 ~ 的边频;99
 ~ 的频谱分析;92
 ~ 的频谱;97
FM 信号;边频;94 ~ 95
Friis 公式;461
Friis 自由空间公式
 ~ 的定义;456 ~ 457
 ~ 的应用;457 ~ 458
FSK;见频移键控
FSK 信号;336
GMSK
 ~ 是一种特殊形式的二进制频率调制;347
 性能损失;347
GMSK 调制器
 ~ 的整形脉冲;346
 ~ 与 ISI;347
 GMSK 信号;347
 ~ 的功率谱;347
 频谱收敛性;347 ~ 348
Gold 定理;442
Gold 序列
 ~ 的实现;442
 ~ 的相关特性;443 ~ 444
Gram-Schmidt 正交化
 过程;274 ~ 275
GSM;见全球移动通信系统
I 信道;82
ISI;223 ~ 225;347
 ~ 与比特差错;213
 ~ 的条件;244 ~ 245
 在设计者控制下的 ~ ;231
 ~ 作为一种主要损耗;241
 ~ 的影响;253;256
 ~ 作为干扰的一种形式;256 ~ 257
 减小 ~ ;245
 ~ 和噪声的存在;254
 ~ 峰值失真;249
 ~ 与定时误差;229
JPEG 图像编码标准;6
Kraft-McMillan 不等式;505
Kummer 微分方程;650
Lempel-Ziv 编码;508 ~ 509
Lempel-Ziv 算法;6;540
 ~ 与霍夫曼编码对比;509
 ~ 的定义;508
 按 ~ 执行的编码过程;508
 用于文件压缩的 ~ 标准;509
Lloyd-Max 量化器;170;172
LPC;见线性预测码
Mersenne 素数长度序列;424
MPEG-1 视频编码标准;7
MPEG-1 音频编码标准;7
 ~ 的信道容量;201
 ~ 的运行;202
 ~ 的性能;201
MPEG 音频编码标准;200;203
MSK;337
 ~ 的信号空间图;339 ~ 340
MSK 发射机;344
MSK 接收机;344
 MSK 系统;341
 MSK 信号;341
 ~ 的特征;345 ~ 346
 ~ 的检测;344
 ~ 的差错概率;342
 ~ 的产生;344;346
 ~ 的可能形式;340
 ~ 的功率谱;345;347
 ~ 的优点;345 ~ 346
M 进制 FSK 信号
 ~ 的带宽效率;350 ~ 351
 ~ 所需的信道带宽;350
 ~ 正交信号;350 ~ 351
 ~ 的功率谱;350
 ~ 的谱分析;350
M 进制 PAM 系统
 ~ 的设计复杂性;239
 ~ 的功率要求;238 ~ 239
M 进制 PSK
 ~ 的似然函数;394
 ~ 的功率带宽需求;366
 ~ 的信号星座图;319;366
 ~ 相似谱与带宽特性;366
 ~ 特例;318
 ~ 信号的持续时间;318
 ~ 的符号差错概率;318
 ~ 的带宽效率;320
 ~ 的基带功率谱密度;320
 ~ 的功率谱;320
M 进制 PSK 系统;392
M 进制 QAM 信号;323
M 进制 FSK;349 ~ 351
 ~ 的描述;350
 跳频系统的 ~ ;438
 ~ 的特性;349
M 进制数字调制技术;366
M 进制系统;239
M 进制相干相移键控;318
M 进制信号;351
M 进制 QAM
 ~ 的检测;323
 ~ 中的函数;321
 ~ 的性能;366

- ~ 的符号差错概率;323
 ~ 中的发射能量;323
 M 进制 QAM 信号;321~324
 M 进制正交幅度调制;406
 m 序列
 ~ 的自相关函数;422
 ~ 的平衡特性;422
 ~ 定义;422
 ~ 的性质;422~424
 M 元信号图谱;300
 N 维欧氏空间
 ~ 中的角度;271
 ~ 矢量长度;271
 ~ 矢量;270;271
 N 维矢量;270
 PM;见相位调制
 PM 信号;92
 PAM;16;161~164,203
 ~ 的定义;161
 ~ 和调制器的设计;239
 ~ 与自然抽样;161
 PAM 系统;164
 PAM 信号
 ~ 的产生;161~162
 ~ 的性能;164
 ~ 的抽样;162
 ~ 的传输;163
 ~ 的波形;161~162
 PAM 信号;374
 PCM;166;538
 ~ 的优点;186;203
 ~ 的带宽问题;187
 ~ 的基本过程;166
 ~ 优点的代价;187
 ~ 的定义;172
 ~ 作为一种数字脉冲调制;172
 ~ 的相关性;194
 推荐采用;16
 PCM 接收机;223
 PCM 链路;181
 PCM 系统;187
 ~ 的基本步骤;172
 ~ 的特点;181
 ~ 的描述;172~179
 ~ 与干扰;181
 ~ 的噪声考虑;180~181
 ~ 的操作;182
 PCM 信号;177
 PDM;16;164~165;203
 PPM;16;164
 PPM 系统
 ~ 与 FM 系统对比;165
 ~ 的噪声分析;165
 ~ 的性能;165
 PSK;19
 ~ 与相干系统;429~430
 M 进制 ~ ;300
 PSTN;184;367
 ~ 是线性模拟网络;367
 ~ 上的失真;248
 ~ 有效使用;371
 PVC 护套;13
 QAM;见正交幅度调制
 QPSK 信号
 ~ 幅度波动;315
 ~ 与二进制 PSK 信号;312
 ~ 常用的星座图;315
 ~ 的滤波;314
 干扰结果;345
 ~ 的结论;312~314
 ~ 的相位转换;314
 ~ 的功率谱;313~314
 QPSK 发射机;312~313
 QPSK 接收机;312~313
 QPSK;307~314
 ~ 的特性;308
 ~ 的差错概率;310~312
 使用 ~ 的目的;445
 ~ 的信号空间图;308
 Q 信道;82
 RAKE 接收机;481~482
 ~ 的基本概念;481
 ~ 的构成;481
 是一种分集接收机;481
 ~ 的技术;489
 RS-232 标准;5~6
 RSA 加密算法;667
 RSA 算法;665~667
 RSA 系统;665
 RSA 陷门单向函数;666
 sinc 函数;226
 SNR;见信噪比
 SSB 调制;83~84;114;138
 ~ 的基本操作;87~88
 ~ 的定义;79
 频分复用中的 ~ ;89~90
 SSB 已调信号;83~84
 Strowger 开关;21
 TDMA;见时分多址
 Tomlinson-Harashima 预编码;375
 Turbo 编码;591
 Turbo 编码器;598
 Turbo 码的译码;594~596;598
 Turbo 译码器
 ~ 的基本结构;594~595
 ~ 的复杂度;600
 ~ 的详细情况;598~599
 Turbo 码
 ~ 的组成;591
 ~ 的发展;591
 ~ 的性能;593~594
 ~ 的性质;598;607~608
 ~ 的结束方法;593
 V.32 调制解调器标准
 替换调制方案;367
 ~ 的特点;367~368
 V.32 调制解调器与非冗余编码;369~370
 ~ 中的相位变化;368
 转换为 QPSK 模式;370
 ~ 与网格编码;370
 V.90 调制解调器;375
 VDSL;见超高速数字用户线
 VLSI;见超大规模集成电路
 VSB 调制
 ~ 的定义;84
 ~ 与它在商用电视广播中的作用;86~87
 VSB 调制;84~86;139

