

高等院校电子信息类专业系列教材

数字通信原理

韩玉兵 李 鹏 编著

$$C = W \log_2 (1 + S/N)$$



南京大学出版社

内容简介

数字通信技术是现代电子信息领域较为重要的技术之一,本书从系统性和实用性出发,较全面地介绍了数字通信技术的理论和应用。全书共分9章,主要内容有数字通信的数学基础、各种数字调制技术、AWGN信道最佳接收机、带限信道数字通信、载波与符号同步、信息论基础与信道容量、通信系统链路预算和无线衰落信道等。

本书着重基本概念、基本分析方法、主要算法和工程应用,内容的讨论力求深入浅出和简明扼要,可作为信息与通信工程、电子科学与技术学科各有关专业的研究生和高年级本科生教材,也可供从事通信、雷达、导航与电子对抗等领域的广大工程技术人员学习与参考。为了更好地阅读本书,读者应具备基本的微积分、信号与系统以及概率和随机过程的预备知识等背景知识。

图书在版编目(CIP)数据

数字通信原理 / 韩玉兵, 李鹏编著. — 南京: 南京大学出版社, 2025. 12. — ISBN 978-7-305-29658-1
I. TN914.3
中国国家版本馆 CIP 数据核字第 20254U0X10 号

南京大学出版社

出版发行 南京大学出版社
社 址 南京市汉口路22号 邮 编 210093
书 名 数字通信原理
SHUZI TONGXIN YUANLI
编 著 韩玉兵 李 鹏
责任编辑 吴 华 编辑热线 025-83596997

照 排 南京南琳图文制作有限公司
印 刷
开 本 787 mm×1092 mm 1/16 开 印张 15.25 字数 390 千
版 次 2025 年 12 月第 1 版 2025 年 12 月第 1 次印刷
ISBN 978-7-305-29658-1
定 价 49.80 元

网 址: <http://www.njupco.com>
官方微博: <http://weibo.com/njupco>
微信服务号: NJUYUNSHU
销售咨询热线: (025) 83594756

-
- * 版权所有,侵权必究
 - * 凡购买南大版图书,如有印装质量问题,请与所购图书销售部门联系调换

前 言

随着信息技术的迅猛发展,数字通信作为现代通信技术的核心,已经成为信息与通信工程、电子科学与技术等学科的重要研究领域。数字通信技术不仅在传统的通信、雷达、导航和电子对抗等领域中发挥着关键作用,还在新兴的物联网、5G/6G 通信、卫星通信等应用中展现出巨大的潜力。为了满足广大研究生和工程技术人员对数字通信知识的迫切需求,我们编写了这本《数字通信原理》教材。

本教材旨在系统性地介绍数字通信的基本理论、关键技术及其在实际工程中的应用。全书内容涵盖了数字通信的数学基础、各种数字调制技术、AWGN 信道最佳接收机、带限信道数字通信、载波与符号同步、信息论基础与信道容量、通信系统链路预算以及无线衰落信道等核心内容。通过对这些内容的深入讲解,读者将能够全面掌握数字通信的基本原理和设计方法,并具备解决实际工程问题的能力。同已有数字通信方面的教材相比,本教材具有如下特色:

(1) 系统性强,条理清晰,理论推导严谨

本教材在内容编排上注重系统性,力求做到条理清晰、逻辑严密。围绕信号空间维数、信号向量表示、等效低通表示、最佳检测等核心概念,系统阐述数字调制和解调、最佳检测等。每章内容都经过精心设计,确保前后章节之间的衔接自然流畅。在理论推导方面,我们力求简洁而不失严谨,确保每个结论通俗易懂且具有坚实的数学基础。通过这种方式,读者不仅能够理解数字通信的基本概念,还能够掌握其背后的数学原理,从而为后续的深入研究和工程应用打下坚实的基础。

(2) 内容选择精,包含最新的通信技术进展

数字通信技术日新月异,新的理论和方法不断涌现。为了确保教材内容的时效性和前瞻性,我们在编写过程中特别注重吸收最新的研究成果和技术进展。例如,本书不仅涵盖了传统的数字调制技术(如 PAM、FSK、PSK、QAM、MSK 等),还介绍了近年来在 5G 通信中广泛应用的 OFDM(正交频分复用)和 OTFS(正交时频空间调制)技术。此外,本书还讨论



了无线衰落信道中的最新研究进展,包括多天线 MIMO 技术、时空频分集技术等,帮助读者了解数字通信领域的最新动态。

(3) 创新程度高,贴近工程实际

本教材不仅注重理论知识的传授,还特别强调理论与工程实践的结合。例如,在介绍 AWGN 信道最佳接收机时,结合工程实际详细阐述了数字通信中带宽效率、功率效率和差错性能的权衡。在通信系统链路预算部分,结合实际 Iridium 卫星通信系统和 Starlink 卫星通信系统,详细介绍了链路预算方法,帮助读者掌握如何在实际工程中进行通信系统的设计和性能评估。

本教材适用于“信息与通信工程”和“电子科学与技术”学科的研究生,以及通信、雷达、导航与电子对抗等领域的广大工程技术人员。对于研究生而言,本书可以作为数字通信课程的教材,帮助他们系统地掌握数字通信的基本理论和关键技术。对于工程技术人员而言,本书可以作为参考书,帮助他们解决实际工程中的问题,并了解数字通信领域的最新进展。

在本书编写过程中,我们得到了许多同行和专家的帮助与支持。感谢各位审稿专家对本书提出的宝贵意见和建议。特别感谢与我共同完成教材的李鹏副教授,李兴海博士,刘潇维、张源广、王瑾、朱江林、吴可妍、陈虹瑾、高骏华、蒋中俊等硕士研究生。此外,我们还要感谢出版社的编辑团队,他们的辛勤工作使得本书得以顺利出版。

由于数字通信领域的发展迅速,书中难免存在不足之处,恳请广大读者批评指正。我们希望通过本书的出版,能够为数字通信领域的研究和工程实践贡献一份力量。

编者:韩玉兵

2025 年 7 月

目 录

第1章 绪论	1	3.2 无记忆调制	46
1.1 引言	1	3.3 有记忆调制	56
1.2 数字通信发展历程	1	3.4 扩频调制	63
1.3 数字通信系统组成	3	3.5 多载波调制	72
1.4 通信系统的性能指标	5	3.6 小结	83
1.5 教材内容与学习引导	6	3.7 思考题	83
参考文献	7	参考文献	84
第2章 数字通信的数学基础	8	第4章 AWGN信道最佳接收机	85
2.1 引言	8	4.1 引言	85
2.2 确定信号分析	11	4.2 最佳检测理论	85
2.3 随机信号分析	18	4.3 最佳接收机设计	91
2.4 带通信号与带通系统的表示	26	4.4 典型调制差错性能分析	95
2.5 信号空间与矢量表示	36	4.5 数字调制带宽效率、功率效率和差错性能的权衡	111
2.6 小结	42	4.6 小结	115
2.7 思考题	42	4.7 思考题	116
参考文献	44	参考文献	117
第3章 数字调制	45	第5章 带限信道数字通信	118
3.1 引言	45	5.1 引言	118



5.2 带限信道特性·····	118	解释·····	178
5.3 带限信道信号设计·····	125	7.4 信道容量·····	180
5.4 存在 ISI 时 AWGN 信道的最佳接收机 ·····	135	7.5 小结·····	190
5.5 线性均衡·····	143	7.6 思考题·····	190
5.6 小结·····	150	参考文献 ·····	191
5.7 思考题·····	150	第 8 章 通信系统链路预算 ·····	192
参考文献 ·····	151	8.1 引言·····	192
第 6 章 载波与符号同步 ·····	153	8.2 通信链路预算·····	193
6.1 引言·····	153	8.3 通信系统的其他性能·····	203
6.2 载波同步·····	154	8.4 小结·····	207
6.3 符号同步·····	165	8.5 思考题·····	207
6.4 载波和符号联合同步·····	169	参考文献 ·····	208
6.5 小结·····	170	第 9 章 无线衰落信道 ·····	209
6.6 思考题·····	170	9.1 引言·····	209
参考文献 ·····	171	9.2 无线信道物理建模·····	211
第 7 章 信息论基础与信道容量 ·····	172	9.3 多径衰落信道特性·····	221
7.1 引言·····	172	9.4 信号分集技术·····	227
7.2 信息论基础·····	172	9.5 小结·····	235
7.3 信源编码和信道编码的信息论		9.6 思考题·····	236
		参考文献 ·····	237

第1章 绪论

1.1 引言

数字通信(Digital Communications)是用数字信号作为载体来传输消息,或用数字信号对载波进行数字调制后再传输的通信方式。它可以传输电报、数字数据等数字信号,也可以传输经过数字化处理的语声和图像等模拟信号。

数字通信技术是现代电子信息技术较为重要的领域之一,特别是在现代社会中,人类生活所需的各种信息主要依靠数字通信技术及其设施来处理 and 传输,数字通信技术已渗透到社会许多行业的职业岗位中。《数字通信原理》教材将介绍设计和分析数字通信系统的基本原理,以数字通信技术为主线,对通信系统的基本概念和基本组成,数字通信系统的基带传输、频带传输、同步系统、最佳接收和链路预算等主要技术进行全面、系统的阐述,同时对一些较新的调制解调和同步技术做介绍,其内容与现阶段通信技术发展相吻合。

本书采用信号空间、等效低通和随机过程等分析方法,根据最佳接收准则,先后讨论并分析了在加性高斯白噪声(Additive White Gaussian Noise, AWGN)信道、带限信道(线性滤波信道)以及多径衰落信道等三种基本的典型信道条件下的数字信号可靠且高效传输及其最佳接收问题。从信号传输角度主要介绍了通信信号表示、数字调制、自适应均衡、最佳接收和同步等内容;从信息传输角度介绍了信息论基础、信道容量等内容。

本讲义取材新颖,讨论问题系统全面、逐步深入、概念清晰,理论分析严谨、逻辑性强,参考资料丰富,适合作为电子信息领域“信息与通信工程”和“电子科学与技术”等学科的研究生教材,对相关专业的教师、学生以及科技工作者也是一本很好的参考书。

1.2 数字通信发展历程

有线电报是最早的电通信形式,是现代通信技术的开端。有线电报就是一个数字通信系统,是由 S. 莫尔斯(Samuel Morse)研制并在 1837 年进行了演示试验。莫尔斯设计出一种可变长度的二进制码——莫尔斯码,其中英文字母用点划线的序列(码字)来表示。这种码将较频繁使用的字母用短码字表示,不常使用的字母则用较长的码字来表示。因此,莫尔斯码是可变长度信源编码方法的先驱。

差不多在 40 年之后,1875 年, E. 博多(Emile Baudot)设计出一种电报码——博多码,其中每一个字母编成一个固定长度为 5 的二进制码字。在博多码中,二进制码的元素是等长度的,且指定为传号(二进制的 1)和空号(二进制的 0)。



虽然莫尔斯在研制第一个电的数字通信系统(电报)中起了重要的作用,但是现在所指的现代数字通信系统起源于奈奎斯特(Nyquist, 1924)的研究。奈奎斯特研究出了在给定带宽的电报信道上,无符号间干扰的最大符号传输速率——2倍带宽每秒。现在把这个速率称为奈奎斯特速率。通过采样脉冲 $g(t) = (\sin 2\pi Wt)/2\pi Wt$ 可以达到此速率,其中 W 是带宽。这个研究成果等价于带限信号抽样定理的一种形式,后来香农(1948)准确阐述了抽样定理:带宽为 W 的信号可以由它的奈奎斯特速率采样值通过下面的插值公式来重构:

$$s(t) = \sum_n s\left(\frac{n}{2W}\right) \frac{\sin\left[2\pi W\left(t - \frac{n}{2W}\right)\right]}{2\pi W\left(t - \frac{n}{2W}\right)} \quad (1.2-1)$$

在通信技术发展中,另一个有重大意义的进展是科尔莫哥洛夫(Kolmogorlov, 1939)和维纳(Winer, 1942)的研究,他们研究了在加性噪声 $n(t)$ 存在的情况下,根据对接收信号 $r(t) = s(t) + n(t)$ 的观测来估计期望的信号波形 $s(t)$ 的问题,这个问题出现在信号检测中。维纳得出一个线性滤波器,其输出是对期望信号 $s(t)$ 最好的均方近似。这个滤波器称为最佳线性(科尔莫哥洛夫-维纳)滤波器。

奈奎斯特和哈特利(Hartley, 1928)的关于数字信息最大传输速率的研究成果是香农(Shannon, 1948)研究工作的先导,香农奠定了信息传输的数学基础,并导出对数字通信系统的基本限制。香农在他的开拓性研究中采用了信息源和通信信道的概率模型,以统计术语将可靠的信息传输基本问题表示成公式。在这些统计公式中,他对信源的信息含量使用对数进行度量。香农证明了发送机的功率限制、带宽限制和加性噪声的影响可以和信道联系起来,合并成一个单一的参数,称为信道容量。例如,在加性高斯白(平坦频谱)噪声干扰情况下,一个带宽为 W 的理想带限信道所具有的信道容量为

$$C = W \log_2 \left(1 + \frac{P}{WN_0} \right) \text{ bit/s} \quad (1.2-2)$$

式中, P 是平均发射功率, N_0 是加性噪声的功率谱密度。信道容量的意义如下:如果信源的信息速率 $R < C$, 那么采用适当的编码在信道上实现可靠(无差错)的传输在理论上是可能的。如果 $R > C$, 则不管在发送机和接收机中采用什么样的信号处理,都不可能实现可靠的传输。因此,香农建立了对信息通信的基本限制,并开创了一个新的领域,现在称之为信息论。

对数字通信领域做出重要贡献的另一位科学家是科捷利尼科夫(Kotelnikov, 1947),他用几何的方法提出了对各种各样的数字通信系统的相干分析。科捷利尼科夫的方法后来由沃曾克拉夫特和雅各布斯(Wozencrat & Jacobs, 1965)进一步推广。

在香农的研究成果公布之后,接着就是汉明(Hamming, 1950)对纠错码的经典研究工作,用来克服信道噪声的不利影响。在随后的多年中,许多研究工作者发现了多种新的功能更强的纠错码,如分组码、级联码、卷积码、网格编码、Turbo码和LDPC码等,其中许多码仍用于现代通信系统中。

1973年,美国摩托罗拉工程师马丁·库帕(Martin Lawrence Cooper)发明了单人可携带的移动电话(手机),开启了从1G到6G不断发展的移动通信新时代,其中1G是模拟通信,2G到6G是数字通信。

2G时代以GSM和CDMA为核心技术,首次实现数字信号传输,支持语音通话和短信

功能。3G 采用 WCDMA、CDMA2000 和 TD-SCDMA 等技术,通过宽带 CDMA 实现更高速率的数据传输,为移动互联网奠定了基础。4G 以 OFDMA 和 MIMO 为核心技术,通过 LTE 标准实现全 IP 网络架构,大幅提升数据传输速率,支持高清视频和移动互联网应用。5G 采用毫米波通信、大规模 MIMO、网络切片和边缘计算等创新技术,在增强移动宽带(eMBB)、超高可靠低时延通信(URLLC)和海量机器类通信(mMTC)三大场景实现突破。

正在研发的 6G 技术将太赫兹通信、智能超表面、AI 原生网络和天地一体化网络作为核心技术方向,致力于实现“一极致三融合”的能力,即把 5G 的能力进一步推向极致,并且实现通信与感知一体化(通感融合)、通信与人工智能融合(通智融合)、无人机卫星地面网络融合(空天地融合),为全息通信、数字孪生和元宇宙等未来应用提供技术支撑。每一代通信技术的演进都通过核心技术的突破,推动着通信能力和应用场景的跨越式发展。

在过去的三四十年中,对数据传输需求的增长以及更复杂的集成电路的发展,导致了非常有效的且更可靠的数字通信系统的发展。在这个发展过程中,香农关于信道最大传输速率极限及所达到的性能界限的最初结论及其推广已作为通信系统设计的基准。由香农和其他研究人员导出的理论极限对信息论的发展做出了贡献,并成为设计和开发更有效的数字通信系统不断努力的最终目标。

最后列举一下通信领域的一些著名应用:

- 1837 年,美国人莫尔斯(Morse)发明了莫尔斯电码和有线电报。
- 1876 年,美国人亚历山大·贝尔(Alexander Bell)申请了电话专利,成为了电话之父。而关于电话发明者尚存争议。
- 1896 年,意大利人伽利尔摩·马可尼(Guglielmo Marchese Marconi)实现了人类历史上首次无线电通信,通信距离为 30 米(次年达到 2 英里)。
- 1965 年,美国贝尔公司成功生产了世界上第一台商用存储程序控制交换机,即“程控交换机”。
- 1965 年,美国宇航局发射了半试验、半实用的地球同步轨道卫星“晨鸟”(Early Bird,后改名为“国际通信卫星-I”,Intelsat-I),用于欧美间的商用卫星通信,从此卫星通信进入了实用阶段。
- 1973 年,美国摩托罗拉工程师马丁·库帕(Martin Lawrence Cooper)发明了单人可携带的移动电话(手机),开启了从 1G 到 6G 不断发展的移动通信新时代。
- 1984 年,美国科学家本内特(Charles Henry Bennett)和加拿大科学家布拉萨德(Gilles Brassard)首次提出了利用量子比特作为信息载体、使用量子密钥进行具有绝对安全特性的量子通信。

1.3 数字通信系统组成

图 1.1 展示了一个数字通信系统的功能性框图和基本组成部分。信源输出可以是模拟信号(如音频或视频信号),或者数字信号(如计算机的输出),数字信号在时间上是离散的并且具有有限个输出字符。在数字通信系统中,由信源产生的消息被转换成二进制数字序列。理论上,应当用尽可能少的二进制数字表示信源输出(消息)。换句话说,就是要寻求一种信



源输出的有效的表示方法,使其很少产生或不产生冗余。通常将模拟或数字信源的输出有效地变换成二进制数字序列的处理过程称为信源编码或数据压缩。

由信源编码器输出的二进制数字序列称为信息序列,它被传送到信道编码器。信道编码器的目的是在二进制信息序列中以受控的方式引入一些冗余,以便于在接收机中用来克服信号在信道中传输时受到的噪声和干扰的影响。因此,所增加的冗余是用来提高接收数据的可靠性以及改善接收信号的失真度的,这些冗余有助于接收机译出期望的信息序列。例如,二进制信息序列的一种基本编码形式就是将每个二进制数字简单地重复 m 次(m 为一个正整数)。更复杂的编码则一次取 k 个信息比特并将每种由 k 个比特组合成的序列映射成唯一的 n 比特序列,该序列称为码字。以这种方式对数据进行编码所引入的冗余度的大小是由比率 n/k 来度量的。该比率的倒数,即 k/n ,称为码的速率,简称为码率。

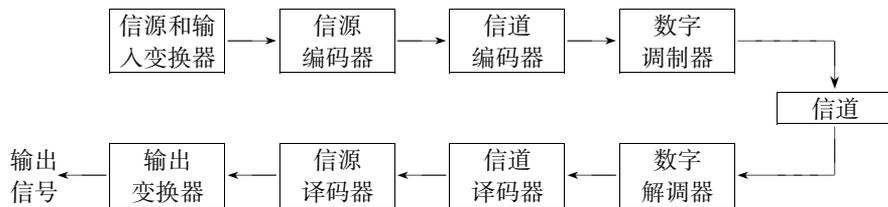


图 1.1 数字通信系统的基本组成部分

信道编码器输出的二进制序列送至数字调制器,数字调制的主要目的是将二进制信息序列映射成适合在信道中传输的信号波形。数字调制可以分为二进制调制和多进制调制。假定已编码的信息序列以某个均匀速率 R 比特/秒一次传输一个比特,数字调制器可以简单地将二进制数字“0”映射成波形 $s_0(t)$,将二进制数字“1”映射成波形 $s_1(t)$ 。在这种方式中,每一比特是分别调制传输的,称为二进制调制。另一种方式中,调制器一次传输 b 个已编码的信息比特,其方法是采用 $M=2^b$ 个不同的波形 $s_i(t), i=0,1,2,\dots,M-1$,分别对应 2^b 种由 b 比特组合成的序列,这种方式称为 M 进制调制($M>2$)。注意,每 b/R 秒就有一个新的 b 比特序列进入调制器。因此,当信道比特率 R 固定时, M 进制调制波形 $s_i(t)$ 的时长是二进制调制波形 $s_0(t)$ 的 b 倍。

通信信道是信号从发射机传输到接收机的物理媒质。在无线传输中,信道通常是大气(自由空间)。而电话信道通常使用各种各样的物理媒质,包括有线线路、光缆和无线(微波)等。无论用什么物理媒质来传输信息,其基本特点是发送信号随机地受到各种可能机理的影响而恶化,例如由电子器件产生的加性热噪声、人为噪声及环境噪声等。