- VSB 已调信号
 生成方式;84~85
 ~的正交分量;86
- VSB 滤波器
 ~的频率响应;86
 ~的幅度响应;84~85
- Wiener-Hopf 等式;192
- δ 函数/冲激函数;51
 ~的性质;628
 ~的移位特性;162;276
- δ 函数的重复性;628
- μ 律;173
- $\pi/4$ 移相 DQPSK 符号;316
- $\pi/4$ 移相 DQPSK 信号;317
- $\pi/4$ 移相 QPSK 方案;315
- $\pi/4$ 移相的 QPSK 信号
 ~的解调;317
 ~位于八个可能的相位状态的
 种;315
- 3 电平输出;231~232
- 埃尔米特共轭变换;386
- 安全通信
 在敌对环境下的 ~;420
 ~的需要;652
- 安全信道;653
- 八相 PSK;318
- 白噪声;51~53
 ~的自相关函数;51~52
 ~的特征;51~52
 ~的数学特性;51
 ~的功率谱密度;51~53
- 白噪声过程;52
- 半双工链路;550~551
- 半周期余弦脉冲;339
- 半周期正弦脉冲;340
- 包络
 ~的定义;640
 ~与相位分量;56~58
 ~的类型;640
- 包络检波;86~87;111
- 包络检波器;104;121
 ~的组成;78
 ~的创立;78
- ~中的信息丢失;117
~的性能;116
信号比较;119~120
- 包络失真;76~77
- 包络时延;13
- 保护频带;450
- 保护时间;479
- 保护时隙;385
- 保密;655
- 北美数字时分复用系列;184
- 贝塞尔方程;645
- 贝塞尔函数;645~649
 ~的性质;96~97
 ~与调制指数相对比;96
 ~的特性;646~647
- 贝叶斯准则;512
- 倍频;102
- 倍频器;
 ~的组成;101
 ~的框图;101~102
- 倍频项;134
- 本地回路;367
- 本地振荡器;82
- 本原 BCH 码;572
- 本原多项式;442
- 比特;175;498
- 比特持续时间;218
- 比特率的降低;187
- 比特信噪比;334
- 比特填充;184
- 闭环优化
 过程;483
- 边带;上和下;78
- 边带的发送;138
- 边缘概率分布;510
- 边缘密度;623
- 边界
 预测 ~;288
 ~的应用;288~291
- 编译码器;483
- 编码过程
 ~的操作;7
 ~的步骤;483
- ~的应用;175
- 编码理论;22;593
- 编码脉冲
 模拟调制中的 ~;186
 在数字脉冲调制中的 ~;157
 ~的应用;158
- 编码器
 最优化 ~ 的条件;171
 ~的功能单元;202
 ~的主要部分;482
 ~的运行;566
 ~的状态;578;583
- 编码文本;5
- 编码效率;503
- 编码增益;371
- 遍历过程;34;41
- 变频;87;88~89
- 贝塞尔
 修正贝塞尔方程;647~648
 修正贝塞尔函数;648~649
- 标量量化(均匀和非均匀)
 ~的形式;166
 ~的应用;166
- 标量量化器
 ~的最优化条件;170~172
 ~的设计;170
 ~是一个简单的信号压缩器;
 539
- 标准调制技术;420
- 标准幅度调制;138
- 标准压缩算法;6
- 标准正交矩阵;387
- 并串转换器;388
- 并行编码方案;593
- 波分复用;17
- 波特;238
- 波形;18;
 重要线路码的 ~;175~176
 调制中的 ~;430~431
- 波形失真;86~87
- 坡印廷矢量;454
- 泊松分布;49
- 泊松和式;629

- 捕获;432
 捕获效应;126
 不变性;287
 不归零电平编码;312
 不规则 LDPC 码;607
 不规则 Turbo 码;606;607
 不规则交织器;606~607
 不规则码;606
 不可能事件;618
 不平坦信道;474
 不确定的分解;498
 不相关定理;279;279~280
 步长;166
 步长参数;193
 部分响应信号;233;237
 部分响应信号方案;237
 ~的实现;237
 ~的分类;237
 ~的有用特性;237
 参数函数;45
 抽样;172~173
 抽样定理;157;172
 带限信号的~;159~160
 ~的推导;159~160
 ~的本质;157
 脉冲调制系统的~;159~160
 抽样过程;203
 ~与数字信号处理和数字通信;157
 二进制 PCM 信号生成中的~;166
 ~与脉冲调制系统;157;203
 ~适用于;158
 抽样速率;158;172
 抽样周期;158
 有色噪声信道的信息容量;532
 ~与高斯信道;522~523
 ~的含义;526~527
 ~的系统参量;524
 ~的注水解释;534
 参考天线
 ~的定义;454
 是 -各向同性源;454
 参考信号;487
 残留 AM;94
 层;9
 层与层之间的接口;9
 插入损耗;13
 插值公式;160
 插值函数;160;372
 差错概率;287;435~437
 ~的不变性;287
 噪声信道的~;516
 ~的联合边界;288~291
 条件~;305~306
 差错控制;549
 差错控制编码
 ~的理论;424
 为可靠通信的~;497
 ~技术;549
 ~的技术;607
 编码过程;550
 差错控制码理论;424
 差错率;218
 差错门限;180~181
 差分编码器
 二进制波形的~;361
 ~的组成;368~369
 编码方法;176
 应设定一个参考比特;176
 差分检测器
 ~的组成部分;317
 正切型~;317
 差分量化方法;196
 差分脉冲编码调制;194~197
 ~的基本思想;195
 ~与数字脉冲调制;203
 系统对比;196
 ~与发射机;196
 差分相位编码器;362
 差分 PM;368
 差分相移键控;361~364
 ~的误比特率;364
 ~的产生和检测;362
 接收机;363
 发射机;362~363
 差频项;134
 长途连接;13
 场功率图;456
 超大规模集成电路
 ~的发展;22
 超高速数字用户线
 ~的优点;389~390
 ~的使用;389~390
 超均方误差;251
 超外差;108
 超外差接收机;21;108~109
 ~的组成;108
 频率调制与幅度调制的差异;109~110
 超正交信号;见单工信号
 乘积调制器;80;83;94;429
 乘法器;133;566
 乘积密码;658
 承载信息的信号;25;75
 ~以数字方式;239
 ~乘上 PN 信号;427
 充分统计;279
 冲激函数;51
 冲激响应;574;630
 处理增益
 ~定义;196;435
 ~的得到;196
 触发器;566
 传播;3
 传播时延;470
 传播影响;466~469
 传输安全;429
 传输带宽
 ~的定义;100
 ~瞬时扩展;437
 传输的安全性;429
 传输路径;182
 传输时延;224
 传统电话服务;243
 传统相干二进制 FSK
 ~的误比特率的表示;364
 单比特译码的~;364
 传真机

- ~的基本原理;5
~的作用;5
工作在接收模式下;367
串并转换;389
串话;17
~的产生;241
~的定义;439
损耗;241
~的类型;241
次要信息;197
错误传播
~的可能性消除;236
~现象;234
~性质;654
错误检测比特;4
错误图样;557
代换密码
~的描述;658~659
~的使用;658~659
代价函数;488
代数译码器;551~552
代数码
~的类型;608
带宽;631~634
~的定义;631~632
~效率;302
带宽时间乘积;633
带宽受限信道;2;13
带宽效率;302
~的定义;302
~的图形;526~527
两个独立因素作用的结果;302
带宽噪声权衡;165
带通滤波器;84
带通通信信道;302
带通系统
~的分析;641~644
~的冲激响应;641
带通信号
~的组成部分;637;640~641
希尔伯特变换和~;641
~的表示法;95;637~638
带通信号波形;639
带限高斯白噪声;532
带限信号
~的定义;372
~的抽样定理;160
戴维南等效电路;50
单工信号;297
单极性不归零信号;176
单极性归零信号
~的缺点;176
~的特征;176
单键值(压控)振荡器;335
单密钥加密技术;652
单频干扰;445
单位矩阵;286
单位时延;188~189
单位延时元件;566
单一距离;657
单音频 FM 信号;95
单音频调制
~与窄带 FM 信号;92
~与宽带 FM 信号;93
单字母失真度量;536
导频;84
等概率符号;500
等效噪声温度;51;460
低码重码字;599
低密度奇偶校验
~码;599~602
~的优点;599
~的分组长度;604
~的构造;600
~的译码算法;605
~的译码;605
~的初始化;605
~的最小距离;604
~的共有特性;608
~的统计分析;604
~的步骤;605~606
低噪声放大器;452
狄拉克 δ 函数;见 δ 函数
笛卡儿积;322
递归定理;598
递归对称卷积码;592
递归科斯塔环;396
~的收敛性能;403~404
~的工作过程;400~401
~的相位捕获特性;401
用于相位同步的~;396
递归算法
用于相位恢复的~;396
用于定时恢复的~;399~400
递归早~晚延迟同步器;404~405
点对点通信;2
电报;21
电报码;21
电磁干扰;14
电话电路与频率范围;2
电话手柄;13
电话网;22
电路交换网络的一个例子;8
商业电话;22
~的主要目的;12
电话信道;12~15
电路;8
电路交换;8~9
电路交换网络;8
~控制;8
建立连接;8
电视图像;4~5
电视信号;86~88
~的调制方式;86~87
~是宽带信号;6
电信环境;3~7
电子束;4
电子数字计算机;22
叠加原理;630
定时同步;397
定时误差;228
动圈接收机;13
动态多径环境;467
抖动;178
对等进程;9
对角阵;387
对偶码;562
对数似然函数;395
似然函数;279~280

- ~ 的定义;280
 AWGN 信道的 ~ ;282
 ~ 的关系;280
多波束天线;450
多级编码;302;329
多对一映射;7
多回路反馈电路;421
多径
 物理现象;449
 ~ 的存在;449
多径成分;481
多径强度分布;472
多径现象
 ~ ;466 ~ 467
 移动无线环境中的 ~ ;14
 ~ 的本质;466 ~ 467
多径信道
 ~ ;484
 ~ 的分类;474
 ~ 频率选择性;474
 ~ 的模型;60 ~ 61
 ~ 的统计特征;469 ~ 474;485
 ~ 平衰落;61
多径自相关函数;471
复用;89
 ~ 的定义;16
 数字信号的 ~ ;183
 ~ 的类型;16 ~ 17
复用系统;450
复用信号;106 ~ 107
复接器;389
多脉冲激励的线性预测
 编码;482 ~ 483
多频干扰;445
多普勒扩展;472 ~ 474
多普勒频谱;473 ~ 474
多普勒频移;469
多普勒频移平均值;474
多相码;176
多信道传输系统;381
多信道调制;376
 ~ 的基本思想;376
 ~ 的形式;406
多信道数据传输系统;377 ~ 379
多用户通信;449
 ~ 环境;346
多址
 ~ 的基本类型;450
 ~ 与复用相比较;450
多址干扰;480
多址技术
 ~ 的共同特点;450
 ~ 定义;450
 ~ 基本思想;451
 ~ 与通信资源共享;450
多址系统
 ~ 的目的;480
 ~ 的干扰;480
恶性码;585
二阶数字滤波器;397
二次接收机
 ~ 的方程;353
 ~ 的形式;353
二分图;600
二级子空间过程;486
二阶反馈系统;136
二进制
 ~ 的效率;169
 ~ 的符号;176
二进制 BCH 码
 ~ 的一般类型;572 ~ 573
 ~ 与非二进制码的对比;573
二进制 FSK 发射机;334 ~ 335
二进制 FSK 系统;331 ~ 337
二进制 DPSK;362
二进制编码脉冲;166
二进制对称信道;222;551;585
二进制加性流密码;654
二进制假设检验;353
二进制检测问题;352
二进制脉冲;编码;166
二进制 PCM 信号波形;166
二进制 FM;346
二进制 FSK
 ~ 的误比特率;334
 ~ 的差错概率;332 ~ 334
二进制 FSK 信号
 ~ 的基带功率谱密度;336
 连续相位 ~ ;335 ~ 336
 ~ 的检测;337
 ~ 的表示;336
 ~ 的功率谱;335 ~ 336
二进制数据传输系统;247
二进制数据序列;312
二进制数字;175
二进制数字通信系统;351 ~ 353
二进制 PSK;303 ~ 307
 ~ 的差错概率;304 ~ 306
 线性操作 ~ ;430
 ~ 的调制器;429
 ~ 信号;306
 ~ 与扩频调制;481
 ~ 发射机;305 ~ 306
二进制信号
 ~ 的误比特率;475 ~ 476
 ~ 的方案;355
二进制循环冗余校验码
 ~ 的容量;572
 ~ 和误码检测;572
二维匹配滤波器;329 ~ 330
二维最佳接收机;329
发射机
 模拟通信系统的 ~ ;75 ~ 76
 两个基本操作;361
 ~ 位于;1
 ~ 的功率限制;523
 ~ 的作用;2
 ~ 中预加重的使用;131 ~ 132
发射天线
 ~ 的功能;454
 ~ 的安装;14 ~ 15
 点源;454
 ~ 和功率密度;457
发射功率
 ~ 的定义;2
发射滤波器;224
发射信号;241 ~ 242
发射信号点;280 ~ 281
发生概率;498

- 发送的 FH/MFSK 信号;437
发送的电视信号;86
发送码矢量;579
发送脉冲幅度;218
发送脉冲波形;225
反馈系统;二阶;136
反馈移位寄存器;421
反比例;455
反射天线;457
反向链路
 ~与前向链路;489
 ~的子频带;479
反正切计算器;317
泛函
 ~的定义;45
 ~与函数对比;45
方差;507
方向性;456
方向增益;456
放大;108
非对称比;244
非对称调制解调器
 ~的配置;371
 ~的优化设计;371
非对称数字用户线(ADSL)
 将 DMT 用于 ~;389
 ~的目的;244
 ~支持的服务;243
 ~的应用;389
非均匀量化;173 ~ 174
非冗余编码;368
非数据辅助同步;391
非系统码;574
非线性
 ~的基本形式;107
 ~的影响;107
 ~的引起;107
非线性调制;139
非线性预加重和去加重技术;133
非相干 M 进制 FSK 检测器;439
非相干二进制 DPSK;355
非相干二进制 FSK;355;360
 ~的误比特率;361
 ~信号的定义;361
非相干接收机;354;356
非相干匹配滤波器;354
非相干正交调制;355 ~ 356
 加性高斯白噪声;355
 ~的接收机;355
 ~的特例;361
分辨率;4
分布函数
 ~的性质;621
 平稳随机过程的 ~;27
分层结构;9
分段线性近似;174
分集技术;477 ~ 478
 ~是非常有效的;478
 特殊技术;489
分集接收技术;476
分解公式;386
分离定理;597
分离器;243
分离系统;106
分路器;181
分析合成编译码器;483
分组交换
 ~网络;9
 ~的原理;9
分组码;516;553
 ~的特征;560
 ~的码率;560
峰值脉冲信噪比
 ~定义;214 ~ 216
 匹配滤波器的 ~;216
峰值失真;249
蜂窝无线电
 ~的理想模型;465
 ~的主要传播问题;466
 在 ~ 环境的无线通信;465
佛纳姆密码;656
服务质量;12
符号;1
符号波形函数
 定义;306
 ~的能量谱密度;336;345
符号定时;397;400;404 ~ 405
符号定时恢复;404 ~ 405
符号函数;635
符号能量与噪声谱密度的比值;439
符号速率;439
符号同步;391
符号差错概率;291;311
 条件 ~;289
 平均 ~;239
 ~的确定;324
 估算 ~;301 ~ 302
 ~的公式;350
 信号星座图的 ~;292
 与误比特率对比;291 ~ 292
 ~的计算;310
 ~的定义;180;269
 ~的公式;221 ~ 222
 ~的最小化;269;301
 用比值表示 ~;312
 ~的递减;291
幅度谱;637
幅度失真;164
幅度响应;37;532;631
幅移键控(ASK)
 ~调制;299
幅域中的量化;203
辐射强度;455
辐射强度图样;455
辐射效率因数;456
负向咔嗒声;128
复包络;301;637;638;639;640
复傅里叶系数;629
复合调制信号;104
复合信号;89
复指数傅里叶级数;629
复最小均方算法
 ~的优点;489
 ~的局限性;489
 ~;见最小均方算法
傅里叶变换
 ~的定义;627
 ~反变换;627

- 傅里叶分析;627~631
 傅里叶级数展开;275
 傅里叶反变换;43;627
 改进的双二进制码;243
 改进双二进制编码器
 ~的响应;235;236
 ~的有用特性;235~236
 改进双二进制变换滤波器;236
 概率
 ~公理;618~619
 ~的基本性质;618~619
 误比特率;334
 正确判决~;311
 差错~;357
 符号差错概率;285;305
 概率分布;510
 概率概念;617~621
 概率译码器;552
 概率论;617;618
 概率码;608
 概率密度函数;56~57;220;520;
 621~622;623
 概率系统;618
 干扰
 ~的平均功率;433
 扩频码对~的影响;429
 ~和衰落;60~61
 ~的强度;126
 干扰波形;427
 干扰容限;437
 干扰;单频;445
 干扰;多音频;445
 干扰;脉冲噪声;409
 干扰;阻塞噪声;445
 干扰信号;427
 干扰者;433
 ~的策略;433
 ~的类型;445
 ~的波形;445
 感知;3
 感觉编码;7
 高斯白噪声;52
 分布相同;290
 过程;342
 高斯分布;45;61~62
 高斯分布随机变量;45
 高斯过程;45~48
 ~的定义;45
 ~数学证明;46
 ~的有用的性质;46~48
 ~的优点;46
 高斯函数;346
 高斯假设;436
 高斯滤波的MSK;345~349
 高斯滤波器;347
 高斯滤波最小频移键控
 高斯模型;46
 高斯随机变量;45;48
 高斯特性;65
 高斯信道;522
 高性能的CAP系统;326
 格雷码二位组;315
 格雷编码方案;369
 个人电脑;5
 跟踪;432
 跟踪滤波器;131
 公开密钥加密技术;652;663~665
 公开密钥加密系统;664
 公开密钥系统;667
 功率;等效的;51
 功率定理;455
 功率控制
 CDMA系统中的~;480
 ~的使用;480
 功率谱;37;301
 功率谱密度;36~38;301
 ~与幅度谱;41~43
 ~的低频和高频部分;132
 ~的图形总结;64~65
 ~的性质;38~39
 随机过程的~;41;43
 ~的物理意义;37
 功率受限信道;2
 功率增益;456
 天线的~;456
 ~的概念;456
 ~的定义;456
 共信道蜂窝
 ~的测定;465
 最近的~;465
 干扰;484
 固定的点对点的;15
 固定调制方式;549
 固定散射器;470
 固定信道输出;510
 固定信道输入;510
 观察空间;281;282
 观察矢量;279~287
 观察元素;279
 光传输系统;12
 光电探测电路;49
 光电阴极;4
 光栅扫描;4
 光通信;23
 光纤
 ~的优点;14
 ~的组成;14
 ~的特性;14
 ~是一种介质波导;14
 广播方式;2
 广播系统;108
 归一化传输带宽;139
 规则Turbo码;607
 规则脉冲激励;483
 滚降因子;229
 不同~的频率响应;229~230
 不同~的时间响应;229~230
 国家电视系统委员会;4
 过带宽因子;385
 过调幅;77
 汉明距离;558~559;580;584
 汉明重量;558;583
 好码;591
 合成公式;386
 合成分析原理;482
 合成滤波器;483
 ~的组成;483
 ~作为编码器的一部分;482
 和积算法;606