在数字通信系统的接收端,数字解调器对受到信道恶化的发射波形进行解调处理,将该波形还原成数字序列,该序列表示发送数字符号的估计值(二进制或 M 进制)。这个数字序列被送至信道译码器,它根据信道编码器的编码方式及接收数据所含的冗余度来试图重构初始的信息序列。同样,信源译码器也根据信源编码方式进行译码,还原出原始信号。

度量解调器和译码器工作性能好坏的一个指标是译码序列中发生错误的频度,即译码器输出的平均比特错误概率(误比特率)。一般地,错误概率是下列各种因素的函数:码特征、用来在信道上传输信息的波形的类型、发送功率、信道的特征(即噪声的大小、干扰的性质等)以及解调和译码的方法。在后续章节中,将详细讨论这些因素及其对差错性能的影响。

输出变换器作为最后一步,将数字信息转换成所需的输出形式。由于信道译码的差错

以及信源编码器可能引入的失真,在信源译码器输出端的信号只是原始信源输出信号的一个近似。原始信号与重构信号之间的信号差或信号差的函数是数字通信系统引入了多少失真的一种度量指标。

数字通信的发展速度明显超过模拟通信,已成为当代通信的主流。与模拟通信相比,数字通信更能适应现代社会对通信技术越来越高的要求。数字通信的优点在于以下几个方面:

(1) 抗干扰能力强:由于在数字通信中,传输的信号幅度是离散的,以二进制为例,信号的取值只有两个,这样接收端只需判别两种状态。信号在传输过程中受到噪声的干扰,必然会使波形失真,接收端对其进行抽样判决,以辨别是两种状态中的哪一个。只要噪声的大小不足以影响判决的正确性,就能正确接收(再生)。而在模拟通信中,传输的信号幅度是连续变化的,一旦叠加上噪声,即使噪声很小,也很难消除它。数字通信抗噪声性能好,还表现在微波中继通信时,它可以消除噪声积累。这是因为数字信号在每次再生后,只要不判决错码,它仍然像信源发出的信号一样,没有噪声叠加在上面。因此,中继站再多,数字通信仍具有良好的通信质量。而模拟通信中继时,只能增加接收信号能量(对接收信号放大),而不能消除噪声。

(2) 差错可控:数字信号在传输过程中出现的错误(差错),可通过纠错编码技术进行检测和纠正,以提高传输的可靠性。

(3) 易加密:数字信号与模拟信号相比,更容易加密和解密。模拟信号加密需复杂变换,而数字信号可直接通过数学算法加密。因此,数字通信保密性好。

(4) 易于与现代技术相结合:由于计算机技术、数字存贮技术、数字交换技术以及数字处理技术等现代技术飞速发展,许多设备、终端接口均是数字信号,因此,极易与数字通信系统相连接。

1.4 通信系统的性能指标

通信系统的性能指标是衡量其信息传输能力与质量的关键参数,涉及可靠性、有效性、适应性、经济性、标准性、可维护性等。从研究信息传输的角度来说,可靠性和有效性是通信系统的主要性能指标。可靠性和有效性是互相矛盾又对立统一的两个性能,具有此消彼长的可互换特性。

1.4.1 可靠性指标

可靠性是衡量通信系统性能的一个重要指标。它是指在给定信道内接收到的信息的可靠程度,衡量的是系统抗干扰和正确传输的能力。在数字通信中可靠性指标主要指误码率和误比特率。

(1) 误码率

误码率又叫作误符号率,定义为错误码元数与总传输码元数的比值。

$$P_e = \frac{\text{错误码元数}}{\text{总传输码元数}} \quad (1.4-1)$$

(2) 误比特率

误比特率又叫作误信率,定义为错误比特数与总传输比特数的比值。



$$P_b = \frac{\text{错误比特数}}{\text{总传输比特数}} \quad (1.4-2)$$

1.4.2 有效性指标

有效性衡量系统在特定资源下传输信息的能力,即资源利用效率。有效性指标主要包括传输速率、带宽效率和功率效率。

(1) 传输速率

传输速率包括码元传输速率 R_s 和信息传输速率 R 。

码元传输速率又叫作传码率、符号率、波特率,定义为每秒传送的码元个数。单位为符号/s,或者波特(Baud)。若每个码元时间长度为 T_s ,则码元传输速率 R_s 为

$$R_s = \frac{1}{T_s} \quad (1.4-3)$$

信息传输速率又叫作传信率、比特率,定义为每秒传送的比特数,即信息量,或者是每秒传送的二进制码元个数。单位为 bit/s,简记为 bps。若码元符号的熵为 H ,则信息传输速率 R 为

$$R = R_s \times H \quad (1.4-4)$$

若 M 进制的码元符号等概率分布,其一个码元携带 $\log_2 M$ 比特的信息量,此时有

$$R = R_s \times \log_2 M \quad (1.4-5)$$

(2) 带宽效率

带宽效率定义为单位带宽内的传输速率,表示对频带的利用效率,单位为 Baud/Hz 或 bps/Hz。若通信带宽记为 W ,则带宽效率为

$$r = \frac{R_s}{W} \text{ (Baud/Hz)} \quad (1.4-6)$$

$$r_b = \frac{R}{W} \text{ (bps/Hz)} \quad (1.4-7)$$

(3) 功率效率

在通信系统中,功率效率(Power Efficiency)通常用于衡量系统如何有效地利用发射功率来可靠地传输信息。其常用单位比特能量(E_b)与噪声功率谱密度(N_0)的比值(E_b/N_0)来表示,单位为 dB,其值越低,说明系统在给定噪声环境下以更少的功率实现可靠传输。

$$\frac{E_b}{N_0} = \frac{P_{\text{avg}}}{RN_0} \quad (1.4-8)$$

其中, P_{avg} 为平均发射功率。

1.5 教材内容学习引导

本书适合作为信息与通信工程、电子科学与技术等学科一年级研究生课程的教材,也适



合从事数字通信系统设计的工程师作为自学课本和参考书。为了更好地阅读本书,读者应具备基本的微积分、线性系统理论以及概率和随机过程的预备知识等背景知识。

第1章是本书主题的导引,介绍了数字通信的基本含义与背景,包括数字通信的发展历程、数字通信系统组成、通信系统的性能指标以及全书的章节介绍。

第2章主要介绍后续章节所需的基础知识,包括傅里叶变换、确定信号分析、随机信号分析、带通信号与带通系统的等效低通表示以及信号空间与矢量表示方法。

第3章介绍了各种类型的数字通信调制方式,将数字序列映射成与信道特征相匹配的信号。本章将其分类为无记忆调制、有记忆调制、扩频调制和多载波调制等,分别介绍了每种数字调制类别中常见的调制方式。

第4章在加性高斯白噪声(AWGN)信道传输环境下,研究最佳接收机的理论设计和性能特征,并分析各类数字调制系统的可靠性。本章内容包括最佳检测理论、最佳接收机设计和典型调制方式下的差错性能分析。

第5章研究带宽受限情况下的信号设计。在介绍带限信道的一般特性基础上,重点讨论带限信道的信号设计、受控符号间干扰(ISD)的数据检测、存在ISI时AWGN信道对应的最佳接收机设计以及信道均衡问题。

第6章研究了在接收机中的载波和符号同步问题。主要介绍了极大似然载波相位估计和锁相环两种载波同步方法、极大似然符号定时估计和符号定时环路两种符号同步方法以及载波和符号联合同步方法。

第7章主要介绍信息论基础知识,并对一些重要的信道模型进行讨论,分析其信道容量。信息论研究通信系统有效性、可靠性等性能指标,为系统的优化提供理论指导,指出通信系统的极限。

第8章介绍了通信系统的链路预算。通过对系统中下行(或前向)和上行(或反向)信号传播途径中各种影响因素进行考察,对系统的链路余量进行估计,获得保持一定通信质量下链路所允许的最大传输速率。

第9章研究了无线信道衰落问题。围绕无线信道的传播特性展开分析,重点介绍无线信道的物理建模、多径衰落信道特性以及信号分集抗衰落技术。

参考文献

- [1] Proakis, J. G., & Salehi, M. Digital Communications (5th ed.) [M]. McGraw-Hill Science/Engineering/Math, 2007.
- [2] 樊昌信,曹丽娜. 通信原理(第7版)[M]. 北京:国防工业出版社,2012.
- [3] Joseph Boccuzzi 等. 通信信号处理[M]. 刘祖军,田斌,易克初译. 北京:电子工业出版社,2010.
- [4] 张杭,张邦宁等. 数字通信技术[M]. 北京:人民邮电出版社,2010.

第 2 章 数字通信的数学基础

2.1 引言

本章主要介绍后续章节所需的基础知识。首先在本章的引言部分介绍傅里叶变换和三角公式这两个非常有用的数学工具,接着在后续各节中分别对确定信号和随机信号进行分析,介绍它们的主要性质和统计特性等内容。然后研究信号的不同表示方法,给出带通信号与带通系统的等效低通表示,以及信号波形集的信号空间与矢量表示方法。

2.1.1 傅里叶变换

傅里叶变换是一种实用性很强的数学工具,在信号处理、图像处理、通信等多个领域都有广泛应用,主要用于将信号从时域(时间域)转换到频域(频率域),反之为傅里叶逆变换。傅里叶变换的核心思想是将一个信号分解为一系列不同频率的正弦波的叠加。通过傅里叶变换,可以分析信号在不同频率上的强度,从而更好地理解信号的特性。以下是傅里叶变换(从时域到频域)和傅里叶逆变换(从频域到时域)的基本定义:

$$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} dt \quad (2.1-1)$$

$$x(t) = \int_{-\infty}^{+\infty} X(f) e^{j2\pi ft} df \quad (2.1-2)$$

其中, $x(t)$ 为时域信号, $X(f)$ 为对应的频域信号。

表 2.1 总结了傅里叶变换的性质,表 2.2 列举了一些常用的傅里叶变换对。在这些表中使用到的信号定义如下。

矩形函数(门函数) $\Pi(t)$:

$$\Pi(t) = \begin{cases} 1 & |t| < \frac{1}{2} \\ \frac{1}{2} & t = \pm \frac{1}{2} \\ 0 & \text{其他} \end{cases} \quad (2.1-3)$$

辛格函数 $\text{sinc}(t)$:

$$\text{sinc}(t) = \begin{cases} \frac{\sin(\pi t)}{\pi t} & t \neq 0 \\ 1 & t = 0 \end{cases} \quad (2.1-4)$$



符号函数 $\text{sgn}(t)$:

$$\text{sgn}(t) = \begin{cases} 1 & t > 0 \\ -1 & t < 0 \\ 0 & t = 0 \end{cases} \quad (2.1-5)$$

三角形函数 $\Lambda(t)$:

$$\Lambda(t) = \Pi(t) * \Pi(t) = \begin{cases} t+1 & -1 \leq t < 0 \\ -t+1 & 0 \leq t < 1 \\ 0 & \text{其他} \end{cases} \quad (2.1-6)$$

其中, * 表示卷积。

单位阶跃信号 $u_{-1}(t)$:

$$u_{-1}(t) = \begin{cases} 1 & t > 0 \\ \frac{1}{2} & t = 0 \\ 0 & t < 0 \end{cases} \quad (2.1-7)$$

表 2.1 傅里叶变换的性质

性质	信号	傅里叶变换
线性	$\alpha x_1(t) + \beta x_2(t)$	$\alpha X_1(f) + \beta X_2(f)$
对偶性	$X(t)$	$x(-f)$
共轭性	$x^*(t)$	$X^*(-f)$
标度变换($a \neq 0$)	$x(at)$	$\frac{1}{ a } X\left(\frac{f}{a}\right)$
时移性	$x(t-t_0)$	$e^{-j2\pi f t_0} X(f)$
频移性	$e^{j2\pi f_0 t} x(t)$	$X(f-f_0)$
卷积	$x(t) * y(t)$	$X(f)Y(f)$
乘积	$x(t)y(t)$	$X(f) * Y(f)$
微分	$\frac{d^n}{dt^n} x(t)$	$(j2\pi f)^n X(f)$
频域微分	$t^n x(t)$	$\left(\frac{j}{2\pi}\right)^n \frac{d^n}{df^n} X(f)$
积分	$\int_{-\infty}^t x(\tau) d\tau$	$\frac{X(f)}{j2\pi f} + \frac{1}{2} X(0)\delta(f)$
帕塞瓦尔定理	$\int_{-\infty}^{\infty} x(t)y^*(t)dt = \int_{-\infty}^{\infty} X(f)Y^*(f)df$	
瑞利能量定理	$\int_{-\infty}^{\infty} x(t) ^2 dt = \int_{-\infty}^{\infty} X(f) ^2 df$	



表 2.2 常见的傅里叶变换对

时域	频域
$x(t) = \int_{-\infty}^{+\infty} X(f) e^{j2\pi ft} df$	$X(f) = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} dt$
$\delta(t)$	1
1	$\delta(f)$
$\delta(t-t_0)$	$e^{-j2\pi ft_0}$
$e^{j2\pi f_0 t}$	$\delta(f-f_0)$
$\cos(2\pi f_0 t)$	$\frac{1}{2}[\delta(f-f_0)+\delta(f+f_0)]$
$\sin(2\pi f_0 t)$	$\frac{1}{2j}[\delta(f-f_0)-\delta(f+f_0)]$
$\Pi(t)$	$\text{sinc}(f)$
$\text{sinc}(t)$	$\Pi(f)$
$\Lambda(t)$	$\text{sinc}^2(f)$
$\text{sinc}^2(t)$	$\Lambda(f)$
$e^{-\alpha t} u_{-1}(t), \alpha > 0$	$\frac{1}{\alpha + j2\pi f}$
$t e^{-\alpha t} u_{-1}(t), \alpha > 0$	$\frac{1}{(\alpha + j2\pi f)^2}$
$e^{-\alpha t }, \alpha > 0$	$\frac{2\alpha}{\alpha^2 + (2\pi f)^2}$
$e^{-\alpha t^2}$	$e^{-\pi f^2}$
$\text{sgn}(t)$	$\frac{1}{j\pi f}$
$u_{-1}(t)$	$\frac{1}{2}\delta(f) + \frac{1}{j2\pi f}$
$\frac{1}{2}\delta(t) + j\frac{1}{2\pi t}$	$u_{-1}(f)$
$\delta'(t)$	$j2\pi f$
$\delta^{(n)}(t)$	$(j2\pi f)^n$
$\frac{1}{t}$	$-j\pi \text{sgn}(f)$
$\sum_{n=-\infty}^{\infty} \delta(t-nT_0)$	$\frac{1}{T_0} \sum_{n=-\infty}^{\infty} \delta\left(f-\frac{n}{T_0}\right)$

2.1.2 三角公式

三角公式是三角函数的基本公式,用于描述角度和边长之间的关系,在通信系统的调制和解调数学表述时非常有用。下面给出一些常用的三角公式。



积化和差公式:

$$\begin{aligned}\cos A \cos B &= \frac{1}{2} [\cos (A+B) + \cos (A-B)] \\ \sin A \sin B &= -\frac{1}{2} [\cos (A+B) - \cos (A-B)] \\ \sin A \cos B &= \frac{1}{2} [\sin (A+B) + \sin (A-B)] \\ \cos A \sin B &= \frac{1}{2} [\sin (A+B) - \sin (A-B)]\end{aligned}\quad (2.1-8)$$

和差化积公式:

$$\begin{aligned}\cos A + \cos B &= 2 \cos \left(\frac{A+B}{2} \right) \cos \left(\frac{A-B}{2} \right) \\ \cos A - \cos B &= -2 \sin \left(\frac{A+B}{2} \right) \sin \left(\frac{A-B}{2} \right) \\ \sin A + \sin B &= 2 \sin \left(\frac{A+B}{2} \right) \cos \left(\frac{A-B}{2} \right) \\ \sin A - \sin B &= 2 \cos \left(\frac{A+B}{2} \right) \sin \left(\frac{A-B}{2} \right)\end{aligned}\quad (2.1-9)$$

倍角公式:

$$\begin{aligned}\cos (2A) &= 2 \cos^2 A - 1 \\ \sin (2A) &= 2 \sin A \cos A\end{aligned}\quad (2.1-10)$$

半角公式:

$$\begin{aligned}\cos^2 A &= \frac{1}{2} [1 + \cos (2A)] \\ \sin^2 A &= \frac{1}{2} [1 - \cos (2A)]\end{aligned}\quad (2.1-11)$$

2.2 确定信号分析

确定信号是指数学模型确定的信号,即在任意时刻,信号的值都是确定和可预知的,并且可以通过数学表达式或图形来精确描述。确定信号的特点是信号的值不随时间随机变化,而是遵循一定的规律。确定信号分析是信号处理领域中的一个重要分支,主要从确定信号的傅里叶级数、傅里叶变换、频谱密度、功率谱密度和相关函数等方面展开研究。这些分析有助于理解确定信号在时域和频域的特性,对通信领域具有重要意义。



2.2.1 确定信号的类型

1. 周期信号和非周期信号

按照是否具有周期性,确定信号可以分为周期信号和非周期信号。如果信号 $s(t)$ 的信号值在规律的时间间隔内重复出现,即满足

$$s(t) = s(t + T_0), \quad -\infty < t < \infty \quad (2.2-1)$$

其中, T_0 为一常数,且 $T_0 > 0$,则称信号 $s(t)$ 为周期信号,否则为非周期信号。将满足式(2.2-1)的最小 T_0 称为信号 $s(t)$ 的周期, $f_0 = \frac{1}{T_0}$ 称为基频。

2. 能量信号和功率信号

按照能量是否有限,确定信号可以分为能量信号和功率信号。如果信号 $s(t)$ 在整个时间轴上的能量是一个正的有限值,即

$$0 < E = \int_{-\infty}^{\infty} |s(t)|^2 dt < \infty \quad (2.2-2)$$

则称信号 $s(t)$ 为能量信号,能量信号通常在有限的时间内存在。定义信号的平均功率为

$$P = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |s(t)|^2 dt \quad (2.2-3)$$

则能量信号的平均功率为零,因为能量信号的平方的积分有限,在被趋于无穷大的时间 T 除后,得到的平均功率趋近于零。

在实际的通信系统中,信号都具有有限的功率、有限的持续时间,因而具有有限的能量。但是,如果信号的持续时间非常长,如广播信号,则可以近似认为它具有无限长的持续时间。此时,认为由式(2.2-3)定义的信号平均功率是一个有限的正值,但是其能量近似等于无穷大,称这样的信号 $s(t)$ 为功率信号。

2.2.2 确定信号的频域性质

在通信系统中,确定信号的频域性质包括其频率成分、幅度谱、相位谱、带宽和功率谱密度等,这些性质的研究对于理解信号如何在不同频率上分布能量和信息至关重要。通过分析这些频域特性,能够设计出高效的调制解调和滤波技术,优化信号传输,减少干扰,提高通信质量,并确保信息在复杂信道中的可靠传输。本节主要研究功率信号的频谱、能量信号的频谱密度、能量信号的能量谱密度和功率信号的功率谱密度。

1. 功率信号的频谱

设一个具有周期性的功率信号 $s(t)$ 的周期为 T_0 ,则定义其频谱函数为

$$C_n = C(nf_0) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} s(t) e^{-j2\pi n f_0 t} dt \quad (2.2-4)$$

其中, $f_0 = \frac{1}{T_0}$, n 为整数,且 $-\infty < n < +\infty$ 。

由傅里叶级数理论可知,式(2.2-4)中 C_n 就是周期性函数 $s(t)$ 展开成傅里叶级数的



系数, 即当周期信号 $s(t)$ 满足狄利克雷条件时, 其可以展开成如下的傅里叶级数

$$s(t) = \sum_{n=-\infty}^{\infty} C_n e^{j2\pi n t / T_0} \quad (2.2-5)$$

一般情况下, 式(2.2-4)中频谱函数 C_n 是一个复数, 代表在频率 nf_0 上信号分量的复振幅, 可以写作:

$$C_n = |C_n| e^{j\theta_n} \quad (2.2-6)$$

其中, $|C_n|$ 为频率 nf_0 上信号分量的振幅, θ_n 为频率 nf_0 上信号分量的相位。

式(2.2-6)表明周期性功率信号的频谱函数 C_n 是离散的, 只在 f_0 的整数倍上取值。由于 n 可以取负值, 所以在负频率上 C_n 也有值。通常称 C_n 为双边(频)谱。但是, 双边谱中的负频谱仅在数学上有意义, 在物理上, 并不存在负频率。