- 和声;3
 恒定包络;94
 恒定角速度;91
 横向均衡器;247
 后检波滤波器;121
 互补误差函数;220;221;290
 互功率谱密度;43
 互相关函数;33;43
 互信息;511~512
 连续集的~;520~522
 ~定义;522
 ~的性质;512~513;522
 香农模型中的~;655
 划分;587
 环路滤波器;133;135~136
 环路增益参数;135
 回波抵消;240
 方案对比;240~241
 ~的操作模式;240
 回波抵消器;240
 收发信机中的~;240~241
 ~的使用;452
 回扫;见水平回扫
 汇合型超几何函数;650~651
 混叠;160
 混叠的频谱;160
 混合调制过程;326
 混频器
 ~的框图;88
 ~的功能;109
 ~的操作;89
 混淆;658
 喉门;3
 霍夫曼编码;506~508;540
 ~的基本思想;506
 相比于 Lempel-Ziv 算法;508
 ~与数据压缩;7
 作为熵编码的~;7
 用于综合的~算法;506
 作为前缀码的一类;506
 ~的缺点;508
 ~的非唯一性;507
 霍夫曼编码过程;506;507
 霍夫曼译码;7
 机载交换;452
 积分
 ~的好处;190
 ~是一种线性运算;190
 积分域;288
 基本不等式;500
 基本函数;393
 基本事件;618
 基带;75
 基带 M 进制 PAM 传输;238~239
 基带二进制 PAM 系统;223
 二进制 PSK 信号的基带功率谱密
 度;307
 计算~;302
 基带二进制数据传输系统;224
 基带扩频系统;428
 基带脉冲;325~326
 基带脉冲传输系统
 ~和固有特性;257
 ~的性能;256
 ~和信噪比;257
 造成比特差错的原因;223~224
 基带信道
 ~的均衡;213
 ~与数字数据传输;213
 基带信号;75;80
 基带信噪比;131
 基频;629
 基本速率;184
 基数;499;505
 基站;465
 奇偶校验多项式
 ~定义;564
 ~的倒数;565
 奇偶校验方程;556
 奇偶校验矩阵;556
 激光;23
 激励发生器;482;483
 激励时间;630
 级联
 ~结构的整体噪声;461
 双端网络的~;460
 几何表示的信噪比;380
 几何平均;380
 计算机产生的数据;6
 计算机通信;1
 既约多项式;442
 加法器;566
 加密技术;652
 加密技术的香农模型
 ~混淆方法;658
 ~扩散方法;658
 ~设计方法;658
 加密术;652
 加密术;652;653
 加权矢量;485
 加性编码调制干扰;432
 加性高斯白噪声;329
 ~干扰;329
 ~和接收机设计;292
 ~下的信号检测;303
 加载方案;382
 加载问题;382
 间接 FM;101~102
 间隙;376
 检测器;283;303
 校验节点;600
 渐近编码增益;585;589
 鉴频器;102;105;122
 ~的组成;102~103
 ~输入;126
 ~的要求;84
 鉴频器输出;123;132
 鉴相器;82
 交换;527
 交换中心;移动的;464
 交替极性反转信号;176
 交织场;4
 校验比特;5;553
 校正子
 ~的计算;566~567
 ~的重要性;557
 ~的性质;557~558
 校正子多项式;567
 校正子计算器;567

- 校正子译码;559~560
 校正子矢量;558
 交织器;数字复接器
 角度;91
 角度调制;16;75
 - ~ 的分类;139
 - ~ 的积分形式;91~92
 - ~ 的重要特点;91
 - ~ 的定义;91
 角度调制信号;337
 - ~ 的定义;91
 - ~ 的波形;75
 角频率;91
 接口;9
 接收机
 - 模拟通信系统的 ~ ;75
 - ~ 的假设条件;429
 - ~ 中的去加重;131
 - ~ 模型;112
 - 动圈 ~ ;13
 - ~ 的噪声性能;337
 最优化的最大似然检测器;356
 - ~ 与预处理接收信号;53
 接收机模型;112
 接收滤波器;224
 接收矢量;579
 接收天线;454
 接收信号
 - ~ 的分量;25
 - ~ 能量的平均值;476
 接收信号点;280
 解调;16;75
 - ~ 方法;84
 - 两步 ~ ;430
 解调类型;111
 解调输出;78
 解密;652
 解析信号;见包络
 近端串话;241;330
 晶体管;49
 晶体振荡器;101
 精确同步;432
 静态图像;3
 精神视觉冗余;7
 静噪作用;124
 镜像干扰;109
 纠错码的种类;549
 纠正 t 个错误的 RS 码;573
 纠正单个错误的汉明码;561
 局部噪声干扰;445
 矩;624
 矩形函数;226
 矩阵变换电路
 - 差信号;106
 - 和信号;106
 距离转移函数;584
 卷积编码;585
 卷积编码器
 - ~ 的码树;576~577
 - ~ 的输入输出关系;578
 - ~ 的状态;576~577
 - ~ 的网格;576~578
 - 采用 ~ ;370
 卷积积分;35~36
 卷积码;574~575
 - ~ 的约束长度;574
 - ~ 的距离特性;583
 - ~ 与分组码;550
 - ~ 的最大似然译码;579~581
 - ~ 的性能;583
 - ~ 的应用;574~575
 绝对熵;520
 均方根(rms)持续时间;633
 均方根带宽;632
 均方失真;171
 均方误差;247
 - ~ 的定义;245
 均方误差准则
 - 接收机设计的 ~ ;245
 均衡器;164
 均匀量化器;168
 均值;28
 卡逊准则
 - 近似计算的 ~ ;139
 - ~ 与标准曲线;101
 开放系统互联模型;9
 开关电平编码器;334
 坎贝儿定理;50
 抗干扰特性;436~437
 抗混叠滤波器;160~161
 科斯塔环
 - ~ 的一般说明;396
 - 用于相位恢复的 ~ ;396
 科斯塔接收机
 - ~ 的组成;82
 - 在 ~ 中的相位控制;83
 - ~ 的工作原理;82
 颗粒噪声;189
 - ~ 和失真;189~190
 - ~ 与量化噪声相对比;190
 可变长度编码;502
 可变节点;600
 克拉默 - 拉奥边界
 - ~ 的定义;403
 - ~ 的修正;404
 空分多址(SDMA);452
 空分多址接入;450
 空间抽样;4
 空间分集;477~478
 空间分集技术;478
 空间现象;468
 空 - 时处理器;486
 孔径效应;163
 控制符
 - ASCII 码中的 ~ ;5
 - 通信用 ~ ;5
 - 字符打印用 ~ ;5
 块密码;653~654
 快速傅里叶算法;387
 快跳频;441~442
 宽带 FM;95~96
 宽带 FM 信号;100
 宽带发射信号;429
 宽带通信信道;187
 宽带网络;11~12
 宽带信号;6
 宽带综合业务数字网;11~12
 - ~ 中的信元;12
 - ~ 的服务质量;12

- ~与电话网;11
扩频
 ~通信;445
 ~的重要属性;427
 ~概念;427~429
扩频调制
 ~的定义;420
 军用目的;420
 ~的原理;420
 防止多径现象;420
 安全的通信;420
扩频技术
 直接序列扩频;429
 在宽带传输中的~;429~430
 ~相比于标准的调制技术;420
扩频通信系统
 ~的主要优点;420
 ~的干扰抑制;420
 ~的要求;432
扩散;658
扩张律;174
扩展码
 ~的平均码字长度;506
 ~的应用;505
扩展码
 具有伪随机性质;429
 ~的使用;429
扩展门限解调器;129;130
扩展前缀码;506
扩展源;501
扩张器;174
拉格朗日因子法;381
莱布尼兹准则;222
莱斯分布
 ~的表示;69~70
 ~的标准化形式;69~70
莱斯衰落信道;470
累积分布函数;621
累积增量;189~190
离散多音(DMT);376;384~387
 ~的应用;389
 ~的基本思想;360
 ~与多信道调制;390~391
离散傅里叶变换(DFT);388
 ~的定义;386
 ~与数字信号处理;387
离散傅里叶反变换(IDFT);386;388
离散脉冲调制;223
离散PAM;223
离散时间傅里叶变换;158
离散随机变量;621
离散无记忆信道;509~511
 ~的信道容量;514~515
 ~的定义;509~510
离散无记忆信源;499
 ~的扩展;501
 ~的性质;498
离散信源;538
离散余弦变换;6
离散余弦变换系数;6~7
里德-索罗蒙(RS)码;573;608
理想的包络检波器;114
理想抽样信号;158
理想的延迟单元;231
理想基带脉冲传输;226
理想鉴频器;105
理想奈奎斯特信道;226~228;229
 ~的困难;227~228
 ~的应用;227~228
理想微分电路
 ~的特点;102~103
 ~的频率响应;103
理想系统;526
理想窄带滤波器;37
立体声复用
 在FM无线广播中的~;105
 ~是一种频分复用技术;105
连续波调制;16
 ~的接收机性能;110
 两种基本的~;75
 ~的原理;138
 ~技术;110
连续波调制系统;75
 ~的比较;111
 ~的组成部分;75
 ~中的噪声;110
 ~的性能;139
连续AWGN信道;276
连续扫描
 图像的~;3~4
 ~过程;3~5
连续随机变量;520;621
连续相位频移键控(CPFSK);332;337
连续相位频移键控(CPFSK)信号
 ~的分量;339
 ~的偏移率;338
 ~的相位;338
 ~的表示;338
连续信号源;539
联合边界;292
 ~的举例说明;289
 ~的简化;291
 ~是实用的上界;288~289
 ~的应用;350
联合分布函数;27;622
联合概率;619;620
联合概率分布;510
联合概率密度函数;622~623
联合矩;626
联合事件;289
链路预算;453
链路预算分析;453
链路预算平衡表;453
两步频谱分析;92~93
两两差错概率;289
亮度信号;5
量化
 ~的应用;167;171~174
 ~与编码;7
 ~的目的;6
 ~的类型;166~167
 ~函数;168
量化表;7
量化过程;166~167;203
 二进制PCM信号产生中的;142
 ~的非线性;170
 ~会导致;16
量化级;166