对于物理可实现的实信号 $s(t)$, 由式(2.2-4)有

$$C_{-n} = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} s(t) e^{+j2\pi n f_0 t} dt = \left[\frac{1}{T_0} \int_{-T_0/2}^{T_0/2} s(t) e^{-j2\pi n f_0 t} dt \right]^* = C_n^* \quad (2.2-7)$$

即频谱函数的正频率部分和负频率部分间存在复数共轭关系。也就是说, 负频谱和正频谱的模是偶对称的, 相位是奇对称的, 如图 2.1 所示。

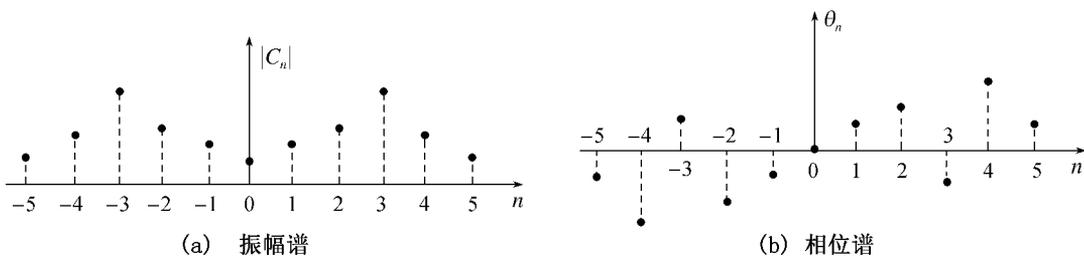


图 2.1 周期性信号的频谱

2. 能量信号的频谱密度

设一个能量信号为 $s(t)$, 将它的傅里叶变换 $S(f)$ 定义为它的频谱密度, 即

$$S(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi f t} dt \quad (2.2-8)$$

能量信号一般时长有限, 不具备周期性, 功率信号一般时长无限, 可以具有周期性。时域的周期性对应频域的离散性, 时域的非周期性对应频域连续性。能量信号的频谱密度 $S(f)$ 和周期性功率信号的频谱 C_n 主要有以下两点区别: 第一, $S(f)$ 是连续谱, C_n 是离散谱; 第二, $S(f)$ 的单位是伏/赫(V/Hz), 而 C_n 的单位是伏(V)。能量信号的能量有限, 并且分布在连续频率轴上, 所以在每个频率点 f 上信号的幅度是无穷小; 只有在一小段频率间隔 df 上才有确定的非零振幅。功率信号的功率有限, 但能量无限, 它在无限多的离散频率点上有确定的非零振幅。

对于实能量信号 $s(t)$ 有:

$$S(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi f t} dt = \left[\int_{-\infty}^{\infty} s(t) e^{+j2\pi f t} dt \right]^* = [S(-f)]^* \quad (2.2-9)$$



由式(2.2-9)可知,实能量信号的频谱密度和实功率信号的频谱有一个共同的特性,即其负频谱和正频谱的模偶对称,相位奇对称。也即实能量信号的频谱密度的正频率部分和负频率部分也成复数共轭关系。

3. 能量信号的能量谱密度

设一个能量信号为 $s(t)$, 它的能量 E 为

$$E = \int_{-\infty}^{\infty} |s(t)|^2 dt \quad (2.2-10)$$

若此信号的傅里叶变换(频谱密度)为 $S(f)$, 则由帕塞瓦尔(Parseval)定理可知

$$E = \int_{-\infty}^{\infty} |s(t)|^2 dt = \int_{-\infty}^{\infty} |S(f)|^2 df \quad (2.2-11)$$

式(2.2-11)表示 $|S(f)|^2$ 在频率轴 f 上的积分等于信号的能量, 因此, 称 $|S(f)|^2$ 为能量谱密度, 它表示在频率 f 处宽度为 df 的频带内的信号能量, 也可以看作是单位频带内的信号能量。

若将式(2.2-11)改写为

$$E = \int_{-\infty}^{\infty} G(f) df \quad (2.2-12)$$

则 $G(f) = |S(f)|^2$ (J/Hz) 为能量谱密度。若信号 $s(t)$ 是一个实函数, 则 $|S(f)|$ 是一个偶函数。因此, 对于实信号, 式(2.2-12)可改写为

$$E = 2 \int_0^{\infty} G(f) df \quad (2.2-13)$$

4. 功率信号的功率谱密度

由于功率信号具有无穷大的能量, 式(2.2-10)的积分不存在, 所以不能计算功率信号的能量谱密度。但是, 可以求出它的功率谱密度。首先, 将信号 $s(t)$ 截短为一个长度为 T 的截短信号 $s_T(t)$, $-T/2 < t < T/2$ 。此时, $s_T(t)$ 为一个能量信号。对于这个能量信号, 有

$$E = \int_{-T/2}^{T/2} |s_T(t)|^2 dt = \int_{-\infty}^{\infty} |S_T(f)|^2 df \quad (2.2-14)$$

当 $T \rightarrow \infty$ 时, 将

$$\lim_{T \rightarrow \infty} \frac{1}{T} |S_T(f)|^2 \quad (2.2-15)$$

定义为信号的功率谱密度 $P(f)$, 即

$$P(f) = \lim_{T \rightarrow \infty} \frac{1}{T} |S_T(f)|^2 \quad (2.2-16)$$

信号功率则为

$$P = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} |S_T(f)|^2 df = \int_{-\infty}^{\infty} P(f) df \quad (2.2-17)$$

若此功率信号具有周期性, 则可以将 T 选作等于信号的周期 T_0 , 并且用傅里叶级数代替傅里叶变换, 求出信号的频谱。这时, 式(2.2-3)信号的平均功率变成



$$P = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |s(t)|^2 dt = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} |s(t)|^2 dt \quad (2.2-18)$$

并且由周期函数的帕塞瓦尔定理得知

$$P = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} |s(t)|^2 dt = \sum_{n=-\infty}^{\infty} |C_n|^2 \quad (2.2-19)$$

其中, C_n 为此周期信号的傅里叶级数的系数。因此 $|C_n|^2$ 即周期功率信号的功率谱密度。

2.2.3 确定信号的时域性质

在本节中,我们主要讨论确定信号在时域中的自相关函数和互相关函数的定义和基本性质。

1. 能量信号的自相关函数

性质 2.1 能量信号的自相关函数 $R(\tau)$ 和能量谱密度 $|S(f)|^2$ 构成一对傅里叶变换。

证明 能量信号 $s(t)$ 的自相关函数描述同一个信号在不同时刻之间的相关程度, 定义为

$$R(\tau) = \int_{-\infty}^{\infty} s(t+\tau)s^*(t) dt, \quad -\infty < \tau < \infty \quad (2.2-20)$$

其中, τ 是时间差。这个定义表明自相关函数是信号的共轭与其自身在不同时间延迟下的乘积的积分。自相关函数 $R(\tau)$ 与时间 t 无关, 只与时间差 τ 有关。当 $\tau=0$ 时, 能量信号 $s(t)$ 的自相关函数 $R(0)$ 等于其能量, 即

$$R(0) = \int_{-\infty}^{\infty} |s(t)|^2 dt = E \quad (2.2-21)$$

其中, E 为能量信号的能量。

能量信号 $s(t)$ 的自相关函数 $R(\tau)$ 满足共轭对称性, 即

$$R(\tau) = R^*(-\tau) \quad (2.2-22)$$

对定义式(2.2-20)求傅里叶变换得

$$\begin{aligned} \int_{-\infty}^{\infty} R(\tau) e^{-j2\pi f\tau} d\tau &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(t+\tau)s^*(t) e^{-j2\pi f\tau} dt d\tau \\ &= \int_{-\infty}^{\infty} s^*(t) dt \left[\int_{-\infty}^{\infty} s(t+\tau) e^{-j2\pi f(t+\tau)} d(t+\tau) \right] e^{+j2\pi ft} \end{aligned} \quad (2.2-23)$$

令 $t'=t+\tau$, 并代入式(2.2-23), 得

$$\begin{aligned} \int_{-\infty}^{\infty} R(\tau) e^{-j2\pi f\tau} d\tau &= \int_{-\infty}^{\infty} s^*(t) dt \left[\int_{-\infty}^{\infty} s(t') e^{-j2\pi ft'} dt' \right] e^{+j2\pi ft} \\ &= S(f) \int_{-\infty}^{\infty} s^*(t) e^{+j2\pi ft} dt = S(f)S^*(f) \end{aligned} \quad (2.2-24)$$

其中, $S(f) = \int_{-\infty}^{\infty} s(t) e^{-j2\pi ft} dt$ 为能量信号 $s(t)$ 的频谱密度。一般情况下, $S(f)$ 是复函数, 令



$$S(f) = A(f) + jB(f) \quad (2.2-25)$$

其中, $A(f)$ 和 $B(f)$ 为实函数。因此式(2.2-24)变为

$$\begin{aligned} \int_{-\infty}^{\infty} R(\tau) e^{-j2\pi f\tau} d\tau &= S(f) S^*(f) = [A(f) + jB(f)][A(f) - jB(f)] \\ &= A^2(f) + B^2(f) = |S(f)|^2 \end{aligned} \quad (2.2-26)$$

将式(2.2-26)和式(2.2-11)进行比较可知, 能量信号的自相关函数的傅里叶变换就是其能量谱密度。反之, 能量信号的能量谱密度的逆傅里叶变换就是能量信号的自相关函数, 即

$$R(\tau) = \int_{-\infty}^{\infty} |S(f)|^2 e^{j2\pi f\tau} df \quad (2.2-27)$$

也就是说, 能量信号的自相关函数 $R(\tau)$ 和能量谱密度 $|S(f)|^2$ 构成一对傅里叶变换。

2. 功率信号的自相关函数

性质 2.2 周期性功率信号的自相关函数 $R(\tau)$ 和其功率谱密度 $P(f)$ 之间是傅里叶变换关系。

证明 功率信号 $s(t)$ 的自相关函数描述信号与其自身在不同时间点上的相似性或相关性, 定义为

$$R(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s(t+\tau) s^*(t) dt, \quad -\infty < \tau < \infty \quad (2.2-28)$$

当 $\tau=0$ 时, 功率信号的自相关函数 $R(0)$ 等于信号的平均功率, 即

$$R(0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |s(t)|^2 dt = P \quad (2.2-29)$$

其中, P 为信号的功率。

与能量信号的自相关函数类似, 功率信号的自相关函数也满足共轭对称性, 即

$$R(\tau) = R^*(-\tau) \quad (2.2-30)$$

对于周期性功率信号 $s(t)$, 其自相关函数的定义可以改写为

$$R(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} s(t+\tau) s^*(t) dt, \quad -\infty < \tau < \infty \quad (2.2-31)$$

周期性功率信号的自相关函数和其功率谱密度之间也有简单的关系, 即

$$R(\tau) = \int_{-\infty}^{\infty} P(f) e^{j2\pi f\tau} df \quad (2.2-32)$$

式(2.2-32)表明周期性功率信号的自相关函数 $R(\tau)$ 和其功率谱密度 $P(f)$ 之间是傅里叶变换关系, 即 $P(f)$ 的逆傅里叶变换是 $R(\tau)$, 而 $R(\tau)$ 的傅里叶变换是功率谱密度 $P(f)$, 即

$$P(f) = \int_{-\infty}^{\infty} R(\tau) e^{-j2\pi f\tau} d\tau \quad (2.2-33)$$

3. 能量信号的互相关函数

性质 2.3 能量信号的互相关函数和互能量谱密度也是一对傅里叶变换。



证明 能量信号的互相关函数用来衡量两个不同信号之间的相似性或一致性,假设两个不同的能量信号分别为 $s_1(t)$ 和 $s_2(t)$, 它们的互相关函数定义为

$$R_{12}(\tau) = \int_{-\infty}^{\infty} s_2(t+\tau)s_1^*(t)dt, \quad -\infty < \tau < \infty \quad (2.2-34)$$

由式(2.2-34)可知,互相关函数反映了一个信号和延迟 τ 后的另一个信号间相关的程度。互相关函数 $R_{12}(\tau)$ 和时间 t 无关,只和时间差 τ 有关。另外,能量信号的互相关函数通常与两个信号相乘的前后次序有关,即

$$R_{21}(\tau) = R_{12}^*(-\tau) \quad (2.2-35)$$

对于互相关函数和信号能量谱密度的关系,由定义式(2.2-34),有

$$\begin{aligned} R_{12}(\tau) &= \int_{-\infty}^{\infty} s_2(t+\tau)s_1^*(t)dt \\ &= \int_{-\infty}^{\infty} s_1^*(t)dt \int_{-\infty}^{\infty} S_2(f)e^{j2\pi f(t+\tau)}df \\ &= \int_{-\infty}^{\infty} S_2(f)df \int_{-\infty}^{\infty} s_1^*(t)e^{j2\pi f(t+\tau)}dt \\ &= \int_{-\infty}^{\infty} S_1^*(f)S_2(f)e^{j2\pi f\tau}df \\ &= \int_{-\infty}^{\infty} S_{12}(f)e^{j2\pi f\tau}df \end{aligned} \quad (2.2-36)$$

其中, $S_{12}(f) = S_1^*(f)S_2(f)$ 称为互能量谱密度。式(2.2-36)表明 $R_{12}(\tau)$ 是 $S_{12}(f)$ 的逆傅里叶变换。所以, $S_{12}(f)$ 是 $R_{12}(\tau)$ 的傅里叶变换,即

$$S_{12}(f) = \int_{-\infty}^{\infty} R_{12}(\tau)e^{-j2\pi f\tau}d\tau \quad (2.2-37)$$

也就是说,互相关函数和互能量谱密度也是一对傅里叶变换。

4. 功率信号的互相关函数

性质 2.4 周期性功率信号的互功率谱 C_{12} 是其互相关函数 $R_{12}(\tau)$ 的傅里叶级数的系数。

证明 功率信号的互相关函数衡量两个功率信号之间的相似性,两个不同的功率信号 $s_1(t)$ 和 $s_2(t)$ 的互相关函数定义为

$$R_{12}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s_2(t+\tau)s_1^*(t)dt, \quad -\infty < \tau < \infty \quad (2.2-38)$$

同样,功率信号的互相关函数 $R_{12}(\tau)$ 也和时间 t 无关,只和时间差 τ 有关,并且互相关函数与两个信号相乘的前后次序有关,即

$$R_{21}(\tau) = R_{12}^*(-\tau) \quad (2.2-39)$$

若两个周期性功率信号的周期相同,则其互相关函数的定义可以写为

$$R_{12}(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} s_2(t+\tau)s_1^*(t)dt, \quad -\infty < \tau < \infty \quad (2.2-40)$$

其中, T 为相同的信号周期。



功率信号的互相关函数和其功率谱之间也有简单的傅里叶变换关系,即

$$R_{12}(\tau) = \sum_{n=-\infty}^{\infty} [C_{12}] e^{j2\pi n f_0 \tau} \quad (2.2-41)$$

其中, $C_{12} = (C_n)_1^* (C_n)_2$ 称为信号的互功率谱。式(2.2-41)表明周期性功率信号的互功率谱 C_{12} 是其互相关函数 $R_{12}(\tau)$ 的傅里叶级数的系数。

2.3 随机信号分析

随机信号分析也是信号处理领域的一个重要分支,它主要研究信号中的随机成分及其统计特性。随机信号,也称为随机过程,是在时间上或空间上变化的随机变量的集合。这类信号在自然界和工程实践中普遍存在,如噪声信号、生物电信号等。通信系统中发送的信号对于接收方来说具备一定的不可预知性,或者说随机性,这样信号才具有传输的价值。另外,接收系统中的干扰与噪声、信道特性的起伏,也是随机变化的。通信系统中的信号与噪声都具有一定的随机性,需要用随机过程来描述,因此,随机信号分析在通信系统领域具有重要的作用。

本节在介绍随机过程的分布和数字特征等基本概念的基础上,重点讨论通信系统中常见的平稳随机过程的统计特性,并介绍宽平稳随机过程的功率谱密度,提供一种在频率域上分析随机过程的方法。此外,着重研究随机过程及其统计特性通过线性时不变系统后的情况,并对具有离散时间、离散值的随机过程——马尔可夫链(Markov Chain)进行简要分析。

2.3.1 随机过程的基本概念

随机过程是一类随时间作随机变化的过程,它不能用确定的时间函数描述。一方面,随机过程可以看成是不同随机试验结果的时间函数的集合。而试验结果的每一个记录,都是一个确定的时间函数 $x_i(t)$,我们把它称作样本函数(Sample Function)或随机过程的一次实现(Realization)。全部样本函数构成的总体 $\{x_1(t), x_2(t), \dots, x_n(t)\}$ 就是一个随机过程。

另一方面,随机过程可以看作是在时间进程中处于不同时刻的随机变量的集合。在任一给定时刻 t_1 上,每一个样本函数 $x_i(t)$ 都是一个确定的数值 $x_i(t_1)$,但是每个 $x_i(t_1)$ 都是不可预知的,这正是随机过程随机性的体现。所以,在一个固定时刻 t_1 上,不同样本的取值是一个随机变量。换言之,随机过程在任意时刻的值是一个随机变量。将随机过程看作是不同时刻随机变量的集合更方便我们对随机过程理论进行精确的数学描述。

2.3.2 随机过程的分布函数与概率密度

随机过程 $X(t)$ 的统计特性可以用累计分布函数(Cumulative Distribution Function, CDF)和概率密度函数(Probability Density Function, PDF)来描述。

1. 累计分布函数

我们用概率测度 $P(\cdot)$ 表示一个事件发生的可能性大小,将随机变量 X 小于或等于某



一数值 x 的概率称为随机过程的累计分布函数 $F(x)$, 记作

$$F(x) = P\{X \leq x\} \quad (2.3-1)$$

一个随机过程的一维累计分布函数可以记作

$$F_X(x, t) = P\{X(t) \leq x\} \quad (2.3-2)$$

但是, 一维累计分布函数对随机过程的描述非常不充分, 它仅仅能描述随机过程在任一瞬间的统计特性。进而, 对于任意固定的时刻 t_1 和 t_2 , 我们把 $X(t_1) \leq x_1$ 和 $X(t_2) \leq x_2$ 同时出现的概率称为该随机过程的二维累计分布函数, 即

$$F_X(x_1, x_2; t_1, t_2) = P\{X(t_1) \leq x_1 \text{ 且 } X(t_2) \leq x_2\} \quad (2.3-3)$$

同理, 对于任意给定的 $t_1, t_2, \dots, t_n \in T$, 随机过程 $\{X(t), t \in T\}$ 的 n 维累计分布函数为

$$F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = P\{X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_n) \leq x_n\} \quad (2.3-4)$$

显然, 当 n 取值越大, 对随机过程统计特性的描述将越充分。

2. 概率密度函数

如果一个随机变量的累计分布函数 $F(x)$ 对 x 的导数 $f(x)$ 存在, 我们就称 $f(x)$ 为该随机变量的概率密度函数, 即有

$$f(x) = \frac{dF(x)}{dx} \quad (2.3-5)$$

类似的, 随机过程 $\{X(t), t \in T\}$ 的一维概率密度函数记作

$$f_X(x, t) = \frac{\partial}{\partial x} F_X(x, t) \quad (2.3-6)$$

则对于任意固定的时刻 t_1 和 t_2 , 其二维概率密度函数为

$$f_X(x_1, x_2; t_1, t_2) = \frac{\partial^2}{\partial x_1 \cdot \partial x_2} F_X(x_1, x_2; t_1, t_2) \quad (2.3-7)$$

同理, 对于任意给定的 $t_1, t_2, \dots, t_n \in T$, 其 n 维概率密度函数为

$$f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = \frac{\partial^n}{\partial x_1 \cdot \partial x_2 \cdot \dots \cdot \partial x_n} F_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \quad (2.3-8)$$

2.3.3 随机过程的数字特征

随机过程的累计分布函数和概率密度函数是随机过程统计特性的完整描述, 但在实际中往往很难或不需要确定随机过程的分布函数或概率密度, 这时可以用随机过程的数字特征来部分地描述和分析随机过程的主要统计特性, 从而更方便地理解和分析随机过程的行为, 为信号处理、通信、控制等领域提供重要的统计信息。随机过程的数字特征由随机变量的数字特征推广得到, 常用的数字特征有均值、方差、相关函数、协方差和相关系数等。



1. 均值(数学期望)

随机过程 $X(t)$ 的均值也称为数学期望, 表征随机过程在 t 时刻取值的平均特征, 定义为

$$\mu(t) = E[X(t)] = \int_{-\infty}^{\infty} x f_X(x, t) dx \quad (2.3-9)$$

由式(2.3-9)可知, 随机过程 $X(t)$ 的均值 $E[X(t)]$ 是时间的确定函数, 记作 $\mu(t)$, 表示随机过程的 n 个样本函数曲线 $\{x_1(t), x_2(t), \dots, x_n(t)\}$ 的摆动中心。

2. 方差

随机过程 $X(t)$ 的方差反映随机过程在时刻 t 相对于其均值的偏离程度或分散程度, 常记作 $\sigma^2(t)$, 定义为

$$\sigma^2(t) = \text{Var}[X(t)] = E\{[X(t) - \mu(t)]^2\} \quad (2.3-10)$$

3. 自相关函数和自协方差函数

随机过程的均值和方差仅与一维概率密度函数有关, 只能描述随机过程在各个孤立时刻的特征, 而不能反映随机过程内在的联系。为了揭示随机过程在不同时刻的相互依赖关系, 引入自相关函数和自协方差函数两个数字特征。

随机过程 $X(t)$ 的自相关函数定义为

$$R(t_1, t_2) = E[X(t_1)X(t_2)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 \cdot x_2 \cdot f_X(x_1, x_2; t_1, t_2) dx_1 dx_2 \quad (2.3-11)$$

其中, $X(t_1)$ 和 $X(t_2)$ 分别是在 t_1 和 t_2 时刻观测 $X(t)$ 得到的随机变量, $f_X(x_1, x_2; t_1, t_2)$ 为 $X(t)$ 的二维概率密度函数。因此, 有如下关系式

$$E[(X(t_1) + X(t_2))^2] = R(t_1, t_1) + 2R(t_1, t_2) + R(t_2, t_2) \quad (2.3-12)$$

随机过程 $X(t)$ 的自协方差函数定义为

$$\begin{aligned} C(t_1, t_2) &= \text{Cov}(X(t_1), X(t_2)) \\ &= E[(X(t_1) - \mu(t_1)) \cdot (X(t_2) - \mu(t_2))] \end{aligned} \quad (2.3-13)$$

其中, $\mu(t_1)$ 和 $\mu(t_2)$ 分别是在 t_1 和 t_2 时刻得到的 $X(t)$ 的均值。

这里介绍两种特殊的随机过程:

(1) α 依赖过程: 随机过程 $X(t)$ 为 α 依赖过程, 其统计特性随着时间间隔的增加而逐渐变得独立, 即当 $|t_1 - t_2| > \alpha$ 时, $C(t_1, t_2) = 0$;

(2) 白噪声: 随机过程 $X(t)$ 为白噪声, 其自协方差函数 $C(t_1, t_2) = 0 (t_1 \neq t_2)$, 这意味着白噪声在不同时刻的值之间没有相关性。

随机过程 $X(t)$ 的相关系数定义为

$$\rho(t_1, t_2) = \frac{C(t_1, t_2)}{\sqrt{C(t_1, t_1) \cdot C(t_2, t_2)}} \quad (2.3-14)$$

易知, 随机过程 $X(t)$ 的数字特征之间有如下关系:



$$C(t_1, t_2) = R(t_1, t_2) - \mu(t_1) \cdot \mu(t_2) \quad (2.3-15)$$

$$R(t, t) = \sigma^2(t) + [\mu(t)]^2 \quad (2.3-16)$$

$$C(t, t) = \sigma^2(t) \quad (2.3-17)$$

4. 互相关函数和互协方差函数

为了描述两个不同随机过程之间的相关性,我们把自相关函数和自协方差函数的概念引伸到两个随机过程中,得到互相关函数和互协方差函数。设 $X(t)$ 和 $Y(t)$ 分别表示两个不同的随机过程,则互相关函数定义为

$$R_{XY}(t_1, t_2) = E[X(t_1) \cdot Y(t_2)] \quad (2.3-18)$$

如果两个随机过程 $X(t)$ 和 $Y(t)$ 满足

$$R_{XY}(t_1, t_2) = 0 \quad (2.3-19)$$

那么,称这两个随机过程是正交的。

互协方差函数定义为

$$\begin{aligned} C_{XY}(t_1, t_2) &= \text{Cov}(X(t_1), Y(t_2)) \\ &= E[(X(t_1) - \mu_X(t_1)) \cdot (Y(t_2) - \mu_Y(t_2))] \end{aligned} \quad (2.3-20)$$

$C_{XY}(t_1, t_2)$ 与 $R_{XY}(t_1, t_2)$ 之间有如下关系:

$$C_{XY}(t_1, t_2) = R_{XY}(t_1, t_2) - \mu_X(t_1) \cdot \mu_Y(t_2) \quad (2.3-21)$$

如果两个随机过程 $X(t)$ 和 $Y(t)$ 满足

$$C_{XY}(t_1, t_2) = 0 \quad (2.3-22)$$

那么,称这两个随机过程是不相关的。

2.3.4 平稳随机过程

平稳随机过程是一类应用非常广泛的随机过程,通信系统中的信号与噪声大多可视为平稳随机过程。因此,平稳随机过程的研究具有重要的理论与实际意义。平稳随机过程主要分为严平稳(Strict Sense Stationary, SSS)和宽平稳(Wide Sense Stationary, WSS)两种类型。

1. 严平稳随机过程

若随机过程 $X(t)$ 的统计特性不随时间起点的改变而变化,即时间平移不改变其任何统计特性,则称该随机过程是在严格意义下的平稳随机过程,简称严平稳随机过程。严平稳随机过程的任意有限维概率密度函数与时间起点无关,即不随自变量的公共时间平移而改变,则对于任意正整数 n 和所有实数 h ,有

$$f_X(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) = f_X(x_1, x_2, \dots, x_n; t_1 + h, t_2 + h, \dots, t_n + h) \quad (2.3-23)$$

因此,严平稳随机过程的一维概率密度函数与时间 t 无关,即

$$f_X(x, t) = f_X(x) \quad (2.3-24)$$



而严平稳随机过程的二维概率密度函数只与时间差 $\tau = t_2 - t_1$ 有关,即

$$f_X(x_1, x_2; t_1, t_2) = f_X(x_1, x_2; \tau) \quad (2.3-25)$$

2. 宽平稳随机过程

若随机过程 $X(t)$ 仅满足一阶和二阶统计平稳,即均值 $\mu(t)$ 为常数且自相关函数 $R(t, t+\tau)$ 仅与时间差 τ 有关,即同时满足式(2.3-26)和式(2.3-27),则称该随机过程为宽平稳随机过程。

$$\mu(t) = E[X(t)] = \mu \quad (2.3-26)$$

$$R(t, t+\tau) = E[X(t)X(t+\tau)] = R(\tau) \quad (2.3-27)$$

宽平稳随机过程的自协方差函数和相关系数分别为

$$C(\tau) = R(\tau) - |\mu|^2 \quad (2.3-28)$$

$$\rho(\tau) = \frac{C(\tau)}{C(0)} \quad (2.3-29)$$

白噪声过程和 α 依赖过程是宽平稳随机过程中的重要特例,它们各自的自相关或自协方差函数结构具有鲜明的特点,常被用来建模或分析通信系统中的随机现象。对于白噪声这种理想化的随机信号,其在不同时间点的值是完全独立的,因此它的自协方差函数

$$C(\tau) = q\delta(\tau) \quad (2.3-30)$$

其中, q 是白噪声的方差。式(2.3-30)表明白噪声的自协方差函数只在 $\tau=0$ 时不为 0,因此,白噪声的自协方差函数被近似为一个非常窄的脉冲,其宽度接近于零,高度等于白噪声的方差。

而 α 依赖过程的自协方差函数在 $|\tau| \geq \alpha$ 时为 0,即

$$C(\tau) = 0, \quad |\tau| \geq \alpha \quad (2.3-31)$$

其中, α 被称为关联时间或记忆时间,表示随机过程保持相关性的最大时间长度。当时间延迟 τ 超过这个关联时间时,随机过程在不同时刻的值之间的相关性消失,即自协方差为 0。

2.3.5 功率谱密度

宽平稳随机过程 $X(t)$ 的功率谱密度 $S(\omega)$ 描述了随机过程在频率域上的能量分布,其中 $\omega = 2\pi f$ 是角频率。功率谱密度提供了一种在频率域上分析随机过程的方法,有助于理解随机过程的频域特性,其单位是 W/Hz。维纳-辛钦定理阐明宽平稳随机过程的功率谱密度 $S(\omega)$ 和自相关函数 $R(\tau)$ 是一组傅里叶变换对,证明过程见 2.2.3 小节,即

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-j\omega\tau} d\tau \quad (2.3-32)$$

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{j\omega\tau} d\omega \quad (2.3-33)$$

对于实平稳过程,有

$$S(\omega) = 2 \int_0^{\infty} R(\tau) \cos(\omega\tau) d\tau \quad (2.3-34)$$



$$R(\tau) = \frac{1}{\pi} \int_0^{\infty} S(\omega) \cos(\omega\tau) d\omega \quad (2.3-35)$$

当一个功率谱密度为 $S_{xx}(\omega)$ 的宽平稳随机过程 $X(t)$ 通过一个冲激响应为 $h(t)$, 传输函数为 $H(\omega) = F[h(t)]$ 的线性时不变系统后 ($F[\cdot]$ 表示傅里叶变换), 输出为随机过程 $Y(t)$, 其功率谱密度为 $S_{yy}(\omega)$, 则有如下关系式:

$$S_{yy}(\omega) = H(\omega) S_{xx}(\omega) H^*(\omega) = |H(\omega)|^2 S_{xx}(\omega) \quad (2.3-36)$$

即存在一个重要的结论: 输出过程的功率谱密度是输入过程的功率谱密度乘以系统频率响应模值的平方, 则输出过程的功率 $P[Y(t)]$ 可由下式计算得到:

$$P[Y(t)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{xx}(\omega) |H(\omega)|^2 d\omega \quad (2.3-37)$$

下面给出两个实例:

(1) 积分系统

$$y(t) = \frac{1}{2T} \int_{t-T}^{t+T} x(\alpha) d\alpha \quad (2.3-38)$$

单位冲激响应

$$H(\omega) = \frac{\sin(T\omega)}{T\omega} \quad (2.3-39)$$

功率谱密度

$$S_{yy}(\omega) = S_{xx}(\omega) \frac{\sin^2(T\omega)}{T^2\omega^2} \quad (2.3-40)$$

自相关函数

$$R_{yy}(\tau) = \frac{1}{2T} \int_{-2T}^{2T} \left(1 - \frac{|\alpha|}{2T}\right) R_{xx}(\tau - \alpha) d\alpha \quad (2.3-41)$$

(2) 微分系统

$$y(t) = \frac{d}{dt} x(t) \quad (2.3-42)$$

单位冲激响应

$$H(\omega) = i\omega \quad (2.3-43)$$

功率谱密度

$$S_{yy}(\omega) = \omega^2 S_{xx}(\omega) \quad (2.3-44)$$

2.3.6 系统的随机分析

通信过程主要是信号通过系统传输的过程。研究随机过程及其统计特性通过系统后的情况有助于了解系统在随机因素影响下的行为和性能, 从而为通信系统设计和优化提供重要依据。



1. 系统的分类

系统的分类错综复杂,可以分为线性系统和非线性系统、时不变系统和时变系统、因果系统和非因果系统等。在本节中,我们主要研究线性系统、时不变系统和因果系统等。

一般来说,由线性元件(电阻、电感、电容)组成的系统称为线性系统。线性系统具有线性特性,即同时具有叠加性和均匀性(或称为齐次性)。叠加性是指当几个激励信号同时作用于系统时,总的响应信号等于每个激励单独作用所产生的响应之和。而均匀性(或称为齐次性)是指当输入信号乘以某常数 a 时,响应也乘以相同的常数。线性系统的性质可以通过下面的式子进行简单表述:

$$y(t) = \text{System}\{x(t)\} \Rightarrow ay(t) + b = \text{System}\{ax(t) + b\}, \quad \forall a, b \quad (2.3-45)$$

如果系统的参数不随时间而变化,则称这样的系统为时不变系统(或称为非时变系统)。对于时不变系统,由于系统参数本身不随时间变化,因此,在起始条件为零的条件下,系统的响应与激励施加于系统的时刻无关。时不变系统的时不变特性可以简单表述为:

$$y(t) = \text{System}\{x(t)\} \Rightarrow y(t - \tau) = \text{System}\{x(t - \tau)\}, \quad \forall \tau \quad (2.3-46)$$

如果 $t < t_0$ 时,系统的激励信号等于零,相应的响应信号在 $t < t_0$ 时也等于零,则这样的系统称为因果系统,否则即为非因果系统。因果系统的输出只依赖于当前及过去的输入和过去的输出,而不依赖于未来的输入和未来的输出。因此,因果系统符合实际物理系统的特性,是实际可实现的系统。而非因果系统虽然在理论上可以定义和分析,但由于物理不可行性,通常无法实际实现。因果系统的因果性可以简单表述为:

$$y(t_0) = \text{System}\{x(t)\}, \quad t \leq t_0 \quad (2.3-47)$$

2. 线性时不变系统的随机分析

同时具有线性性质与时不变性质的系统称为线性时不变系统。线性时不变系统可由其单位冲激响应 $h(t)$ 表征。设 $x(t)$ 为输入信号, $y(t)$ 为相应的输出信号,则输入、输出与单位冲激响应之间的关系可以表示为卷积,即

$$y(t) = \text{System}\{x(t)\} \Rightarrow y(t) = h(t) * x(t) \quad (2.3-48)$$

设随机过程 $X(t)$ 通过线性时不变系统后的输出随机过程为 $Y(t)$, 线性时不变系统的单位冲激响应为 $h(t)$, 则输入与输出随机过程也应满足式(2.3-48), 即

$$Y(t) = \text{System}\{X(t)\} \Rightarrow Y(t) = h(t) * X(t) \quad (2.3-49)$$

因此,随机过程的均值和自相关函数通过线性时不变系统后为

$$E[Y(t)] = E[h(t) * X(t)] = h(t) * E[X(t)] \quad (2.3-50)$$

$$R_{yy}(t_1, t_2) = h(t_1) * R_{xx}(t_1, t_2) * h^*(t_2) \quad (2.3-51)$$

下面给出两个实例:

(1) 积分系统

$$y(t) = \int_0^t x(\alpha) d\alpha \quad (2.3-52)$$

单位冲激响应



$$h(t) = \int_0^t \delta(t) dt = u_{-1}(t) \quad (2.3-53)$$

输出均值

$$E[y(t)] = \int_0^t E[x(t)] dt \quad (2.3-54)$$

输出自相关函数

$$R_{yy}(t_1, t_2) = \int_0^\infty \int_0^\infty R_{xx}(t_1 - \alpha, t_2 - \beta) d\alpha d\beta \quad (2.3-55)$$

(2) 微分系统

$$y(t) = \frac{d}{dt} x(t) \quad (2.3-56)$$

单位冲激响应

$$h(t) = \frac{d}{dt} \delta(t) \quad (2.3-57)$$

输出均值

$$E[y(t)] = \frac{d}{dt} E[x(t)] \quad (2.3-58)$$

输出自相关函数

$$R_{yy}(t_1, t_2) = \frac{\partial^2}{\partial t_1 \partial t_2} R_{xx}(t_1, t_2) \quad (2.3-59)$$

2.3.7 马尔可夫链

马尔可夫链是指具有离散时间、离散值的随机过程,其下一状态的概率分布只由当前状态决定,不受过去状态的影响。在 j 阶马尔可夫链中,当前值取决于最近的 j 个值,即

$$\begin{aligned} P[X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots] \\ = P[X_n = x_n | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_{n-j} = x_{n-j}] \end{aligned} \quad (2.3-60)$$

为方便起见,可将最近的 j 个值集合看成马尔可夫链的状态,则马尔可夫链的当前状态 $S_n = (X_n, X_{n-1}, \dots, X_{n-j+1})$ 仅取决于最近的状态 $S_{n-1} = (X_{n-1}, X_{n-2}, \dots, X_{n-j})$, 即

$$P[S_n = s_n | S_{n-1} = s_{n-1}, S_{n-2} = s_{n-2}, \dots] = P[S_n = s_n | S_{n-1} = s_{n-1}] \quad (2.3-61)$$

随机过程的内部演变取决于状态集和状态之间转移的概率统计规律,如果 $P[S_n | S_{n-1}]$ 独立于时间 n , 则称马尔可夫链是齐次的。在齐次马尔可夫链中,从状态 i 到状态 j 的转移概率与 n 无关, 则定义状态转移矩阵或一步转移矩阵 \mathbf{P} , 其中, \mathbf{P} 的第 i 行第 j 列元素 P_{ij} 表示从状态 i 直接转移到状态 j 的概率, 即

$$P_{ij} = P[S_n = j | S_{n-1} = i] \quad (2.3-62)$$

一步转移矩阵 \mathbf{P} 中的每个元素均具有非负性,且每一行元素和都为 1。 n 步转移矩阵表示由状态 i 经过 n 步移动到状态 j 的概率。对于离散时间齐次马尔可夫链, n 步转移矩阵等于 \mathbf{P}^n 。另外,对于 $m+n$ 步转移矩阵,有这样的结论:过程从状态 i 出发,经时间间隔 $m+n$ 转移到状态 j 的概率应等于过程从状态 i 出发先经 m 时段转移到任意的中间状态 k ,接着再从 k 出发经 n 时段转移到状态 j 的概率的和。这就是切普曼-柯尔莫戈洛夫方程,简



称 C-K 方程,其数学表述为

$$P_{ij}(m+n) = \sum_{k \in S} P_{ik}(m)P_{kj}(n), \quad i, j \in S \quad (2.3-63)$$

其中, S 表示状态空间,且 $S = \{1, 2, 3, \dots\}$ 。

2.4 带通信号与带通系统的表示

数字通信系统关注数字符号映射的波形信号在通信信道中的传输。通常情况下,发送信息序列的频谱特性与通信信道的频谱特性并不完全匹配,这意味着信息序列不能直接通过信道传输。由于大部分信息信号是低频(基带)信号,而通信信道可用的频谱通常位于较高频段,因此在发送端需要将信息信号转换为高频信号,以匹配通信信道的特性。这一过程称为调制,即将基带信号转换为带通已调信号。本节将详细探讨基带信号与带通信号的主要特点与性质,给出将实窄带高频的带通信号表示为等效低通的复低频信号的方法,同时简单介绍带通信号的功率谱密度、带通系统和带通平稳随机过程的等效低通表示。通过这种方法,可以将高频带通信号的相关问题转化为等效的低通信号来处理,大大简化信号处理的算法,并降低采样速率,从而降低系统的复杂性和成本。

2.4.1 带通信号与低通信号

信号的傅里叶变换可以提供信号的频率域信息。实信号 $x(t)$ 的傅里叶变换具有厄米特(Hermitian)对称性,即 $X(f) = X^*(-f)$,于是对于幅度谱有 $|X(-f)| = |X(f)|$,对于相位谱有 $\angle X(-f) = -\angle X(f)$ 。也就是说,实信号 $x(t)$ 的幅度谱是偶对称的,相位谱是奇对称的。由该对称性可知,实信号的全部信息都包含在正(或负)频域中,因此,只由 $X(f) (f \geq 0)$ 就可以完整地重构 $x(t)$ 。基于此,实信号 $x(t)$ 的带宽定义为最小的正频率范围,即其频率支撑集的一半。频率支撑集表示频谱的非零部分。

定义一个信号的正频谱 $X_+(f)$ 和负频谱 $X_-(f)$ 为

$$X_+(f) = \begin{cases} X(f) & f > 0 \\ \frac{1}{2}X(0) & f = 0 \\ 0 & f < 0 \end{cases} \quad (2.4-1)$$

$$X_-(f) = \begin{cases} X(f) & f < 0 \\ \frac{1}{2}X(0) & f = 0 \\ 0 & f > 0 \end{cases} \quad (2.4-2)$$

则

$$X_-(f) = X_+^*(-f) \quad (2.4-3)$$

$$X(f) = X_+(f) + X_-(f) = X_+(f) + X_+^*(-f) \quad (2.4-4)$$

式(2.4-4)表明 $X_+(f)$ 的信息对重构 $X(f)$ 是充分的。

复信号 $x(t)$ 的频谱 $X(f)$ 是不对称的, 因此, 复信号不能仅由正频率的信息来重构。复信号的带宽定义为频谱为非零值的全部频率范围的一半, 即该信号频率支撑集的一半。该定义与实信号的带宽定义是一致的。一般可以认为, 对于所有信号(实或复的), 带宽定义为频率支撑集的一半。

低通信号(或基带信号)在通信系统中非常常见, 通常是指频谱集中在较低频率范围内的信号, 如语音信号、图像信号等。低通信号通常是低频信号, 在时间域内缓慢变化, 没有跳变或突变。实低通信号的频率支持集为 $[-W, +W]$, 则其带宽为最小的正频率 W 。图 2.2 为一实低通(基带)信号的频谱图, 其中实线表示幅度谱 $|X(f)|$, 虚线表示相位谱 $\angle X(f)$ 。