- 量化器
~的特性;166
~的组成;171
是一种信号压缩器;538
~的类型;166
量化器;非均匀;173
量化器的输入;189
量化噪声;168~169;196
受设计者控制的~;204
增量调制中的~;189
~随时间变化;167
PCM系统的量化噪声;180
临界频带;201
零均值高斯白噪声过程;269
零-零带宽;320;632
流密码;653~655
留存路径;580
六边形蜂窝几何特性;465
滤波;108;179
路径损耗;458
路由器
~的定义;11
~与主机设备;8
~的主要用途;8
率失真定理;536~538
率失真函数;536~538;540
~的应用;540
~的定义;536
马尔可夫过程;594
码;恶性的;585
码本;508
码分复用
~的定义;16
码分复用系统;442
码分多址
~的优点;450
码激励线性预测编码
~的特征;483
~;见码激励线性预测码
码率;517;550;574
码片;428;439
码树 576;576~577
码元;175;182
- 码字;550
平均长度;503
二进制形式的~;503
~持续时间;182
码字长度;平均;503
脉冲;4
脉冲波形函数;324
脉冲调制
~类型;157
~的形式;164~165;203
~的有损性质;203~204
~作为信源编码技术;203
数字~的标准形式;16
~的类型;16
脉冲调制过程
~的信息损失;204
信息丢失与设计者控制~;180
脉冲调制器;181
脉冲调制系统;203
脉冲调制信号;203
脉冲解调器;181
脉冲波形;213
脉冲整形滤波器
~的高斯冲激响应;347
曼彻斯特码;176;243
慢跳频;438~440
慢跳频 M 进制 FSK 系统;439
慢跳频 M 进制 FSK 信号;439
每比特发射信号能量;222
美国标准信息交换码;5
门限扩展;130
门限降低;129
门限效应;116~117
幅度调制和角度调制;138
~中听到的咔嗒声;127~128
~的定义;116;126
在包络检波中的~;116
秘密密钥
~的选择;654
秘密密钥系统;667
密码;652
密码;669
密码分析;652
- ~与授权用户;652~653
~的定义;656
~的描述;652
密码术;652
~和鉴定问题;653
~的数据压缩;658
~的基本假设;655
~和重要性;667
~和保密问题;653
密码系统
~的分类;667
~的组成;653
~的定义;652
密钥;664
密钥加密算法;659
密钥流;654;
明文;652
模 2 π 纠错逻辑;317
模拟到数字的过渡;157
模拟电视;3~4
模拟调制解调器
~的噪声性能;374
模拟脉冲调制;16
~的特征;203
最佳的~形式;165
~的传输;164
~的分类;203
模拟通信系统
~的设计;18
~设计的概念;18
~的建立;18
模拟信号;157
模-数转换;172;330;389
莫尔斯码;21;502
内部信息;595
内层导体;13
内积;272
奈奎斯特带宽;226
奈奎斯特间隔;160
奈奎斯特速率;161;226
奈奎斯特准则;226
能量信号;272
能量间隙;84

- 逆映射;516
 诺顿等效电路;50
 欧拉公式;395
 耦合变换器
 ~ 的定义;240
 ~ 的简化电路;240 ~ 241
 判决错误;253
 判决反馈;233
 判决反馈均衡;252 ~ 253;330;375
 判决反馈均衡器
 ~ 的组成;252
 ~ 中的错误传播;253
 ~ 反馈部分;252
 ~ 前馈部分;252
 判决门限;166
 判决模式;251
 判决设备
 ~ 的设计;239
 判决树;504
 判决准则;281;304 ~ 305;310
 最大后验概率准则;281
 以最大似然准则作为~;281
 陪集首;559 ~ 560
 匹配滤波器;213 ~ 217;248
 频域~;216
 ~ 输出;354
 ~ 的性质;216 ~ 217
 匹配滤波器接收机
 相干接收机与~;283
 ~ 的检测器部分;284
 偏移率;336
 ~ 的定义;101
 ~ 与调制指数;101
 频分多址;450
 频分复用
 ~ 的定义;16 ~ 17;89
 ~ 中调制的方式;89
 频分双工;479
 频率参数;108
 频率分集;477 ~ 478
 频带分离
 ~ 过程;83 ~ 84
 ~ 的应用;84 ~ 85
 频率间隔、时间间隔相关函数;471
 频率解调
 ~ 的定义;102 ~ 103
 ~ 的方法;102 ~ 103
 频率平坦;474
 频率平坦信道;474
 频率响应
 选择~的方法;132
 表示~;37
 频偏;93;130
 频谱;2;3
 频谱分析;93
 频谱方案;302
 频谱扩展
 ~ 与相位调制;430
 频谱零点;320
 频谱效率高;351
 频谱有效调制;302
 频谱整形;302
 频移键控
 基本信号调制方式;299
 ~ 与调制解调器的设计;367
 ~ 与 FM;299
 频域;431
 频域表示;631
 平顶抽样;163
 平方律;165
 平滑;190
 平均时延;473
 平均输出信号;117
 平均输出噪声功率
 ~ 的定义;118
 ~ 的方程;119
 平衰落信道;60 ~ 61
 平坦瑞利衰落信道;485
 平稳过程;27 ~ 28
 ~ 相比于严平稳过程;27
 ~ 的不同名称;29
 平移不变性原理
 ~ 的应用;287
 ~ 的描述;287
 屏蔽门限;7;201;202
 迫零;486
 迫零均衡器;244
 切比雪夫不等式;625
 前馈误码纠错
 前向估计;597
 前向链路;479
 前缀编码;503 ~ 504
 前缀码
 ~ 的定义;503 ~ 504
 前缀条件;504
 切割电平;239
 球体打包;524 ~ 525
 去相关时间;30
 全球覆盖;449
 全球移动通信系统;480
 全双工链路;550
 全余弦滚降特性;229
 热噪声;25;50
 人类听觉系统;7;201
 人类交流;3
 容量;497
 容量边界;527
 冗余
 四种基本~;7
 冗余信息;503
 软判决;552
 软判决译码;586
 瑞利分布;57;59;62
 瑞利能量定理
 ~ 的定义;217
 瑞利衰落信道;470;474
 ~ 上的二进制信号;475 ~ 479
 弱信号抑制;120
 三进制码;175
 散弹噪声;49 ~ 50
 散射波束;60
 散射函数;472
 散射物质;60 ~ 61
 桑德 FSK;332;335 ~ 336;338
 桑德 FSK 信号;336
 扫描;3 ~ 5
 扫描点;4
 色彩感应器
 在人眼中的~;5

- 三种 ~ ;5
 色差信号;5
 熵
 条件 ~ ;511
 ~ 的定义;499
 ~ 的表达式;498
 ~ 的性质;499 ~ 500
 条件熵;512
 熵编码冗余;7
 上边界;350
 上变换;88 ~ 89
 上变频器;89
 上行链路;15,451
 上行数据流的传输;243
 上行转换器;391
 升余弦谱
 ~ 的平坦部分;228 ~ 229
 ~ 的滚降部分;228 ~ 229
 生成多项式;566
 生成方程;556
 声道;3
 声音频谱;3
 声音信号;86
 失真;2
 可接收 ~ ;538
 减少的方法;87
 ~ 的产生;86 ~ 87
 不可避免的 ~ ;536
 失真的衡量;171
 施瓦茨不等式
 ~ 的数学结论;215
 ~ 的证明;272 ~ 274
 十字星座;321 ~ 322;324
 时变转移函数;470
 时分复用(TDM);181;182
 ~ 是无线通信系统;479 ~ 481
 ~ 的概念;181
 定义为;16
 ~ 的应用;181 ~ 182
 时间守恒;181 ~ 182
 时间 - 带宽积;632 ~ 633
 ~ 的选择;347
 时间分集;477 ~ 478
 时间平均自相关函数;35,42
 时间响应;229
 时间压缩复用
 ~ 方案;240 ~ 241
 ~ 的应用;240 ~ 241
 时间函数;245
 时 - 频映射;7
 时 - 频映射网络;202
 时隙;450
 时延扩展;472
 定义;473
 时域离散卷积;550
 时域离散无记忆高斯信道;523
 时域连续信道分割;377 ~ 380
 时域描述;632
 矢量;554
 矢量乘积;597
 矢量量化器
 ~ 的优点;539
 ~ 中编码过程;539
 ~ 与标量量化器对比;539
 ~ 的信号与量化噪声比;539
 视觉暂留;4
 视频带宽;4
 视频点播;7;244
 视频信号;4 ~ 5
 收发信机;240
 输出电流;3 ~ 4
 输出信噪比
 AM 接收机的 ~ ;115
 ~ 的计算;128
 ~ 与载噪比相对比;128
 ~ 的定义;110 ~ 111;123
 估算 ~ ;124 ~ 125
 ~ 中的改善因子;132
 ~ 的增大;132
 均匀量化器的 ~ ;169
 输出字符集;510
 输入信噪比;113
 ~ 的定义;111
 输入字符集;509
 数据比特
 ~ 的二进制形式;5
 用于错误检测的 ~ ;5
 数据传输系统
 同步与异步对比;5 ~ 6
 ~ 的性能;253
 数据复接器;6
 数据辅助同步;391
 数据加密;540
 数据加密标准;659 ~ 662
 数据通信;6
 数据网络;9
 数据信号速率;371
 数据压缩;538 ~ 539
 ~ 的形式;6
 ~ 的概念;538
 ~ 的使用原因;538
 ~ 技术;187
 数据压缩器;538
 数控振荡器;400
 数模转换器;389
 数学模型
 ~ 的分类;25
 用概率术语描述的 ~ ;25
 数字传输设备;184
 数字通带传输系统;299
 评价 ~ 的性能;291
 数字电路技术;162
 数字调制方法
 ~ 误码的概率;364
 ~ 的类型;300
 采用单载波的 ~ ;364 ~ 366
 ~ 的优点;301
 数字调制技术
 ~ 的操作;391
 ~ 的类型;300 ~ 301
 数字调制解调器
 双向的 ~ ;374
 ~ 与数据率;374
 ~ 的一个实现;372 ~ 373
 数字复接器;183 ~ 185
 ~ 的设计问题;184
 ~ 的主要分类;184
 数字复接 - 分接原理;184
 数字 PSTN;371