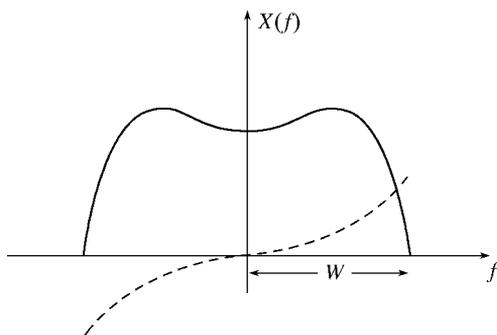


图 2.2 实低通(基带)信号的频谱图

在实际应用中, 由于通信信道的特性, 低通信号通常需要通过调制过程转换为带通信号, 以便在较高的频段上传输, 从而更好地匹配信道的频谱特性。

带通信号是一种带宽远小于载波频率、非零频谱集中在载波频率周围的窄带内的实信号。带通信号通常是高频信号, 在时间域内快速变化。图 2.3 为一实带通信号的频谱图, 其中实线表示幅度谱 $|X(f)|$, 是偶对称的, 虚线表示相位谱 $\angle X(f)$, 是奇对称的。其中心频率 f_0 不必位于带通信号频段的中间。

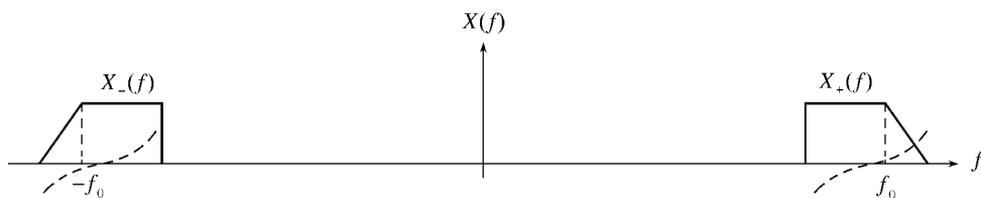


图 2.3 实带通信号的频谱图

2.4.2 带通信号的等效低通表示

1. 带通信号可以等效低通表示

本节将证明, 实窄带高频信号(带通信号)可以用复低频信号(低通信号)来表示, 于是可以把直接处理带通信号替代为处理等效低通信号, 从而大大简化带通信号的处理。因为对低通信号的处理算法要容易得多, 只需要较低的采样速率及相应的采样数据。

首先定义对应实带通信号 $x(t)$ 的解析信号(或预包络)为信号 $x_+(t)$, 其傅里叶变换为



$X_+(f)$ 。该信号仅包含正频率分量,其频谱不是厄米特对称性的,所以 $x_+(t)$ 一般是复信号,有

$$X_+(f) = X(f)u_{-1}(f) \quad (2.4-5)$$

$$\begin{aligned} x_+(t) &= F^{-1}(X_+(f)) \\ &= F^{-1}[X(f)u_{-1}(f)] \\ &= F^{-1}[X(f)] * F^{-1}[u_{-1}(f)] \\ &= x(t) * \left[\frac{1}{2}\delta(t) + j\frac{1}{2\pi t} \right] \\ &= \frac{1}{2}x(t) + \frac{j}{2}\hat{x}(t) \end{aligned} \quad (2.4-6)$$

其中, $u_{-1}(f)$ 是阶跃信号, $\hat{x}(t) = \frac{1}{\pi t} * x(t)$ 是 $x(t)$ 的希尔伯特(Hilbert)变换。

希尔伯特变换是一个对函数产生定义域相同的函数的线性算子,它常用于将一个实值信号转换为一个与之相关的复值信号。具体来说,信号 $x(t)$ 的希尔伯特变换等于该信号通过具有冲激响应 $h(t) = \frac{1}{\pi t}$ 的线性系统后的输出响应 $\hat{x}(t)$, 即

$$\hat{x}(t) = h(t) * x(t), \quad h(t) = \frac{1}{\pi t} \quad (2.4-7)$$

冲激响应 $h(t) = \frac{1}{\pi t}$ 的傅里叶变换,即系统的传输函数为

$$H(f) = \begin{cases} -j = e^{-j\frac{\pi}{2}} & f > 0 \\ 0 & f = 0 \\ j = e^{j\frac{\pi}{2}} & f < 0 \end{cases} \quad (2.4-8)$$

在 $x(t)$ 的正频率分量中引入 $-\frac{\pi}{2}$ 相移,而负频率分量中引入 $\frac{\pi}{2}$ 相移,则可得到 $x(t)$ 的希尔伯特变换。在频率域,有

$$F(\hat{x}(t)) = -j\text{sgn}(f)X(f) \quad (2.4-9)$$

定义实带通信号 $x(t)$ 的等效低通(或复包络) $x_l(t)$ 是由频谱 $2X_+(f+f_0)$ 所确定的信号,即 $x_l(t)$ 的傅里叶变换 $X_l(f) = 2X_+(f+f_0)$, 其中 f_0 是带通信号的中心频率。 $x_l(t)$ 的谱位于零频附近,所以 $x_l(t)$ 一般为复低通信号。图 2.4 给出了图 2.3 中信号的等效低通的频谱。

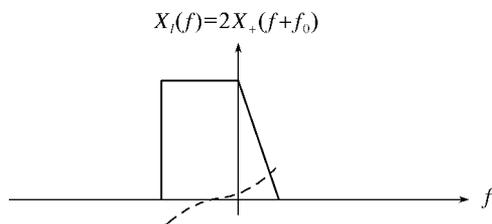


图 2.4 实带通信号的等效低通的频谱



因此,

$$x_l(t) = F^{-1}[X_l(f)] = 2x_+(t)e^{-j2\pi f_0 t} = (x(t) + j\hat{x}(t))e^{-j2\pi f_0 t} \quad (2.4-10)$$

则

$$x(t) = \operatorname{Re}[x_l(t)e^{j2\pi f_0 t}] \quad (2.4-11)$$

该式表示任何实带通信号均可以用其等效低通来表示。

2. 带通信号的正交表示

将复基带信号 $x_l(t)$ 的实部定义为同相分量 $x_i(t)$, 虚部定义为正交分量 $x_q(t)$, 两者都是实低通信号, 则有

$$x_l(t) = x_i(t) + jx_q(t) \quad (2.4-12)$$

又由式(2.4-10)可得

$$\begin{aligned} x_l(t) &= (x(t) \cos 2\pi f_0 t + \hat{x}(t) \sin 2\pi f_0 t) \\ &\quad + j(\hat{x}(t) \cos 2\pi f_0 t - x(t) \sin 2\pi f_0 t) \end{aligned} \quad (2.4-13)$$

所以,

$$\begin{aligned} x_i(t) &= x(t) \cos 2\pi f_0 t + \hat{x}(t) \sin 2\pi f_0 t \\ x_q(t) &= \hat{x}(t) \cos 2\pi f_0 t - x(t) \sin 2\pi f_0 t \end{aligned} \quad (2.4-14)$$

由式(2.4-14)可解得

$$\begin{aligned} x(t) &= x_i(t) \cos 2\pi f_0 t - x_q(t) \sin 2\pi f_0 t \\ \hat{x}(t) &= x_i(t) \sin 2\pi f_0 t + x_q(t) \cos 2\pi f_0 t \end{aligned} \quad (2.4-15)$$

这表明任何实带通信号 $x(t)$ 都可以用两个实低通信号来表示, 即 $x_l(t)$ 的同相分量和正交分量。

3. 带通信号的幅度和相位表示

$x_l(t)$ 除了用其实部和虚部来表示, 还可以在极坐标中用其幅度和相位来表示。定义 $x_l(t)$ 的幅度 $a(t)$ 和相位 $\theta(t)$ 分别为

$$a(t) = \sqrt{x_i^2(t) + x_q^2(t)} \quad (2.4-16)$$

$$\theta(t) = \arctan \left[\frac{x_q(t)}{x_i(t)} \right] \quad (2.4-17)$$

则

$$x_l(t) = a(t)e^{j\theta(t)} \quad (2.4-18)$$

将该结果带入式(2.4-11), 可得

$$x(t) = a(t) \cos[2\pi f_0 t + \theta(t)] \quad (2.4-19)$$

图 2.5 为一带通信号及其包络。由上面的公式可知, $x_l(t)$ 、 $x_i(t)$ 、 $x_q(t)$ 、 $a(t)$ 和 $\theta(t)$ 均独立于中心频率 f_0 。当 f_0 取不同值时就会产生不同的带通信号 $x(t)$, 因此独立于特定中心频率 f_0 定义的带通信号的等效低通更有意义。

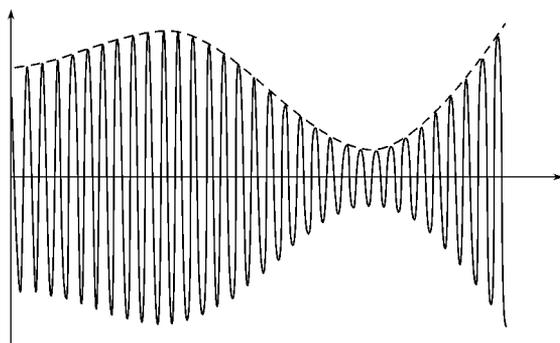


图 2.5 带通信号及其包络(虚曲线)

4. 低通变为带通的处理过程

式(2.4-15)和式(2.4-19)给出了用两个低通信号表示带通信号的两种方法,即用同相分量和正交分量表示,以及用幅度和相位表示。式(2.4-11)和式(2.4-15)是以低通信号定义的调制过程(即低通变为带通的处理过程)来表示带通信号,完成这种处理操作的系统称为调制器,图 2.6(a)和(b)中分别给出了实现式(2.4-11)和式(2.4-15)的一般调制器的结构,图中双线条和双线框表示复值信号和复操作。

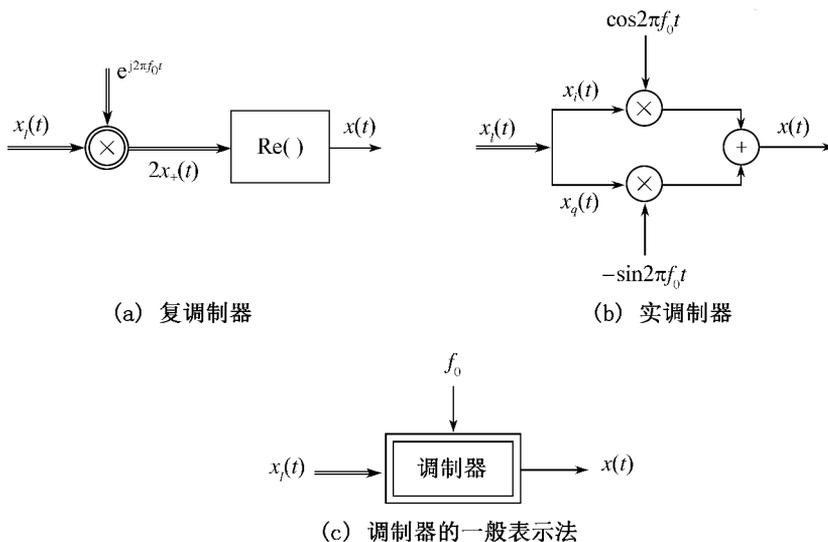


图 2.6 低通变为带通的调制过程

5. 带通变为低通的处理过程

类似地,式(2.4-13)和式(2.4-14)表示如何由带通信号 $x(t)$ 得到 $x_l(t)$ 或 $x_i(t)$ 和 $x_q(t)$,这种处理过程为从带通信号中提取低通信号,称为解调过程,如图 2.7(a)和(b)所示。在方框图中, \mathcal{H} 表示希尔伯特变换。

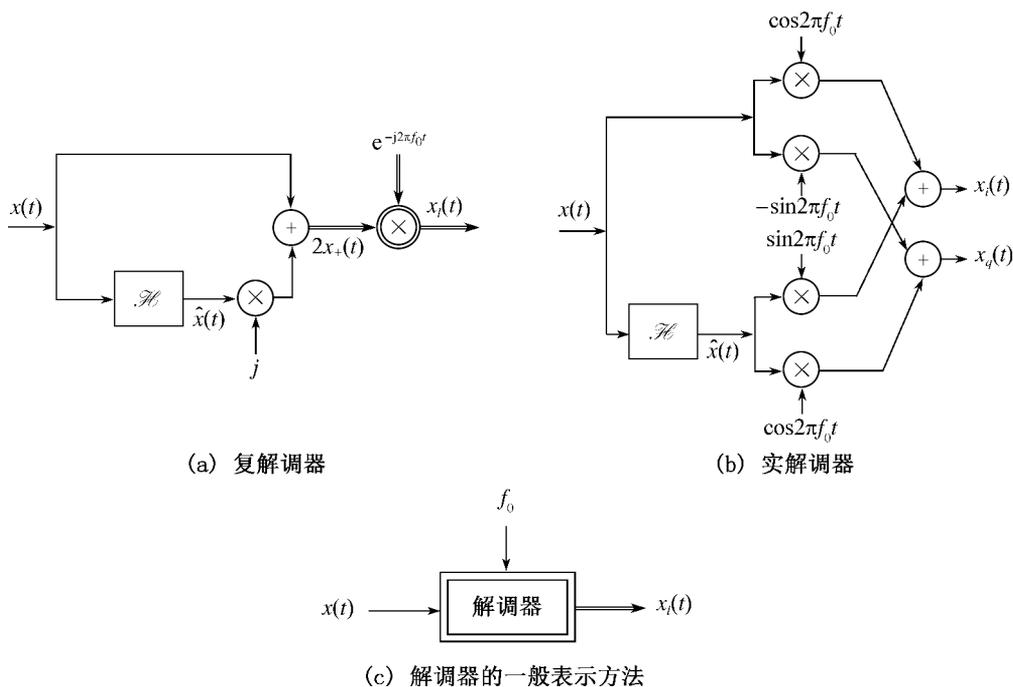


图 2.7 带通变为低通的解调过程

6. 带通信号 $x(t)$ 与其等效低通信号 $x_l(t)$ 的频谱与能量关系

(1) 频谱关系

$$\begin{aligned}
 X(f) &= \int_{-\infty}^{\infty} x(t) e^{-j2\pi ft} dt \\
 &= \int_{-\infty}^{\infty} \operatorname{Re}[x_l(t) e^{j2\pi f_0 t}] e^{-j2\pi ft} dt \\
 &= \frac{1}{2} \int_{-\infty}^{\infty} [x_l(t) e^{j2\pi f_0 t} + x_l^*(t) e^{-j2\pi f_0 t}] e^{-j2\pi ft} dt \\
 &= \frac{1}{2} [X_l(f - f_0) + X_l^*(-f - f_0)]
 \end{aligned} \tag{2.4-20}$$

(2) 能量关系

$$\begin{aligned}
 &\int_{-\infty}^{\infty} x^2(t) dt \\
 &= \int_{-\infty}^{\infty} a^2(t) \cos^2[2\pi f_0 t + \theta(t)] dt \\
 &= \frac{1}{2} \int_{-\infty}^{\infty} a^2(t) \{ \cos[4\pi f_0 t + 2\theta(t)] + 1 \} dt \\
 &= \frac{1}{2} \int_{-\infty}^{\infty} a^2(t) \cos[4\pi f_0 t + 2\theta(t)] dt + \frac{1}{2} \int_{-\infty}^{\infty} a^2(t) dt \\
 &= \frac{1}{2} \int_{-\infty}^{\infty} |x_l(t)|^2 \cos[4\pi f_0 t + 2\theta(t)] dt + \frac{1}{2} \int_{-\infty}^{\infty} |x_l(t)|^2 dt \\
 &= \frac{1}{2} \int_{-\infty}^{\infty} |x_l(t)|^2 dt
 \end{aligned} \tag{2.4-21}$$



即等效低通信号 $x_l(t)$ 中的能量是带通信号 $x(t)$ 中的能量的两倍。

2.4.3 带通信号的功率谱

在数字通信系统设计中,深入理解带通信号的功率谱密度具有重要的实际意义。一方面,当给定信道的传输带宽有限制时,信号的功率谱特性直接决定了哪些数字调制方式能够有效地适配该信道;另一方面,在选定了一种数字调制方式后,其信号的功率谱特性决定了传输该已调信号所需的最小信道带宽。本节基于 2.4.2 节建立的带通信号的等效低通表示理论,推导实带通信号与其等效低通信号的功率谱密度之间的关系,揭示实带通信号功率谱由其等效低通信号功率谱决定的本质。在此基础上,通过分析影响等效低通信号功率谱密度的关键因素,阐述利用预编码技术进行谱成形来适配信道的原理。

由式(2.4-11)可知,实带通信号 $x(t)$ 和其等效低通信号 $x_l(t)$ 满足下列关系:

$$x(t) = \text{Re}[x_l(t)e^{j2\pi f_0 t}] \quad (2.4-22)$$

则实带通信号 $x(t)$ 的自相关函数 $R_x(\tau)$ 和其等效低通信号 $x_l(t)$ 的自相关函数 $R_{x_l}(\tau)$ 之间的关系为

$$R_x(\tau) = \text{Re}[R_{x_l}(\tau)e^{j2\pi f_0 \tau}] \quad (2.4-23)$$

因此,根据维纳-辛钦定理可以推导出实带通信号 $x(t)$ 的功率谱密度 $S_x(f)$ 和其等效低通信号 $x_l(t)$ 的功率谱密度 $S_{x_l}(f)$ 之间的关系为

$$S_x(f) = \frac{1}{2}[S_{x_l}(f-f_0) + S_{x_l}(-f-f_0)] \quad (2.4-24)$$

该式表明,实带通信号 $x(t)$ 的功率谱密度 $S_x(f)$ 完全由其等效低通信号 $x_l(t)$ 的功率谱密度 $S_{x_l}(f)$ 决定。具体而言, $S_x(f)$ 是 $S_{x_l}(f)$ 搬移到载频 $+f_0$ 和搬移到其镜像频率 $-f_0$ 并考虑共轭对称性后的两个分量之和的 $\frac{1}{2}$ 。由于 $x_l(t)$ 是低通信号(其功率谱密度集中在零频附近), $S_x(f)$ 自然呈现出以载频 $\pm f_0$ 为中心的带通特性,且满足实信号功率谱密度的偶对称性。

由式(2.4-24)可以看出,控制带通信号 $x(t)$ 的功率谱密度形状,本质上等价于控制其等效低通信号 $x_l(t)$ 的功率谱密度形状。这在数字调制系统设计中提供了强大的灵活性,使得我们可以通过塑造等效低通信号的功率谱密度来使带通信号更好地适配信道的频谱特性或满足特定的频谱规范。

以线性数字调制信号为例,其等效低通信号 $v_l(t)$ 通常表示为

$$v_l(t) = \sum_{n=-\infty}^{\infty} I_n g(t-nT) \quad (2.4-25)$$

其中, $\{I_n\}$ 为平稳信息序列, $g(t)$ 为基本的调制脉冲, T 是符号间隔。假设信息序列 $\{I_n\}$ 是宽平稳的,则等效低通信号 $v_l(t)$ 的功率谱密度 $S_{v_l}(f)$ 可推导为

$$S_{v_l}(f) = \frac{1}{T} |G(f)|^2 \sum_{k=-\infty}^{\infty} R_I(k) e^{-j2\pi k f T} = \frac{1}{T} |G(f)|^2 S_I(f) \quad (2.4-26)$$

其中, $G(f)$ 为 $g(t)$ 的傅里叶变换, $R_I(k)$ 为信息序列 $\{I_n\}$ 的自相关函数,记



$$S_I(f) = \sum_{k=-\infty}^{\infty} R_I(k) e^{-j2\pi k f T} \quad (2.4-27)$$

则 $S_I(f)$ 表示离散信息序列 $\{I_n\}$ 的功率谱密度。

式(2.4-26)指出了影响等效低通信号功率谱密度形状,即通过式(2.4-24)影响已调带通信号功率谱密度形状的两个主要因素:一个是调制所用的基本脉冲形状,显然该脉冲的形状对已调信号功率谱的形状有重要的影响,较平滑的脉冲将导致更紧凑的功率谱密度。另一个因素是信息序列 $\{I_n\}$ 的功率谱密度,其取决于信息序列的相关特性。因此,通过控制信息序列的相关特性来控制已调信号功率谱密度已经成为一种非常有效的方法,其具体做法是将信息序列通过一个可逆滤波器,该滤波器控制已调信号的相关特性,因为它是可逆的,所以原始信息序列可由此恢复,这种技术称为预编码谱成形(Spectral Shaping by Precoding)。