- 数字交换机;184;389
 数字滤波器;二阶;396~397
 数字脉冲调制
 ~的基本形式;166
 ~的优点;203
 ~的传输;157
 数字签名
 电子邮件系统的~;666~667
 ~的性质;667
 ~的应用;667
 数字数据的基带传输;223
 数字通信
 ~的基本形式;268~269
 ~与误比特率;19
 ~与设计目标;307
 接收机;292
 ~的可靠性;18
 ~的要求;18
 ~与系统设计;18
 设计者的任务;549
 数字卫星通信;366
 数字无线通信系统;482~483
 数字系列;184
 数字信号;184
 零次群~;184
 数字音频广播;391
 数字用户线(DSL)
 用于~的线路码;242~243
 ~的操作模式;239
 ~与双绞线;239;257
 衰落速率;474
 衰落信道
 多径~;469~472
 ~的特征;474
 ~的影响;478
 所分配频带;131;132
 双二进制码;改进的;243
 双二进制编码;233
 双二进制变换滤波器;231~232
 双二进制编码器;231
 双二进制信号方案;231~234
 ~的频率响应;232
 ~技术;235
 双二进制技术;237
 双极性不归零(NRZ)电平编码器;
 306
 双极性不归零(NRZ)信号
 ~的缺点;176
 双极性归零(RZ)信号;176
 双极性码;243
 双绞线对电缆
 ~与同轴电缆对比;13~14
 ~的组成;13
 易受电磁干扰影响;13
 ~的使用;240
 双密钥密码;652
 双向传输;90
 水平回扫;4
 瞬时抽样;158
 瞬时频率;93
 ~的定义;139
 ~的表达式;93
 四进制系统
 ~的眼图;254~255
 四位组;368
 似然函数;279
 随机变量;520,621~623
 ~的定义;621
 ~的描述;621~622
 ~的分布;46
 ~的数学期望;623
 ~的均值;623
 ~的标准偏差;625
 ~的方差;625
 随机二进制序列;422
 随机运动;25
 随机过程
 ~的平均功率;534
 ~的分类;63
 ~的定义;25
 ~的集平均;34
 ~的数学期望;34
 线性系统中的~;35~36
 ~的数学定义;26
 ~参数;63
 ~的属性;25
 通过线性时不变滤波
 器的~;35~36
 随机过程;26
 随机过程的谱分析理论;38
 随机矢量;520
 随机实验
 ~的描述;617
 ~的特征;617
 随机跳频;437
 锁定频率范围;136
 锁相环;102
 ~的复杂度;136
 ~的组成部分;133~136
 ~的缺点;136
 ~中的环路滤波器;133
 ~的模型;134~136
 ~的非线性模型;135
 ~的理解;133;134
 ~的使用;133~136
 锁相环解调器;129
 ~与门限扩展;131
 ~是一种跟踪滤波器;131
 锁相环理论;133
 特征函数;625
 提取;225
 天线
 ~的设计;14
 ~的多波束应用;450
 ~的接收端;14
 天线;多波束;450
 天线阵列;484
 天线中的互易原理;456
 条件差错概率;220;221
 条件概率;619~621
 条件概率密度函数;278~279;334
 条件平均要求;172
 跳频/M进制FSK系统
 跳频;437
 跳频扩频;437;438;
 跳频速率;439
 调制;16
 调制百分比;77
 调制方式;带宽转换;307

- 调制方式;86
 调制过程;16~17
 ~的分类;16
 ~的定义;75
 调制解调器;传真;367
 调制解调器;6
 ~的配置;367
 ~的各部分;367
 调制解调器的对称配置;367~371
 调制器~解调器;367
 调制系统;二进制编码的;169
 调制信号(波);75
 调制指数
 ~的定义;93
 ~的限制;94
 ~很小;100
 ~的值;99~100
 听觉机制;3
 听觉系统;202
 听觉屏蔽;7
 听觉屏蔽现象;201
 停止~等待策略;551
 停止~等待自动重发请求;550
 通带传输模型;302~303
 通带基本函数
 ~的性质;379~380
 ~的时间上的变化;325
 通带脉冲;326
 通带数据传输;299
 ~的替换技术;406
 ~的应用;299
 用于~的通信信道;299
 非线性信道的~;299
 通带数据传输系统
 ~的带宽效率的决定因素;302
 ~的目标;302
 通带调制;640
 通带同相滤波器;326;329
 通带同相脉冲;326
 通带线路码;299
 通带正交脉冲 326
 通~断信号;176
 通信;1
 ~的应用;1
 ~频率;3
 ~的基本原理;1~2
 ~的类型;17~18
 通信过程;1~2
 通信链路
 ~的分析;453
 在电路交换中的~;8
 通信网;8~12
 通信卫星
 在对地静止轨道上的~;15
 ~的历史记录;21~23
 ~的作用;15
 “第二代”~;451
 通信系统
 ~的一般特征;2
 ~的设计;17
 ~的基本部分;1
 ~的噪声分析;2~3;51~54;458
 ~的主要资源;2
 ~的作用;16;75
 通信系统的设计人员;18
 通信信道;6;12~15;239
 ~的分类;2;12~15
 ~的描述;12~15
 ~的使用;75
 同步;391~393;432
 计算方法(现代);392
 ~的典型方法;392
 ~的实现;392
 是统计参数估计问题;392
 TDM 系统中的~;182
 发射机与接收机时钟的~;182
 同步光纤网;12
 同步解调
 同步脉冲;4,182
 同步问题
 解决~的方法;392
 同相分量;79
 ~的功率谱密度;345
 ~的特性;54~55
 ~的表示法;56~58
 同相相干检测器;82
 同轴电缆
 ~的应用;14
 ~的组成;13
 ~与双绞线对比;13
 统计规律;617
 统计平均;623~626
 统计期望算子;623
 突发错误;572
 突发性;6
 突发噪声;389
 突发噪声干扰;445
 图像
 ~与人的视觉系统;3
 ~的感知;3
 ~作为一种信源;3
 特普利茨性质;193
 外部信息;595
 外层导体;13
 完善的保密性;655~657
 万维网;23
 网格;576~579;
 网格编码
 一种差错控制编码技术;375
 一种前向纠错方式;370
 网格编码调制;586~588
 网格编码器;371
 网格码
 带限信道的~;586~587
 ~的设计;586~587
 网络;互联;9
 网络;8
 网络资源;9;12
 微分电路;102
 微分器;121
 微分网络;121
 微蜂窝;465
 惟一译码;502
 维特比算法;580~583
 ~应用中的难题;580
 ~是最大似然译码;581
 ~是最大似然序列估计;581
 伪噪声(PN)序列
 ~的相关性质;443