例如,使用的预编码为 $J_n = I_n + \alpha I_{n-1}$,通过改变 α 的值就能控制已调信号的功率谱密度。一般,引入一个长度为 L 的记忆,预编码的形式为

$$J_n = \sum_{k=0}^L \alpha_k I_{n-k} \quad (2.4-28)$$

产生的已调波形为

$$v_i(t) = \sum_{k=-\infty}^{\infty} J_k g(t - kT) \quad (2.4-29)$$

因为预编码运算是线性运算,其功率谱密度为

$$S_{v_i}(f) = \frac{1}{T} |G(f)|^2 \left| \sum_{k=0}^L \alpha_k e^{-j2\pi k f T} \right|^2 S_I(f) \quad (2.4-30)$$

因此,改变 α_k 的值就可以控制已调信号的功率谱密度。

利用这种预编码谱成形技术,能够实现对数字调制带通信号功率谱密度的精确成形。这使得通信工程师能够优化信号频谱,使其高效地适配可用的信道带宽,满足频谱规范要求,并优化系统性能。

2.4.4 带通系统的等效低通表示

带通系统是指其传递函数位于频率 f_0 (及其镜像 $-f_0$) 附近的系统。更正式的带通系统定义是系统冲激响应 $h(t)$ 为带通信号的系统。因为 $h(t)$ 是带通的,其等效低通记为 $h_l(t)$,则

$$h(t) = \text{Re}[h_l(t) e^{j2\pi f_0 t}] \quad (2.4-31)$$

带通信号 $x(t)$ 通过冲激响应为 $h(t)$ 的带通系统后的输出也为带通信号,记作 $y(t)$ 。则输入与输出的频谱之间存在如下关系

$$Y(f) = X(f)H(f) \quad (2.4-32)$$

其中, $X(f)$ 和 $Y(f)$ 分别为 $x(t)$ 和 $y(t)$ 的傅里叶变换, $H(f)$ 是该带通系统的频率响应,也即冲激响应 $h(t)$ 的傅里叶变换,且 $H(f)$ 具有厄米特对称性,即满足 $H(f) = H^*(-f)$ 。

在频域中,有



$$\begin{aligned}
 Y_l(f) &= 2Y_+(f+f_0) \\
 &= 2Y(f+f_0)u_{-1}(f+f_0) \\
 &= 2X(f+f_0)H(f+f_0)u_{-1}(f+f_0) \\
 &= \frac{1}{2}[2X(f+f_0)u_{-1}(f+f_0)][2H(f+f_0)u_{-1}(f+f_0)] \\
 &= \frac{1}{2}X_1(f)H_1(f)
 \end{aligned} \tag{2.4-33}$$

则在时域中,有

$$y_l(t) = \frac{1}{2}x_l(t) * h_l(t) \tag{2.4-34}$$

式(2.4-33)和式(2.4-34)表明,当带通信号通过带通系统时,其等效低通形式的输入与输出之间的关系与带通形式的输入与输出之间的关系很相似,唯一的差别是在等效低通的等式中引入了 $\frac{1}{2}$ 因子。

2.4.5 带通平稳随机过程

一般情况下,带通随机过程可定义为宽平稳随机过程。若定义一个带通随机过程 $X(t)$ 为零均值的实宽平稳随机过程,其功率谱密度为 $S_X(f)$,且其谱密度的宽度远小于 f_0 。根据2.4.2小节可知, $X(t)$ 可表示为以下三种形式:

$$\begin{aligned}
 X(t) &= A(t)\cos[2\pi f_0 t + \theta(t)] \\
 &= X_i(t)\cos 2\pi f_0 t - X_q(t)\sin 2\pi f_0 t \\
 &= \text{Re}[X_l(t)e^{2\pi f_0 t}]
 \end{aligned} \tag{2.4-35}$$

设该带通随机过程 $X(t)$ 的等效低通 $X_l(t)$ 的实部和虚部,即其同相分量和正交分量分别为 $X_i(t)$ 和 $X_q(t)$,则有

$$\begin{aligned}
 X_i(t) &= X(t)\cos 2\pi f_0 t + \hat{X}(t)\sin 2\pi f_0 t \\
 X_q(t) &= \hat{X}(t)\cos 2\pi f_0 t - X(t)\sin 2\pi f_0 t
 \end{aligned} \tag{2.4-36}$$

对于 $X_i(t)$ 和 $X_q(t)$ 有如下结论:

- (1) $X_i(t)$ 和 $X_q(t)$ 是零均值联合宽平稳随机过程。
- (2) $X_i(t)$ 和 $X_q(t)$ 具有相同的功率谱密度。
- (3) $X_i(t)$ 和 $X_q(t)$ 都是低通随机过程,即它们的功率谱密度位于 $f=0$ 附近。

定义等效低通随机过程 $X_l(t)$ 为

$$X_l(t) = X_i(t) + jX_q(t) \tag{2.4-37}$$

下面推导等效低通随机过程 $X_l(t)$ 的自相关函数 $R_{X_l}(\tau)$ 和功率谱密度 $S_{X_l}(f)$ 的表达式。

因为假设 $X(t)$ 的均值为0,所以其希尔伯特变换 $\hat{X}(t)$ 的均值也为0,这是因为希尔伯特变换只是一种滤波运算。因此, $X_i(t)$ 和 $X_q(t)$ 都是零均值过程。

$X_l(t)$ 的自相关函数



$$\begin{aligned}
 & R_{X_i}(t+\tau, t) \\
 &= E[X_i(t+\tau)X_i(t)] \\
 &= E[(X(t+\tau)\cos 2\pi f_0(t+\tau) + \hat{X}(t+\tau)\sin 2\pi f_0(t+\tau)) \\
 &\quad \times (X(t)\cos 2\pi f_0 t + \hat{X}(t)\sin 2\pi f_0 t)] \\
 &= R_X(\tau)\cos 2\pi f_0(t+\tau)\cos 2\pi f_0 t \\
 &\quad + R_{X\hat{X}}(t+\tau, t)\cos 2\pi f_0(t+\tau)\sin 2\pi f_0 t \\
 &\quad + R_{\hat{X}X}(t+\tau, t)\sin 2\pi f_0(t+\tau)\cos 2\pi f_0 t \\
 &\quad + R_{\hat{X}\hat{X}}(t+\tau, t)\sin 2\pi f_0(t+\tau)\sin 2\pi f_0 t
 \end{aligned} \tag{2.4-38}$$

因为希尔伯特变换是随机过程通过线性时不变系统的结果,因此 $X(t)$ 和 $\hat{X}(t)$ 是联合宽平稳随机过程,所以式(2.4-38)中的所有自相关和互相关仅是关于 τ 的函数,且不难证明

$$R_{X\hat{X}}(\tau) = -\hat{R}_X(\tau), \quad R_{\hat{X}X}(\tau) = \hat{R}_X(\tau), \quad R_{\hat{X}\hat{X}}(\tau) = R_X(\tau) \tag{2.4-39}$$

将上面的式子代入式(2.4-38),并利用三角公式可得

$$R_{X_i}(\tau) = R_X(\tau)\cos(2\pi f_0\tau) + \hat{R}_X(\tau)\sin(2\pi f_0\tau) \tag{2.4-40}$$

同理,

$$R_{X_q}(\tau) = R_{X_i}(\tau) = R_X(\tau)\cos(2\pi f_0\tau) + \hat{R}_X(\tau)\sin(2\pi f_0\tau) \tag{2.4-41}$$

$$R_{X_i X_q}(\tau) = -R_{X_q X_i}(\tau) = R_X(\tau)\sin(2\pi f_0\tau) - \hat{R}_X(\tau)\cos(2\pi f_0\tau) \tag{2.4-42}$$

这些关系式表明, $X_i(t)$ 和 $X_q(t)$ 是零均值联合宽平稳随机过程,具有相等的自相关函数,即具有相等的功率谱密度。下面给出 $X_i(t)$ 和 $X_q(t)$ 的共同的功率谱密度及其互谱密度。

$$\begin{aligned}
 S_{X_i}(f) &= S_{X_q}(f) = F(R_{X_i}(\tau)) = F(R_{X_q}(\tau)) \\
 &= \begin{cases} S_X(f+f_0) + S_X(f-f_0) & |f| < f_0 \\ 0 & \text{其他} \end{cases}
 \end{aligned} \tag{2.4-43}$$

$$\begin{aligned}
 S_{X_i X_q}(f) &= -S_{X_q X_i}(f) = F(R_{X_i X_q}(\tau)) = -F(R_{X_q X_i}(\tau)) \\
 &= \begin{cases} j[S_X(f+f_0) - S_X(f-f_0)] & |f| < f_0 \\ 0 & \text{其他} \end{cases}
 \end{aligned} \tag{2.4-44}$$

式(2.4-43)表明,将 $X(t)$ 的功率谱密度左移、右移 f_0 后再相加,然后去除 $[-f_0, f_0]$ 之外的分量,就得到 $X_i(t)$ 的同相分量和正交分量的共同功率谱密度。该结果也表明 $X_i(t)$ 和 $X_q(t)$ 都是低通过程。式(2.4-44)表明,当 $|f| < f_0$ 时,若 $S_X(f+f_0) = S_X(f-f_0)$, 则 $S_{X_i X_q}(f) = 0$, 从而 $R_{X_i X_q}(\tau) = 0$ 。又因为 $X_i(t)$ 和 $X_q(t)$ 是零均值过程,则在这种情况下, $X_i(t)$ 和 $X_q(t)$ 是不相关的。当 $|f| < f_0$ 时, $S_X(f+f_0) = S_X(f-f_0)$ 成立于 $S_X(f)$ 围绕 f_0 对称的情况下,此时同相分量和正交分量是不相关的过程。

由平稳随机过程的相关知识可知, $X(t)$ 的等效低通 $X_l(t)$ 是平稳随机过程,其自相关函数为

$$\begin{aligned}
 R_{X_l}(\tau) &= 2R_{X_i}(\tau) + 2jR_{X_q X_i}(\tau) \\
 &= 2[R_X(\tau) + j\hat{R}_X(\tau)]e^{-j2\pi f_0\tau}
 \end{aligned} \tag{2.4-45}$$



比较式(2.4-45)和式(2.4-10)可知, $R_{X_I}(\tau)$ 是 $R_X(\tau)$ 的等效低通的 2 倍, 即等效低通过程 $X_I(t)$ 的自相关函数是带通过程 $X(t)$ 自相关函数的等效低通的 2 倍。

对式(2.4-45)两边进行傅里叶变换, 可得

$$S_{X_I}(f) = \begin{cases} 4S_X(f+f_0) & |f| < f_0 \\ 0 & \text{其他} \end{cases} \quad (2.4-46)$$

因此,

$$S_X(f) = \frac{1}{4} [S_{X_I}(f-f_0) + S_{X_I}(-f-f_0)] \quad (2.4-47)$$

带通白噪声是一种带通平稳随机过程。下面简单介绍一下带通白噪声:

白噪声是一种在整个频率范围内具有平坦(恒定)功率谱密度的随机信号。通过频谱平坦的带通滤波器的白噪声过程称为带通白噪声。带宽为 B 的带通白噪声的等效低通信号的功率谱密度为

$$S_{zz}(f) = \begin{cases} N_0 & |f| \leq \frac{B}{2} \\ 0 & |f| > \frac{B}{2} \end{cases} \quad (2.4-48)$$

带通白噪声的等效低通的自相关函数为

$$R_{zz}(\tau) = N_0 B \frac{\sin \pi B \tau}{\pi B \tau} \quad (2.4-49)$$

当 $B \rightarrow \infty$ 时,

$$R_{zz}(\tau) \cong N_0 \delta(\tau) \quad (2.4-50)$$

带通白噪声的等效低通的同相分量和正交分量的互相关函数为

$$R_{xy}(\tau) = 0, \quad \forall \tau \quad (2.4-51)$$

另外, 带通白噪声的等效低通的自相关函数与其同相分量及正交分量的自相关函数是相同的, 即

$$R_{zz}(\tau) = R_{xx}(\tau) = R_{yy}(\tau) \quad (2.4-52)$$

其中, $R_{xx}(\tau)$ 和 $R_{yy}(\tau)$ 是同相分量和正交分量的自相关函数。

2.5 信号空间与矢量表示

在数字已调信号分析中, 信号的信号空间(或矢量)表示法是一种非常有效且有用的工具。本节将阐述任何信号集均可等效为一个矢量集, 信号具有矢量的基本性质。同时, 将矢量空间的概念和矢量的标准正交基表示方法延伸到信号空间中, 着重介绍格拉姆-施密特(Gram-Schmidt)正交化过程, 以及一个信号波形集的信号空间表示法(或信号星座图)。



2.5.1 矢量空间的概念

在 n 维空间中, 矢量 \mathbf{v} 可以用它的 n 个分量 v_1, v_2, \dots, v_n 表征。令矢量 \mathbf{v} 表示一个列矢量, 即 $\mathbf{v} = [v_1, v_2, \dots, v_n]^T$, 其中, $[\cdot]^T$ 表示矩阵的转置。两个 n 维矢量 $\mathbf{v}_1 = [v_{11}, v_{12}, \dots, v_{1n}]^T$ 和 $\mathbf{v}_2 = [v_{21}, v_{22}, \dots, v_{2n}]^T$ 的内积定义为

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \mathbf{v}_1 \cdot \mathbf{v}_2 = \sum_{i=1}^n v_{1i} v_{2i}^* = \mathbf{v}_2^H \mathbf{v}_1 \quad (2.5-1)$$

其中, $[\cdot]^H$ 表示矩阵的厄米特转置(共轭转置), 即先对矩阵转置再共轭其元素。

矢量也可以表示成正交单位矢量或标准正交基 $\mathbf{e}_i, 1 \leq i \leq n$ 的线性组合, 即

$$\mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i \quad (2.5-2)$$

式中, 正交单位矢量 \mathbf{e}_i 的长度为 1, v_i 是矢量 \mathbf{v} 在单位矢量 \mathbf{e}_i 上的投影, 即 $v_i = \langle \mathbf{v}, \mathbf{e}_i \rangle$ 。如果 $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$, 则称矢量 \mathbf{v}_1 与 \mathbf{v}_2 相互正交。更为一般的情况是, 一组 m 个矢量集 $\mathbf{v}_k, 1 \leq k \leq m$, 若对所有 $1 \leq i, j \leq m$ 且 $i \neq j$, 都有 $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$, 则称这组矢量是相互正交的。

矢量 \mathbf{v} 的范数记作 $\|\mathbf{v}\|$, 定义为

$$\|\mathbf{v}\| = (\langle \mathbf{v}, \mathbf{v} \rangle)^{\frac{1}{2}} = \sqrt{\sum_{i=1}^n |v_i|^2} \quad (2.5-3)$$

这在 n 维空间中就是矢量的长度。

如果一组 m 个矢量相互正交且每个矢量具有单位范数, 则称这组矢量为标准(归一化)正交的。如果一组 m 个矢量中没有一个矢量能用其他矢量的线性组合来表示, 则称这组矢量是线性独立的。

任何两个 n 维矢量 \mathbf{v}_1 和 \mathbf{v}_2 满足三角不等式

$$\|\mathbf{v}_1 + \mathbf{v}_2\| \leq \|\mathbf{v}_1\| + \|\mathbf{v}_2\| \quad (2.5-4)$$

同时也满足柯西-施瓦茨不等式

$$|\langle \mathbf{v}_1, \mathbf{v}_2 \rangle| \leq \|\mathbf{v}_1\| \cdot \|\mathbf{v}_2\| \quad (2.5-5)$$

下面介绍格拉姆-施密特正交化过程, 它是从一组 n 维矢量 $\mathbf{v}_i, 1 \leq i \leq m$ 构造一组标准正交矢量的过程。首先, 从这组矢量中任意选择一个矢量, 如 \mathbf{v}_1 , 对它的长度进行归一化, 得到第一个矢量

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \quad (2.5-6)$$

接着, 选择 \mathbf{v}_2 , 先减去 \mathbf{v}_2 在 \mathbf{u}_1 上的投影, 得到

$$\mathbf{u}'_2 = \mathbf{v}_2 - (\langle \mathbf{v}_2, \mathbf{u}_1 \rangle) \mathbf{u}_1 \quad (2.5-7)$$

再将矢量 \mathbf{u}'_2 归一化成单位长度, 得到

$$\mathbf{u}_2 = \frac{\mathbf{u}'_2}{\|\mathbf{u}'_2\|} \quad (2.5-8)$$

再选择 \mathbf{v}_3 并减去 \mathbf{v}_3 在 \mathbf{u}_1 和 \mathbf{u}_2 上的投影, 得到 \mathbf{u}'_3 , 并归一化得到 \mathbf{u}_3 , 即



$$\mathbf{u}'_3 = \mathbf{v}_3 - (\langle \mathbf{v}_3, \mathbf{u}_1 \rangle) \mathbf{u}_1 - (\langle \mathbf{v}_3, \mathbf{u}_2 \rangle) \mathbf{u}_2 \quad (2.5-9)$$

$$\mathbf{u}_3 = \frac{\mathbf{u}'_3}{\|\mathbf{u}'_3\|} \quad (2.5-10)$$

继续选择先前未使用的矢量,重复上面的过程,即

$$\begin{cases} \mathbf{u}'_j = \mathbf{v}_j - \sum_{i=1}^{j-1} (\langle \mathbf{v}_j, \mathbf{u}_i \rangle) \mathbf{u}_i \\ \mathbf{u}_j = \frac{\mathbf{u}'_j}{\|\mathbf{u}'_j\|} \end{cases} \quad (2.5-11)$$

则 \mathbf{u}_j 可构成一组 N 个标准正交矢量,其中 $N \leq \min(m, n)$ 。

2.5.2 信号空间的概念

类似矢量的情况,对于一组信号也可以用相似的方法进行处理。定义两个一般的复信号 $x_1(t)$ 和 $x_2(t)$ 的内积为

$$\langle x_1(t), x_2(t) \rangle = \int_{-\infty}^{\infty} x_1(t) x_2^*(t) dt \quad (2.5-12)$$

如果 $x_1(t)$ 和 $x_2(t)$ 的内积为 0,则这两个信号是正交的。信号 $x(t)$ 的范数定义为

$$\|x(t)\| = \left(\int_{-\infty}^{\infty} |x(t)|^2 dt \right)^{\frac{1}{2}} = \sqrt{E_x} \quad (2.5-13)$$

其中, E_x 为信号 $x(t)$ 的能量。对于一个由 m 个信号组成的信号集,如果这 m 个信号相互正交且其范数均为 1,则该信号集是标准正交的。如果没有一个信号能表示成其余信号的线性组合,则该信号集是线性独立的。

任何两个信号均满足三角不等式

$$\|x_1(t) + x_2(t)\| \leq \|x_1(t)\| + \|x_2(t)\| \quad (2.5-14)$$

同时也满足柯西-施瓦茨不等式

$$|\langle x_1(t), x_2(t) \rangle| \leq \|x_1(t)\| \cdot \|x_2(t)\| \quad (2.5-15)$$

2.5.3 格拉姆-施密特过程

假设要从一组能量有限的信号 $\{s_i(t), i=1, 2, \dots, M\}$ 构造一组标准正交信号,可以利用格拉姆-施密特正交化过程来实现。如 2.5.1 节所述,首先,从 $\{s_i(t)\}$ 中任选一个信号,如 $s_1(t)$,对其长度进行归一化,得到第一个标准正交信号

$$\phi_1(t) = \frac{s_1(t)}{\|s_1(t)\|} = \frac{s_1(t)}{\sqrt{E_1}} \quad (2.5-16)$$

其中, $E_1 = \int_{-\infty}^{\infty} |s_1(t)|^2 dt$ 为信号 $s_1(t)$ 的能量。

接着,为构建第二个标准正交信号,可以选择 $s_2(t)$,先通过 $s_2(t)$ 减去 $s_2(t)$ 在 $\phi_1(t)$ 上的投影,得到



$$\phi_2'(t) = s_2(t) - c_{21}\phi_1(t) \quad (2.5-17)$$

其中, $c_{21} = \langle s_2(t), \phi_1(t) \rangle = \int_{-\infty}^{\infty} s_2(t)\phi_1^*(t)dt$ 为 $s_2(t)$ 在 $\phi_1(t)$ 上的投影。此时, $\phi_2'(t)$ 正交于 $\phi_1(t)$, 但是它不具有单位能量, 因此, 令

$$\phi_2(t) = \frac{\phi_2'(t)}{\|\phi_2'(t)\|} = \frac{\phi_2'(t)}{\sqrt{E_2}} \quad (2.5-18)$$

其中, $E_2 = \int_{-\infty}^{\infty} |\phi_2'(t)|^2 dt$ 为信号 $\phi_2'(t)$ 的能量。这样, 就得到第二个标准正交信号。