- 伪噪声(PN)序列;421
~的组成;248
~的产生与性质;421
~作为一种训练序列;248
- 卫星**
通信用~;23
在对地静止轨道上的~;15
~的服务;15
- 卫星通信**;451~453
~的频带;15
全球覆盖;449
~最普遍的频带;451
~是一种多用户用户通信的类型;449
- 卫星通信系统**;15
~的设计;453
~全球覆盖;489
依靠;449
~的信息容量;15
- 卫星信道**
~的覆盖范围;15
遥远地区的链路;15
- 未调载波;100
- 未知相位的信号;351~354
- 翁格博克码;588
8PSK~;588
~的渐近编码增益;589
- 无记忆高斯信道;时域离散;523
- 无记忆量化器;166
- 无记忆信道;279
- 无失真基带二进制传输;225~226
- 无失真数据压缩;6;503
通过~完成;503
- 无数据辅助递归算法;397
- 无数据辅助早-晚延迟同步器;400
- 无损压缩
~的定义;6
数字文本的~;6
~与有损压缩的对比;6
- 无线电传播
自由空间中的~;449
城区内的~;466
- 无线广播信道;14
- 无线链路分析;453~464
- 无线通信;464~469
用~自适应天线阵列;484
~的目标;484
~与正交频分复用;390
~的信源编码;482~484
~是多用户通信的一种;449
~是多用户无线通信系统的一种;464~465
~与有线通信对比;489~490
- 无线通信系统
~的移动性;489
运用MSK的问题;346
- 无线电通信系统;25
- 无线通信系统;479~480
- 无限带宽;526
- 物理层;9
- 误差函数;互补的;220
- 误差信号
~的计算;399~400
~的定义;249;395
用于时钟恢复的~;399~400
~的应用;487
- 误比特率
假设传输路径无误差;179
相干二进制FSK的~;334
数字调制方式的~;364
较低的~;19
信号再生中的~;177
~和符号差错概率;291~292
- 误比特率;364
- 误码矢量;557
- 希尔伯特变换;326;634~636
~的性质;635~636
信号的~;635
希尔伯特变换对;327;635
- 系统的非理想性;2
- 系统记忆时间;630
- 系统设计原则;3
- 系统线性分组码;554
- 下变频;89
- 下变频器;89;391
- 下变频转换器;452
- 下降最快的方向;193
- 下行链路;15
- 下行数据流传输;243~244
- 先验概率;281;510
- 线路码
可用的~;243
~的比较;243
电表示法的~;175~176
~的功率谱;178
~的选择;242~243
~的类型;176
~的应用;175~176
- 线性调制
~的定义;79
~的例子;139
~的形式;79~80
~的类型;79
- 线性调制系统;639
- 线性分集合并结构;478
- 线性分组码
~的基本性质;555
~的分类;572~573
~的译码过程;560
~的定义;553~554
~的数学结构;554~555
~的最小距离;558
~的标准矩阵;559
- 线性函数;45
- 线性接收机;214
~的设计;244
~中的噪声;112
采用相干检测的~;112
- 线性均衡;486
- 线性时不变滤波器
~的定义;214~216
~的冲激响应;37
作为一种匹配滤波器;216
- 线性时不变系统;630~631
- 线性特性;562
- 线性系统;630
- 线性预测;191~194
- 线性预测编码;482
- 线性预加重和去加重滤波器

- ~的应用;133
 线性增量调制器;189;199
 线性阵列信号处理器
 ~的要求;485
 线性自适应预测;193~194
 线性合并器;478
 限幅器;109;121
 陷门单向函数;664
 相对频率近似;617
 相对熵;520~522
 相对相位变化;465~466
 相对于各向同性源的有效辐射功
 率;456
 相干PSK的直接序列扩频;429~
 432
 ~的基本原理;420
 ~系统;436
 相干带宽;473
 相干二进制FSK;331
 相干二进制FSK系统;335
 ~的特征;332
 ~产生与检测;334~335
 ~接收机;334~335
 相干二进制PSK系统;334
 接收机;306
 信号;306
 相干检测;83;112
 和解调;80
 ~的应用;112
 相干接收机;283~284;285
 相干PSK;303
 相干QPSK;307
 相干QPSK系统
 ~信号;312
 相干MSK
 ~误比特率;343
 ~的表示法;364
 相关电平编码;231~234
 ~的基础;231
 ~的一般形式;237
 ~的基本思想;231
 ~的前提;231
 相关函数;28~34
 相关矩阵;33
 相关器;52
 ~的输入;481
 ~的输出;277~279
 相关系数;297;412
 相量;467
 相位差计算器;317
 相位调制;16;;640
 相位调制方案;321
 相位非线性;108
 相位恢复;392
 相位恢复电路;300
 相位解调;430
 相位连续;338
 相位灵敏度;91
 相位判决;344
 相位谱;627
 相位失真
 ~与入耳;84
 发生~;84
 相位树图;338
 相位网格;338~339
 相位误差
 ~的定义;135
 ~的影响;84
 相位误差发生器;401
 相位响应;631
 相位纠错;317
 香农;克劳德容量定理;535
 ~与通信中的数学理论;22
 ~与数字通信的理论基础;22
 香农的信息容量定理;18~19
 香农第二定理;539
 香农第三定理;524
 关于完善保密性的香农基本
 边界;656
 香农容量定理;535
 香农信息论;540
 相继误差;199
 响应时间;630
 像素冗余;7
 小间隔均衡;248
 消息;132
 消息带宽;77
 消息点;280
 消息频谱
 负频率部分的~;88
 ~;84
 消息信号;112
 ~的描述;2
 ~的产生;2
 消息源;302
 消隐脉冲;4
 系统的比例因子;112
 协方差函数;28~29
 协议
 因特网~;9~11
 ~的类型;9~12
 斜率过载失真;189
 谐波失真;94
 心理声学建模;7
 信道;25
 ~特征;268~269
 ~的频率响应;225
 ~与随机噪声;224
 信道;585;586
 信道编码
 ~的设计目标;516
 ~和映射;516
 信道编码;549
 在设计者的控制下;516
 ~的目标;549
 冗余的引入;516
 马尔可夫假设;597
 信道编码定理;516~517;539
 ~应用;516
 ~的香农第二定理;517
 ~的重要性;518
 ~的不能令人满意的特征;553
 信道带宽
 ~的利用;302
 主要的通信资源;78
 北美采用的~;86
 信道译码
 在设计者的控制下;516
 ~的目标;549

- 逆映射操作;516
 信道矩阵;510
 信道码字;18
 信道的冲激响应;252
 信道模型;110
 信道容量;514~516
 ~的概念;552~553
 离散无记忆信道的~;514
 信道容量定理;524
 信道输出;60
 信道输入;523;534
 信道数据率;550
 信道信噪比;113
 AM的~;114
 ~的定义;112
 ~的表达式;124~125
 信道噪声;25
 无噪声环境;196
 ~;214
 ~与 ISI 同时作用;244
 ~的影响;179;256
 PCM 系统中的~;179;218
 降低~的影响;180
 ~的主要来源;25
 由~引起的不确定性;351
 信号
 ~的定义;3
 噪声中的~检测;280~281
 ~的维数;3~4
 接收与发射;1~2
 信号传输编码器;302
 信号传输译码器;283;303
 信号的带宽;2
 信号的几何表示;269
 信号的平均功率;3
 信号调制方式;299
 信号分离器
 接收机中的~;388~389
 信号间隔;498
 信号检测问题;279
 关于~的似然函数;353
 阐述;280~281
 信号空间表示法
- 干扰信号的~;432
 发射信号的~;432
 信号空间的维度;271;432
 信号空间分析;292
 信号-屏蔽比;201
 信号矢量;270
 信号速率;238
 信号星座图;280~281;292
 圆对称的~;290
 从一维 PCM 符号中构造~;374
 ~的最小欧氏距离;290
 信号-量化噪声比;196
 信号跃变;176
 信号再生;177
 信息;1
 信息传输;509
 信息块;501
 信息论
 ~的基本限制;497
 ~的重要结论;517
 ~作为一种数学学科;497
 信息容量;524;539
 信道的~;522~524
 ~的定义;19;524
 ~估算;524
 信息容量定理;539
 信息矢量;554;579
 信元;11~12
 信元丢失率;12
 信元交换技术;12
 信元时延
 定义;12
 ~变化;12
 信源;3
 信源编码;502
 ~的类型;503
 ~长度上的差异;507
 ~的分割;540
 信源编码定理;502~503;536;539
 ~的平均码长;536
 信源编码器;17;502
 ~功能要求;502
 ~的作用;17~18
- 信源译码器;504
 信源码字;18
 信噪比
 ~的基本定义;3;110~111
 设备输出的~;459
 FMFB 接收机的~;131
 ~较高时;225
 信源的~;459
 信噪比间隙;376
 星座图编码器;387
 行波管放大器;452
 行扫描频率;4
 行矢量;554
 性能受信道噪声的影响;180
 性能指数;192
 修正贝塞尔方程;648
 修正贝塞尔函数;647~649
 虚拟通信;9
 旋律;3
 旋转不变性原理
 ~的举例说明;286~287
 ~的描述;287
 旋转噪声矢量;286
 循环码;562~563
 ~的优点;562
 ~的种类;572~574
 ~的性质;562
 系统形式的~;566
 循环前缀;385
 循环冗余校验码
 用于误码检测的~;572
 ~的生成多项式;572
 循环特性;562
 训练模式;251
 压控振荡器;129~130;133~134
 压扩器;174
 压缩电路;174
 压缩律;173~174
 压缩器;174
 时延功率谱;472
 严平稳随机过程;29
 衍射;14
 眼图