继续下去, 假设已经构建得到 $n-1$ 个标准正交信号 $\phi_i(t), i=1, 2, \dots, n-1$, 任意选择之前未使用的信号 $s_n(t)$, 利用 $s_n(t)$ 减去 $s_n(t)$ 在 $\phi_i(t), i=1, 2, \dots, n-1$ 上的投影, 并归一化为单位能量, 则可构建第 n 个标准正交信号, 即

$$\phi_n'(t) = s_n(t) - \sum_{i=1}^{n-1} c_{n,i}\phi_i(t) \quad (2.5-19)$$

$$\phi_n(t) = \frac{\phi_n'(t)}{\|\phi_n'(t)\|} = \frac{\phi_n'(t)}{\sqrt{E_n}} \quad (2.5-20)$$

其中, $c_{n,i} = \langle s_n(t), \phi_i(t) \rangle = \int_{-\infty}^{\infty} s_n(t)\phi_i^*(t)dt$ 为 $s_n(t)$ 在 $\phi_i(t)$ 上的投影, $E_n = \int_{-\infty}^{\infty} |\phi_n'(t)|^2 dt$ 为信号 $\phi_n'(t)$ 的能量。

按照这样的构建方法, 当所有 M 个信号 $\{s_i(t), i=1, 2, \dots, M\}$ 处理完毕, 则可以构建 $N \leq M$ 个标准正交信号。如果所有信号是线性独立的, 即没有一个信号是其他信号的线性组合, 那么信号空间维数 $N=M$ 。

2.5.4 信号的矢量表示

本节将推导一个信号波形的矢量表示法, 并证明信号波形与它的矢量表示之间的等价性。

假定 $s(t)$ 具有有限能量, 即

$$E_s = \int_{-\infty}^{\infty} |s(t)|^2 dt < \infty \quad (2.5-21)$$

并且假定存在一个标准正交信号集 $\{\phi_n(t), n=1, 2, \dots, N\}$, 则信号 $s(t)$ 可以用这些信号的加权线性组合来近似, 即

$$\hat{s}(t) = \sum_{n=1}^N s_n \phi_n(t) \quad (2.5-22)$$

其中, $s_n, 1 \leq n \leq N$ 是信号 $s(t)$ 的近似式中的系数。近似式导致的近似误差为

$$e(t) = s(t) - \hat{s}(t) \quad (2.5-23)$$

近似误差的能量为

$$E_e = \int_{-\infty}^{\infty} |s(t) - \hat{s}(t)|^2 dt = \int_{-\infty}^{\infty} \left| s(t) - \sum_{n=1}^N s_n \phi_n(t) \right|^2 dt \quad (2.5-24)$$



要选择系数 $\{s_n\}$ 使得近似误差的能量最小化, 可以用通过对式(2.5-24)中每一个系数 $\{s_n\}$ 进行微分, 并令一阶导数为零的方法求得。也可以利用估计理论中基于均方误差准则的结论, 即当误差正交于级数展开式中的每一个函数时, 可得到相对于 $\{s_n\}$ 的 E_e 的最小值, 即需要

$$\int_{-\infty}^{\infty} \left[s(t) - \sum_{n=1}^N s_n \phi_n(t) \right] \phi_n^*(t) dt = 0, \quad n=1, 2, \dots, N \quad (2.5-25)$$

又因为 $\{\phi_n(t), n=1, 2, \dots, N\}$ 是标准正交的, 则(2.5-25)可简化为

$$s_n = \langle s(t), \phi_n(t) \rangle = \int_{-\infty}^{\infty} s(t) \phi_n^*(t) dt, \quad n=1, 2, \dots, N \quad (2.5-26)$$

因此, 所求系数可以由信号 $s(t)$ 投影到 $\{\phi_n(t), n=1, 2, \dots, N\}$ 的每个函数上的方法得到。也就是说, $\hat{s}(t)$ 是 $s(t)$ 在 $\{\phi_n(t), n=1, 2, \dots, N\}$ 所架构的 N 维信号空间上的投影, 所以它正交于误差信号 $e(t) = s(t) - \hat{s}(t)$, 即 $\langle e(t), \hat{s}(t) \rangle = 0$ 。最小均方近似误差为

$$\begin{aligned} E_{\min} &= \int_{-\infty}^{\infty} e(t) e^*(t) dt \\ &= \int_{-\infty}^{\infty} |s(t)|^2 dt - \int_{-\infty}^{\infty} \sum_{n=1}^N s_n \phi_n(t) s^*(t) dt \\ &= E_s - \sum_{n=1}^N |s_n|^2 \end{aligned} \quad (2.5-27)$$

当最小均方近似误差 $E_{\min} = 0$ 时,

$$E_s = \sum_{n=1}^N |s_n|^2 = \int_{-\infty}^{\infty} |s(t)|^2 dt \quad (2.5-28)$$

此时

$$s(t) = \sum_{n=1}^N s_n \phi_n(t) \quad (2.5-29)$$

这表明只有在近似误差具有零能量时, $s(t)$ 与其级数展开式 $\hat{s}(t)$ 才相等。

当一个有限能量信号 $s(t)$ 用式(2.5-29)的级数展开且 $E_{\min} = 0$ 时, 标准正交信号集 $\{\phi_n(t), n=1, 2, \dots, N\}$ 称为完备的。

对于 M 进制数字通信系统, 假设采用 M 种不同波形 $\{s_i(t), i=1, 2, \dots, M\}$ 表示不同发送符号, 其中每个信号波形对应一个发送符号。由式(2.5-29)可知, 只要构建好标准正交波形集 $\{\phi_n(t), n=1, 2, \dots, N\}$, 这 M 种信号波形就都能表示成 $\{\phi_n(t)\}$ 的线性组合, 即

$$s_i(t) = \sum_{n=1}^N s_{in} \phi_n(t), \quad i=1, 2, \dots, M \quad (2.5-30)$$

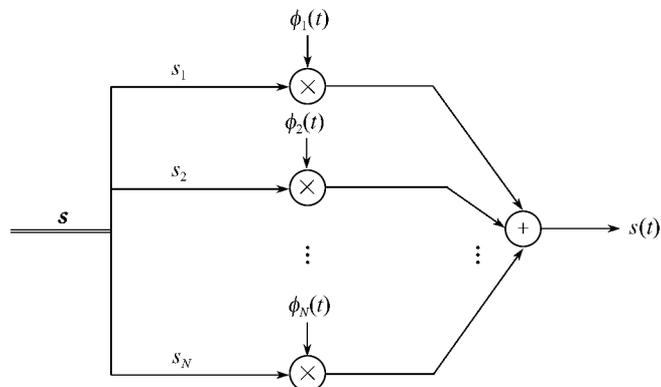
基于式(2.5-30), 每一个信号波形 $s_i(t)$ 都可以表示成矢量

$$\mathbf{s}_i = [s_{i1} \quad s_{i2} \quad \dots \quad s_{iN}]^T \quad (2.5-31)$$

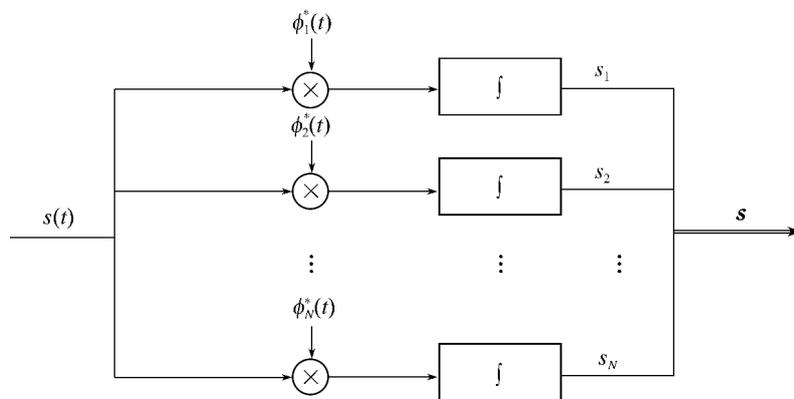
或者等效地表示成 N 维(一般为复数)信号空间的一个点, 其坐标为 $\{s_{in}, n=1, 2, \dots, N\}$ 。因此, 一组 M 个信号集 $\{s_i(t), i=1, 2, \dots, M\}$ 可用 $N(N \leq M)$ 维信号空间的一组 M 个矢量来表示, 相应的矢量集 $\{\mathbf{s}_i, i=1, 2, \dots, M\}$ 称为波形集 $\{s_i(t), i=1, 2, \dots, M\}$ 的信号空间



表示或星座图。如果原信号是实的,则矢量空间表示是在 \mathbb{R}^N 中;如果原信号是复的,则矢量空间表示是在 \mathbb{C}^N 中。图2.8展现了由信号得到等效矢量的过程(信号映射成矢量)及其逆过程(矢量映射成信号)。



(a) 矢量映射成信号



(b) 信号映射成矢量

图 2.8 矢量与信号相互映射过程

由基 $\{\phi_n(t), n=1, 2, \dots, N\}$ 的正交性可得

$$E_i = \int_{-\infty}^{\infty} |s_i(t)|^2 dt = \sum_{n=1}^N |s_{in}|^2 = \|\mathbf{s}_i\|^2 \quad (2.5-32)$$

式(2.5-32)表明第 i 个信号的能量即矢量长度的平方,也等价于 N 维空间中原点到信号点 \mathbf{s}_i 的欧氏距离的平方,因此,任何信号都可以用几何方式表示成由标准正交函数 $\{\phi_n(t)\}$ 构建的信号空间中的一个点。由基的正交性还能得到

$$\langle s_k(t), s_l(t) \rangle = \langle \mathbf{s}_k, \mathbf{s}_l \rangle \quad (2.5-33)$$

式(2.5-33)表明两个信号的内积等于其相应矢量的内积。

基于以上分析,如果存在带通信号集

$$s_i(t) = \text{Re}[s_{il}(t)e^{j2\pi f_0 t}], \quad i=1, 2, \dots, M \quad (2.5-34)$$

其中, $\{s_{il}(t)\}$ 表示带通信号 $\{s_i(t)\}$ 的等效低通信号,则 $s_i(t)$ 的能量为



$$E_i = \int_{-\infty}^{\infty} s_i^2(t) dt = \frac{1}{2} \int_{-\infty}^{\infty} |s_{il}(t)|^2 dt = \|s_i\|^2 \quad (2.5-35)$$

信号 $s_k(t)$ 和 $s_m(t)$ 的相关系数为

$$\rho_{k,m} = \frac{1}{2\sqrt{E_k E_m}} \int_{-\infty}^{\infty} s_{kl}(t) s_{ml}^*(t) dt \quad (2.5-36)$$

$$\rho_r = \frac{1}{\sqrt{E_k E_m}} \int_{-\infty}^{\infty} s_k(t) s_m(t) dt = \frac{\langle s_k, s_m \rangle}{\|s_k\| \cdot \|s_m\|} \quad (2.5-37)$$

其中, $\rho_r = \text{Re}[\rho_{k,m}]$, 则信号 $s_k(t)$ 和 $s_m(t)$ 之间的距离平方为

$$\begin{aligned} (d_{km}^{(e)})^2 &= \|s_k - s_m\|^2 \\ &= \int_{-\infty}^{\infty} [s_k(t) - s_m(t)]^2 dt \\ &= E_k + E_m - 2\sqrt{E_k E_m} \rho_r \end{aligned} \quad (2.5-38)$$

2.6 小结

本章主要讨论了确定信号在时域和频域上的主要性质,介绍了随机过程的分布和数字特征等概念,着重研究了通信系统中常见的平稳随机过程,以及随机过程通过线性时不变系统后的情况。此外,还给出了带通信号与带通系统的等效低通表示方法以及信号波形集的信号空间(或矢量)表示方法。

2.7 思考题

2-1 求图 2.9 所示的矩形脉冲的频谱密度、能量谱密度和自相关函数。

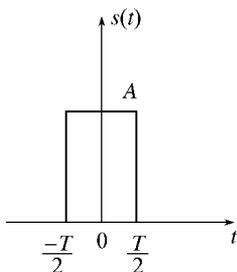


图 2.9 矩形脉冲(门函数)

2-2 在调幅(AM)收音机的接收端,接收到的信号包含一个载波频率为 f_c 的余弦载波信号,该信号有一个随机相位 Θ ,它是均匀分布 $U(0, 2\pi)$ 随机变量的一个样本值。接收到的载波信号是:



$$X(t) = A \cos(2\pi f_c t + \Theta) \quad (2.7-1)$$

则这个随机过程 $X(t)$ 的期望值和自相关函数分别是多少?

2-3 已知随机过程 $z(t) = m(t) \cos(\omega_c t + \theta)$, 其中, $m(t)$ 是宽平稳随机过程且其自相关函数为

$$R_m(\tau) = \begin{cases} 1 + \tau, & -1 < \tau < 0 \\ 1 - \tau, & 0 \leq \tau < 1 \\ 0, & \text{其他} \end{cases} \quad (2.7-2)$$

随机变量 θ 在 $[0, 2\pi]$ 上服从均匀分布, 它与 $m(t)$ 彼此统计独立。

- 证明 $z(t)$ 是宽平稳随机过程;
- 求 $z(t)$ 的自相关函数, 并画出波形;
- 求 $z(t)$ 的功率谱密度及功率。

2-4 给定实带通信号

$$s(t) = m(t) \cos(2\pi f_c t) - \hat{m}(t) \sin(2\pi f_c t), \quad (2.7-3)$$

其中 $\hat{m}(t)$ 为 $m(t)$ 的希尔伯特变换。

- 求其等效低通信号 $s_l(t)$ 的表达式;
- 若 $m(t) = \text{sinc}(200\pi t)$, 求 $s_l(t)$ 的频谱 $S_l(f)$ 。

2-5 若带通白噪声 $n(t)$ 的功率谱密度为 $N_0/2$ (频带 $|f \pm f_c| \leq W$)。

- 求等效低通噪声 $n_{LP}(t)$ 的功率谱密度;
- 证明 $n(t) = \text{Re}[n_{LP}(t) e^{j2\pi f_c t}]$ 且 $E[|n_{LP}(t)|^2] = 2N_0W$;
- 若 $n_{LP}(t) = n_I(t) + jn_Q(t)$, 求 $n_I(t)$ 和 $n_Q(t)$ 的互相关函数。

2-6 给出图 2.10 所示信号波形集 $\{s_i(t), i=1, 2, 3, 4\}$ 的格拉姆-施密特正交化过程及相应的标准正交基。

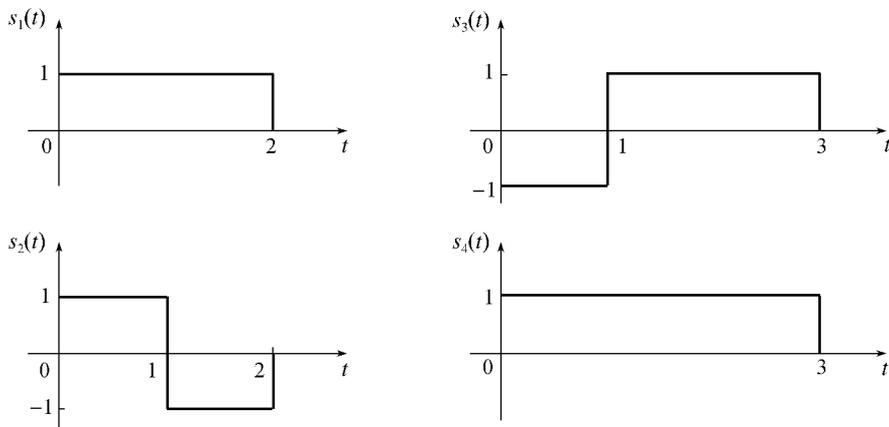


图 2.10 信号 $\{s_i(t), i=1, 2, 3, 4\}$ 的波形图

2-7 给定信号集:

$$s_1(t) = 1, \quad s_2(t) = t, \quad s_3(t) = t^2, \quad 0 \leq t \leq 1 \quad (2.7-4)$$

使用格拉姆-施密特正交化过程构造标准正交基 $\{\phi_1(t), \phi_2(t), \phi_3(t)\}$ 。



(a) 逐步推导正交化过程;

(b) 将信号 $s(t) = 2t + 3$ 表示为该正交基的线性组合。

2-8 确定一组基函数并利用它们来表示图 2.11 所示的四个波形,其中各波形由矢量 s_1, s_2, s_3 和 s_4 标记。

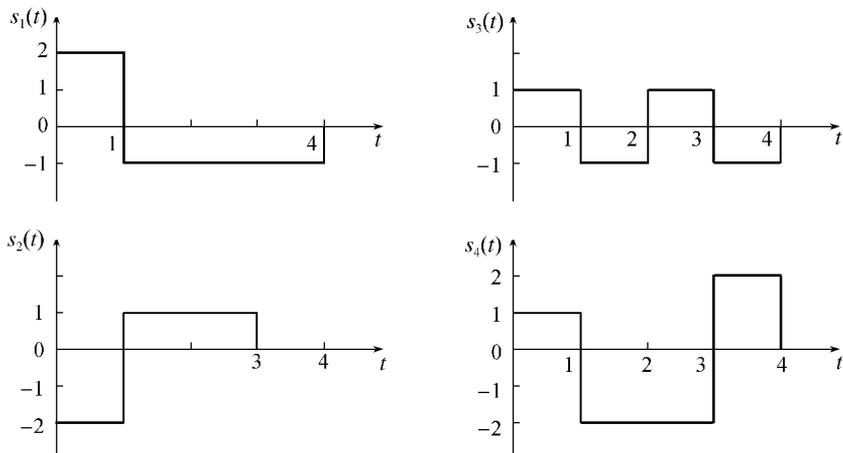


图 2.11 信号 $\{s_i(t), i=1,2,3,4\}$ 的波形图

参考文献

南京大学出版社

- [1] Proakis, J. G., & Salehi, M. Digital Communications (5th ed.) [M]. McGraw-Hill Science/Engineering/Math, 2007.
- [2] 樊昌信, 曹丽娜. 通信原理(第 7 版) [M]. 北京: 国防工业出版社, 2012.
- [3] 罗鹏飞, 张文明. 随机信号分析与处理 [M]. 北京: 清华大学出版社, 2021.
- [4] 钱玲, 谷亚林, 王海青. 信号与系统(第 5 版) [M]. 北京: 电子工业出版社, 2017.