- ~的定义;253
~作为一种评价工具;253
~的结论;253
~与传输系统性能;253
眼睛开口高度;253
样本点;25;618
样本函数;26
样本空间;25;618
一般曲线;100~101
一次填充密码;656
移动性;14
移动无线电;14;464
移动无线电信道
 ~的容量;14
 ~是一种线性时变信道;15
 传播路径;14
已调信号;80
异步传输;6
异步传输模式(ATM);11~12
抑制干扰;126
译码器;389
 最优化的条件;172
 ~的组成;483
 ~的功能;483
 ~的功能单元;201
译码窗口;581
译码错误;579
译码的复杂度;600
译码规则;579
译码过程
 ~的要求;225
译码判决;581
译码球体
 ~的最大个数;525
 球体打包;525
译码算法;594
音乐
 ~是一种信源;3
 ~的结构;3
音乐信号
 ~和信道带宽;3
 ~与话音信号;3
音频调制信号;100~101
- 音频调制解调器;367
 ~与数字用户线对比;389
因特网;9~11
 ~的结构;9~11
 ~的发展;23
 ~的成长;23
 ~协议;9~11
因特网结构
 ~的功能块;9~11
因特网协议;9~11
因特网业务提供商(ISP);367
 ~和与 PSTN 之间的通信;371
 ~与 PSTN;367;371
 ~与音频调制解调器;367
英国广播公司(BBC);22
硬判决解调;586
有色噪声信道;532
有损压缩
 ~的定义;6
 ~更为常用;6
有线传播
 基于~的信道;12
 ~的类型;12
有线电视;14
有限冲激响应(FIR)滤波器;330
有效孔径;456
有意或无意;420
酉矩阵;387
语言交流过程;3
语音编码
 ~的应用;197~198
 ~的设计目的;197
 低比特率的~;197~198
 ~技术;482
PCM~;197
 ~的使用;490
语音的产生;3
语音通信;18
语音信号;3
 ~是双极性的;5
 ~的频带范围;13
预包络
 ~的基本性质;642
- ~的定义;636
~的确定;636
~的正交分量;325
预编码的双二进制方案;233~234
预测;7
预测滤波器;195
预调制低通滤波器;346
预设的门限值;239
原色;5
 用视频信号表示的~;5
 ~的传输;5
圆形星座图;321
远端串话;241
远~近效应;480
约束优化问题
 ~的定义;381
载波;16
载波分量;114
载波同步;391
载波相位;352
载波相位的最大似然估计;396~397
 ~的问题解决;392
载波相位恢复;391;401~404
载频;78
载频调谐;108
载噪比;116~118;122
 定义;117;128
 大小对比;119~120
 ~与信噪比相对比;128
再生;177
再生中继器;177
噪声
 无~;226
 ~的计算;51
 通信系统中的~;49
 ~的定义;2~3
 ~的影响;3
 ~中的信号;280~282
 ~源;3;49
噪声的同相分量;110
噪声等效带宽;633~634
 ~的定义;633~634
噪声分析

- 的比较;139
 接收端的 ~ ;458
 噪声干扰;部分频带;445
 噪声干扰;阻塞;445
 噪声功率;平均
 ~ 的输出计算;128
 ~ 的决定性因素;123 ~ 125
 噪声功率;3;51
 噪声计算;51
 噪声接收机模型;110
 噪声屏蔽;201
 噪声 - 屏蔽比;201
 噪声容限;243
 噪声矢量
 ~ 的协方差矩阵;286
 ~ 的统计特性;286
 噪声系数;458 ~ 460
 噪声信道;有色的;532 ~ 535
 噪声加强;244
 噪声电阻;51
 增量调制(DM);187 ~ 190
 ~ 的优点;188
 ~ 与数字脉冲调制;203
 ~ 的基本原理;187 ~ 188
 ~ 的量化误差;189
 ~ 的简化;190
 ~ 与发射机;196
 增量 - 总和调制;190
 窄带 FM 信号;94 ~ 95
 窄带 FM;93 ~ 95
 窄带过程;54 ~ 55
 窄带相位调制;101
 窄带噪声;53
 ~ 的特性;54
 ~ 的分量;54
 ~ 的影响;54
 ~ 的表示法;53 ~ 55;56 ~ 57
 窄带噪声分析;53 ~ 55
 窄带噪声综合;53 ~ 55
 阵列输出信号;487
 阵列信号处理器;484;486
 振荡器;晶体控制的;101
 整体平均
 ~ 自相关函数;246
 ~ 的估计;34
 用时间平均代替 ~ ;34
 正方形星座图;322
 正交调制;587
 正交分量;79
 ~ 的功率谱密度;336
 ~ 的重要特性;54 ~ 55
 ~ 的作用;79;85
 正交幅度调制(QAM);83
 与 CAP 接收机对比;326
 十字星座图;324
 正交幅度调制器;378 ~ 379
 正交幅度调制(QAM)星座图;322
 正交基本函数;269;274
 MSK 的 ~ ;340
 ~ 的缺点;384
 正交接收机
 ~ 的信道;355 ~ 356
 使用相关器的 ~ ;353
 使用匹配滤波器的 ~ ;353
 正交零化现象;82
 正交零化效应;82
 正交频分复用(OFDM);390 ~ 391
 ~ 的应用;391
 用于广播的 ~ 技术;299
 正交相位相干检测;83
 正交信道;355 ~ 356
 正交载波复用;83;307
 正交载波复用系统;83
 正交噪声分量;110
 正态高斯分布;45 ~ 46
 正弦信号;75
 正弦信号加窄带噪声;58 ~ 59
 正弦调制;94 ~ 95
 正弦调制信号;93
 正弦载波
 ~ 定义;76
 正向咔嗒声;127
 帧;483
 ~ 的组成;4
 ~ 的同步方式;184 ~ 185
 ~ 结构;479
 帧封装;7
 帧 - 封装单元;202
 帧间冗余;7
 直接矩阵倒置;489
 直接判决递归算法;392
 直接序列扩频二进制 PSK 信号;430
 指数律;166
 指针;509
 中频;109
 中心极限定理;61
 ~ 的定义;46
 ~ 和高斯过程;46
 中心矩;625
 重构电平;166
 重构滤波器;161
 重显质量;4
 周期图;42
 周期信号;628 ~ 629
 周期信号的 ~ ;628 ~ 629
 ~ 的性质;627
 ~ 的理论;628
 周期性;628 ~ 629
 逐比特的插入过程;184
 主瓣;320;632
 主机;9
 注水方案;382
 注水解释;534 ~ 535
 著名的香农第三定理;540
 转发器;15
 转发器信道;452
 转接器;181
 转移参量;465
 转移概率;509;539
 转移矩阵;510
 转置密码
 ~ 的描述;659
 ~ 的使用;659
 状态概率;597
 状态图;576
 子网;9
 字节;5
 自动重发请求;550 ~ 551
 自发波动;49

- 自适应;197
自适应差分脉冲编码
 调制(ADPCM);197~199
 ~的描述;197~198
 ~的应用;198
 ~与语音信号;199
自适应接收机;248
自适应均衡
 ~算法;248
 ~的两个选择;330
自适应均衡器
 ~的算法;248~249
 较好的办法;257
 ~的跟踪能力;251
自适应量化;197
 采用后向估计的~;197
 采用前向估计的~;197
 ~要解决的问题;197
自适应量化器;197
自适应滤波;257
自适应天线阵列;487
 ~的应用;489
自适应同步均衡器;248
自适应预测
 采用后向估计的~;198
 ~的缺点;198
 采用前向估计的~;198
 ~的首选方式;198
 多采用APB;197
自适应预测器;193
自相关函数;29~32;36
 ~的定义;29~30
 ~的估算;42
 ~的图形总结;64~65
 ~的性质;30
 ~的重要物理意义;30
自协方差函数;29
自由空间传播模型;454~456
自由空间损耗;458
增量-总和调制;190
阻尼因子;136
阻塞噪声干扰机;445
组合码;599~600
最大功率输出定理;51
最大后验概率检测;594
最大后验概率译码器;594
最大后验概率准则;281
最大似然检测;286;292
最大似然译码准则;579
最大似然判决准则
 AWGN信道的~;282
 ~就是;282
最大似然译码器;280~282
 ~的定义;579~580
 ~的理论;579~580
最大似然准则;281
最大似然检测;301
最佳CAP接收机;329~330
最佳二次接收机;351~353
最佳接收机
 ~是一种相干接收机;283
 ~的设计;269
 ~的检测器部分;284
 ~的子系统;283
最佳接收机子系统;283
最佳均衡器;247
最优判决准则;280~281
最佳同相滤波器;329
最佳线性接收机;244~248
 ~的解释;247
 ~与发射机;247
最佳信息容量;534
最佳正交滤波器;329
最近邻条件;171
最小距离
 ~的考虑;558~559
 ~的定义;558
最小均方(LMS)算法;249~250;487
自适应均衡的~;249
结合使用;257
 ~的等式;249
用于线性自适应预测的~;194
近似于;250
 ~的简化形式;249~250
 ~的总结;250
采用矩阵表示法的~;250
最小均方误差;488
 ~准则;488
 ~均衡器;247
 ~接收机;247~248
最小能量信号;288
最小能量平移;288
最小能量能量;288
最优量化算法;171~172
最优滤波器
 ~的冲激响应;215~216
最优门限;221
最优化问题;382~383