第3章 数字调制

3.1 引言

数字通信中,待传输的信息以二进制比特或多进制数字符号序列形式出现。其核心目标是在给定信道条件下,将这串数字符号可靠地送到接收端。然而,信道本身往往伴随噪声、衰减、失真、衰落和外部干扰等效应,对信号造成损伤。为了既能准确携带数字符号信息,又能适应信道带宽和抗干扰特性,就必须将数字符号序列映射为适合在该信道上传输的模拟波形。所谓“携带数字符号信息”,是指接收端可从所接收的波形中恢复出原始的发送符号;所谓“适应信道带宽和抗干扰特性”,则意味着调制波形的带宽要与信道带宽相当,并具有良好的抗噪声与抗失真能力。不同信道会产生不同的损伤机制,因而针对各类信道设计的调制信号也各不相同。这个将数字符号序列映射成可在信道上传输的信号的过程,即称为数字调制(数字信号传输)。

在数字调制中,数字符号序列到发送波形的映射过程可分为无记忆调制和有记忆调制。在无记忆调制中,每个符号周期内生成的调制波形仅取决于当前输入的符号,与之前发送的符号无关,即调制过程是逐符号独立进行的。而在有记忆调制中,当前符号周期内生成的调制波形不仅取决于当前输入的符号,还依赖于之前发送的符号,即调制波形承载了历史信息。

典型的无记忆调制技术有脉冲幅度调制(Pulse Amplitude Modulation, PAM)、正交幅度调制(Quadrature Amplitude Modulation, QAM)、频移键控(Frequency-Shift Keying, FSK)和相移键控(Phase Shift Keying, PSK)等。例如,在PAM、PSK和QAM调制中,载波的幅度、相位和幅相组合随当前数字符号变化而改变;在FSK中,载波频率随当前数字符号变化而改变。无记忆调制的优点在于解调结构相对简单、实现容易,但在某些信道的非线性效应下性能可能受限,且频谱效率可能不如某些有记忆调制。常见的有记忆调制方式有最小频移键控(Minimum Shift Keying, MSK)、高斯最小频移键控(Gaussian MSK, GMSK)、连续相位调制(Continuous Phase Modulation, CPM)等。MSK和GMSK通过控制相位变化的连续性,使得相位路径平滑,其当前波形的相位不仅取决于当前符号,还受到之前发送符号的影响,以保证相位的连续。有记忆调制通常能获得更紧凑的频谱、更高的带宽效率(频谱利用率)和更好的抗噪性能,但是其调制和解调算法通常比无记忆调制更为复杂。

后续章节将依次分析各类无记忆和有记忆调制方式的原理、信号空间表示及性能特性。



3.2 无记忆调制

一般来讲,用来在通信信道上传输信息的波形 $s_m(t)$ 可以具有多种形式,其区别通常体现在幅度、相位或频率等参数上,也可以是这些参数的组合。下面分别从幅度、相位和频率等角度考察各类线性调制方式,首先介绍脉冲幅度调制 PAM。假设调制器输入的二进制数字序列的速率均为 R bit/s。

3.2.1 脉冲幅度调制(PAM)

幅度调制是用调制波形的不同幅度对应表示每种待传数据符号。在数字 PAM 中,信号波形可以表示为

$$s_m(t) = A_m g(t), 1 \leq m \leq M \quad (3.2-1)$$

式中, $g(t)$ 是持续时间为 T 的成形脉冲, $\{A_m, 1 \leq m \leq M\}$ 表示 M 个可能的幅度集合,它对应于 $M=2^k$ 个可能的比特组的符号。通常,信号幅度 A_m 取离散值

$$A_m = 2m - 1 - M, m = 1, 2, \dots, M \quad (3.2-2)$$

即幅度是 $\pm 1, \pm 3, \pm 5, \dots, \pm(M-1)$ 。波形 $g(t)$ 是实信号脉冲,其形状会影响发送信号的频谱。

1. PAM 信号的能量

PAM 信号 $s_m(t)$ 的能量为

$$\mathcal{E}_m = \int_{-\infty}^{\infty} A_m^2 g^2(t) dt = A_m^2 \mathcal{E}_g \quad (3.2-3)$$

式中, \mathcal{E}_g 是 $g(t)$ 的能量。由此可知,假设各符号等概率发送,则符号平均能量 \mathcal{E}_{avg} :

$$\begin{aligned} \mathcal{E}_{\text{avg}} &= \frac{\mathcal{E}_g}{M} \sum_{m=1}^M A_m^2 \\ &= \frac{2\mathcal{E}_g}{M} [1^2 + 3^2 + 5^2 + \dots + (M-1)^2] \\ &= \frac{2\mathcal{E}_g}{M} \times \frac{M(M^2-1)}{6} = \frac{(M^2-1)\mathcal{E}_g}{3} \end{aligned} \quad (3.2-4)$$

和比特平均能量 $\mathcal{E}_{\text{bavg}}$:

$$\mathcal{E}_{\text{bavg}} = \frac{(M^2-1)\mathcal{E}_g}{3 \log_2 M} \quad (3.2-5)$$

上面描述的是基带脉冲幅度调制(PAM),其中没有载波调制。在许多情况下,PAM 信号被载波调制成带通信号,其等效低通形式 $s_m(t)$ 为 $A_m g(t)$ 。在这种情况下,带通 PAM 信号可表示为

$$s_m(t) = \text{Re}[s_{ml}(t)e^{j2\pi f_c t}] = \text{Re}[A_m g(t)e^{j2\pi f_c t}] = A_m g(t) \cos(2\pi f_c t) \quad (3.2-6)$$



式中, f_c 为载波频率。比较式(3.2-1)和式(3.2-6)可知, 等效低通信号的能量是带通信号能量的 2 倍, 即带通信号的能量 \mathcal{E}_m 为

$$\mathcal{E}_m = \frac{1}{2} A_m^2 \mathcal{E}_g \quad (3.2-7)$$

由式(3.2-4)和式(3.2-5)可得出, 对于带通信号有

$$\mathcal{E}_{\text{avg}} = \frac{(M^2 - 1) \mathcal{E}_g}{6} \quad (3.2-8)$$

和

$$\mathcal{E}_{\text{bavg}} = \frac{(M^2 - 1) \mathcal{E}_g}{6 \log_2 M} \quad (3.2-9)$$

2. PAM 信号的矢量表示

显然, PAM 信号是一维的, 因为所有的信号都是同一基本信号波形, 只有幅度不同。基带 PAM 信号 $s_m(t) = A_m g(t)$ 的基可表示为

$$\phi_g(t) = \frac{g(t)}{\sqrt{\mathcal{E}_g}} \quad (3.2-10)$$

带通 PAM 信号的基可以表示为

$$\phi_g(t) = \sqrt{\frac{2}{\mathcal{E}_g}} g(t) \cos(2\pi f_c t) \quad (3.2-11)$$

于是 PAM 信号可由基信号表示为

$$s_m(t) = A_m \sqrt{\mathcal{E}_g} \phi_g(t) \text{ (基带 PAM)} \quad (3.2-12)$$

$$s_m(t) = A_m \sqrt{\frac{\mathcal{E}_g}{2}} \phi_g(t) \text{ (带通 PAM)} \quad (3.2-13)$$

由上式可见, 这些信号的一维矢量表示形式为

$$s_m = A_m \sqrt{\mathcal{E}_g}, A_m = \pm 1, \pm 3, \dots, \pm(M-1) \text{ (基带 PAM)} \quad (3.2-14)$$

$$s_m = A_m \sqrt{\frac{\mathcal{E}_g}{2}}, A_m = \pm 1, \pm 3, \dots, \pm(M-1) \text{ (带通 PAM)} \quad (3.2-15)$$

图 3.1 所示为 $M=2$ 、 $M=4$ 和 $M=8$ 时 PAM 信号的星座图。

带通数字脉冲幅度调制 PAM 也称为幅移键控 (Amplitude-Shift Keying, ASK)。可以采用多种方法实现 k 个信息比特到 $M=2^k$ 个可能信号幅度的映射或分配, 优选的分配方法是相邻信号的幅度相差一个二进制数字, 如图 3.1 中的二进制数字所示, 这种映射称为格雷编码 (Gray Coding), 它在信号的解调中很重要, 因为噪声最可能引起的差错就是使得通信接收方错选相邻幅度作为发送信号幅度。在这种情况下, 格雷编码的 k 比特序列仅发生 1 个比特差错。

注意, 任何一对信号点之间的欧氏距离是

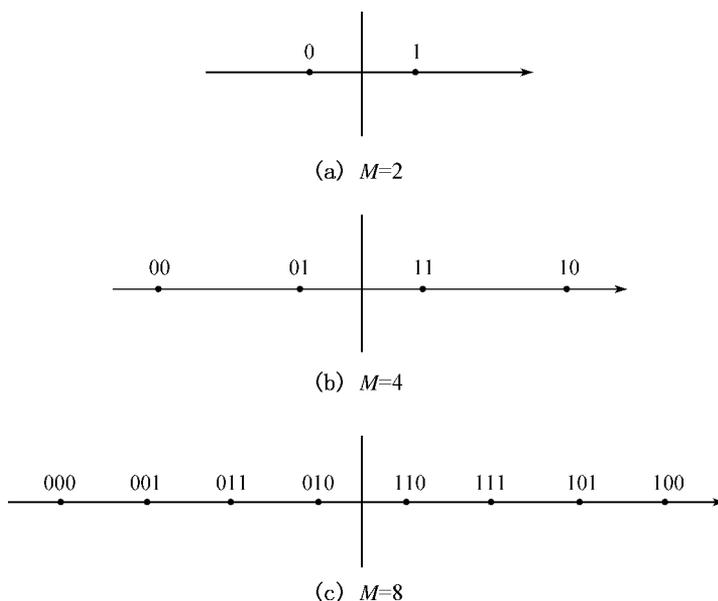


图 3.1 PAM 信号星座图

$$d_{mn} = \sqrt{\|s_m - s_n\|^2} = |A_m - A_n| \sqrt{\mathcal{E}_g} \text{ (基带 PAM)} \quad (3.2-16)$$

$$d_{mn} = \sqrt{\|s_m - s_n\|^2} = |A_m - A_n| \sqrt{\mathcal{E}_g/2} \text{ (带通 PAM)} \quad (3.2-17)$$

相邻信号点的 $|A_m - A_n| = 2$, 因此该星座的最小距离为

$$d_{\min} = 2\sqrt{\mathcal{E}_g} \text{ (基带 PAM)} \quad (3.2-18)$$

$$d_{\min} = \sqrt{2\mathcal{E}_g} \text{ (带通 PAM)} \quad (3.2-19)$$

M 元 PAM 系统的最小距离也可以用其能量 $\mathcal{E}_{\text{bavg}}$ 表示, 分别求解式(3.2-5)和式(3.2-9)可用 $\mathcal{E}_{\text{bavg}}$ 表示基带 \mathcal{E}_g 和带通 \mathcal{E}_g , 再代入式(3.2-18)和(3.2-19), 其结果表达式均为

$$d_{\min} = \sqrt{\frac{12 \log_2 M}{M^2 - 1} \mathcal{E}_{\text{bavg}}} \text{ (基带 PAM)} \quad (3.2-20)$$

$$d_{\min} = \sqrt{\frac{12 \log_2 M}{M^2 - 1} \mathcal{E}_{\text{bavg}}} \text{ (带通 PAM)} \quad (3.2-21)$$

式(3.2-6)表示的载波调制 PAM 信号是双边带(Double-Side Band, DSB)信号, 传输信号时要求两倍的等效低通信号带宽。也可以使用单边带(Single-Side Band, SSB)PAM 信号, 其表达式(下边带或上边带)为

$$s_m(t) = \text{Re}[A_m (g(t) \pm j\hat{g}(t)) e^{j2\pi f_c t}], m=1, 2, \dots, M \quad (3.2-22)$$

式中, $\hat{g}(t)$ 是 $g(t)$ 的希尔伯特变换。可以用傅里叶变换证明 SSB 信号的带宽是 DSB 信号的一半。

图 3.2(a)所示为 4 幅度电平基带 PAM 信号, 图 3.2(b)所示为带通 PAM 信号。



在 $M=2$ (或二进制信号) 时, PAM 信号具有特殊的性质, 即 $s_1(t) = -s_2(t)$, 这两个信号具有相等的能量且互相关系数为 -1 , 这样的信号称为双极性 (Antipodal) 信号, 也称为二进制双极性信号。

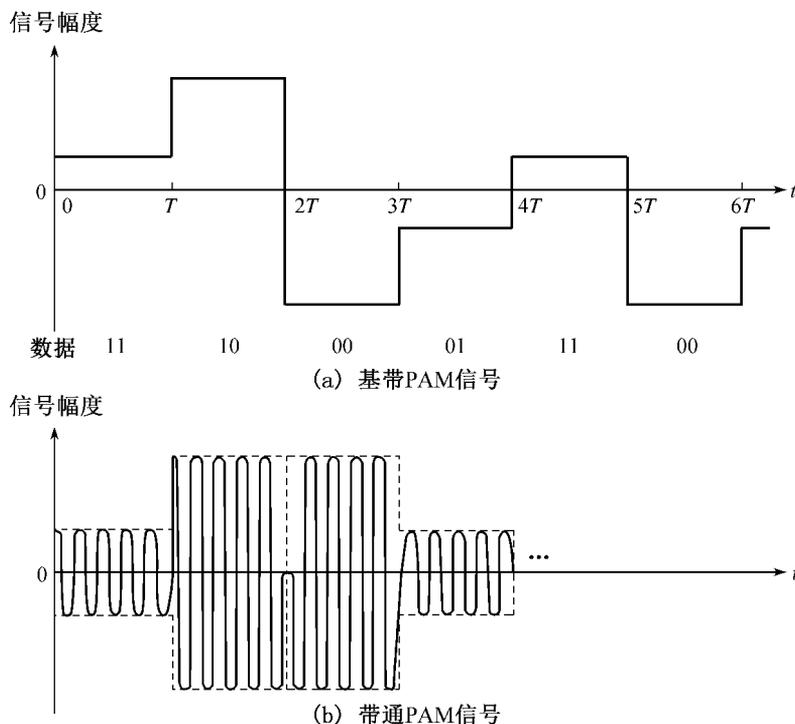


图 3.2 四进制基带与载波调制 PAM 信号

3.2.2 相移键控 (PSK)

相位调制是用调制波形的不同相位对应表示不同的数字符号。在数字相位调制中, M 个信号波形可表示为

$$\begin{aligned}
 s_m(t) &= \operatorname{Re} \left[g(t) e^{j\frac{2\pi(m-1)}{M}} e^{j2\pi f_c t} \right], m=1, 2, \dots, M \\
 &= g(t) \cos \left[2\pi f_c t + \frac{2\pi}{M}(m-1) \right] \\
 &= g(t) \cos \left[\frac{2\pi}{M}(m-1) \right] \cos 2\pi f_c t - g(t) \sin \left[\frac{2\pi}{M}(m-1) \right] \sin 2\pi f_c t
 \end{aligned} \tag{3.2-23}$$

式中, $g(t)$ 是脉冲波形, $\theta_m = \frac{2\pi(m-1)}{M}$ ($m=1, 2, \dots, M$) 是载波的 M 个可能的相位, 用于表征不同的发送符号。

1. PSK 信号的能量

数字相位调制通常称为相移键控 (Phase-Shift Keying, PSK)。注意, 这些信号波形具有相等的能量。由等效低通能量是带通信号能量的 2 倍可知



$$\mathcal{E}_{\text{avg}} = \mathcal{E}_m = \frac{1}{2} \mathcal{E}_g \quad (3.2-24)$$

所以

$$\mathcal{E}_{\text{bavg}} = \frac{\mathcal{E}_g}{2 \log_2 M} \quad (3.2-25)$$

2. PSK 信号的矢量表示

注意 $s_m(t)$ 的表达式(3.2-23)中 $g(t) \cos 2\pi f_c t$ 和 $g(t) \sin 2\pi f_c t$ 是正交的,因此,它包含两个基:

$$\phi_1(t) = \sqrt{\frac{2}{\mathcal{E}_g}} g(t) \cos 2\pi f_c t \quad (3.2-26)$$

$$\phi_2(t) = -\sqrt{\frac{2}{\mathcal{E}_g}} g(t) \sin 2\pi f_c t \quad (3.2-27)$$

$\phi_1(t)$ 和 $\phi_2(t)$ 可用于 $s_m(t)$ ($1 \leq m \leq M$) 的展开式,即

$$s_m(t) = \sqrt{\frac{\mathcal{E}_g}{2}} \cos \left[\frac{2\pi}{M}(m-1) \right] \phi_1(t) + \sqrt{\frac{\mathcal{E}_g}{2}} \sin \left[\frac{2\pi}{M}(m-1) \right] \phi_2(t) \quad (3.2-28)$$

因此,信号空间的维度 $N=2$,信号的矢量表达式为

$$\mathbf{s}_m = \left[\sqrt{\frac{\mathcal{E}_g}{2}} \cos \left(\frac{2\pi}{M}(m-1) \right), \sqrt{\frac{\mathcal{E}_g}{2}} \sin \left(\frac{2\pi}{M}(m-1) \right) \right], m=1, 2, \dots, M \quad (3.2-29)$$

图 3.3 所示为 BPSK ($M=2$)、QPSK ($M=4$) 和 8-PSK ($M=8$) 的信号星座图。注意, BPSK 相当于一维信号,与二进制 PAM 信号相同。这种信号传输方式是前面讨论的二进制双极性信号传输的特例。

与 PAM 的情况一样, k 个信息比特到 $M=2^k$ 个可能的相位的映射或分配可以用多种方法来实现,优选的分配方法是格雷编码,因此,由噪声引起的最大可能的差错是 k 比特符号中单个比特差错。在信号点之间的欧氏距离是

$$d_{mn} = \sqrt{\|\mathbf{s}_m - \mathbf{s}_n\|^2} = \sqrt{\mathcal{E}_g \left[1 - \cos \left(\frac{2\pi}{M}(m-n) \right) \right]} \quad (3.2-30)$$

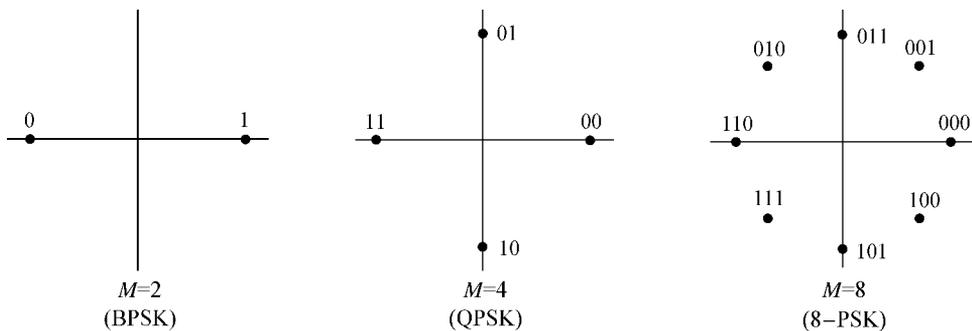


图 3.3 BPSK、QPSK 和 8-PSK 的信号星座图



最小距离为 $|m-n|=1$, 即

$$d_{\min} = \sqrt{\mathcal{E}_g \left(1 - \cos \frac{2\pi}{M}\right)} = \sqrt{2\mathcal{E}_g \sin^2 \frac{\pi}{M}} \quad (3.2-31)$$

求解式(3.2-25)可得 \mathcal{E}_g , 代入上式可得

$$d_{\min} = 2\sqrt{\left(\log_2 M \times \sin^2 \frac{\pi}{M}\right) \mathcal{E}_{\text{avg}}} \quad (3.2-32)$$

当 M 值很大时, $\sin \frac{\pi}{M} \approx \frac{\pi}{M}$, d_{\min} 近似为

$$d_{\min} \approx 2\sqrt{\frac{\pi^2 \log_2 M}{M^2} \mathcal{E}_{\text{avg}}} \quad (3.2-33)$$

$\pi/4$ -QPSK 是 4-PSK(QPSK)的一个变型,是通过在每一个符号间隔的载波相位中引入附加的 $\pi/4$ 相移而得到的,该相移可使符号同步变得容易些。

3.2.3 正交幅度调制(QAM)

PAM 与 PSK 分别通过改变载波的幅度或相位来传递信息。为了能同时利用幅度和相位两个维度,下面介绍正交幅度调制(Quadrature Amplitude Modulation, QAM)。

QAM 的基本原理是利用一对频率相同但相位正交的载波:余弦载波 $\cos(2\pi f_c t)$ 和正弦载波 $\sin(2\pi f_c t)$ 。第 m 种符号首先映射为复数 $A_m = A_{mi} + jA_{mq}$, 其中,实部 A_{mi} (同相分量, In-phase component)用于调制余弦载波的幅度;虚部 A_{mq} (正交分量, Quadrature component)用于调制正弦载波的幅度。最终,将这两路调制后的信号相加,形成 QAM 发送信号 $s_m(t)$:

$$\begin{aligned} s_m(t) &= \text{Re}[(A_{mi} + jA_{mq})g(t)e^{j2\pi f_c t}] \\ &= A_{mi}g(t)\cos 2\pi f_c t - A_{mq}g(t)\sin 2\pi f_c t, m=1, 2, \dots, M \end{aligned} \quad (3.2-34)$$

式中, A_{mi} 和 A_{mq} 是承载信息的正交载波的信号幅度, $g(t)$ 是脉冲信号。

由式(3.2-34)可以看出,类似于 PSK 的情况,由式(3.2-26)和式(3.2-27)确定的 $\phi_1(t)$ 和 $\phi_2(t)$ 可作为 QAM 信号展开式标准正交基。QAM 信号空间的维度 $N=2$, 利用该基可得

$$s_m(t) = A_{mi}\sqrt{\mathcal{E}_g/2}\phi_1(t) + A_{mq}\sqrt{\mathcal{E}_g/2}\phi_2(t) \quad (3.2-35)$$

由此得到的矢量表达式为

$$\mathbf{s}_m = (s_{m1}, s_{m2}) = (A_{mi}\sqrt{\mathcal{E}_g/2}, A_{mq}\sqrt{\mathcal{E}_g/2}) \quad (3.2-36)$$

和

$$\mathcal{E}_m = \|\mathbf{s}_m\|^2 = \frac{\mathcal{E}_g}{2}(A_{mi}^2 + A_{mq}^2) \quad (3.2-37)$$

任意一对信号矢量之间的欧氏距离是

$$d_{mn} = \sqrt{\|\mathbf{s}_m - \mathbf{s}_n\|^2} = \sqrt{\frac{\mathcal{E}_g}{2}[(A_{mi} - A_{ni})^2 + (A_{mq} - A_{nq})^2]} \quad (3.2-38)$$



在特殊情况下,即信号幅度取一组离散值 $\{(2m-1)E_g, m=1, 2, \dots, M\}$,信号空间图是矩形,如图 3.4 所示,在这种情况下,相邻两点的欧氏距离即最小距离,即

$$d_{\min} = \sqrt{2E_g} \quad (3.2-39)$$

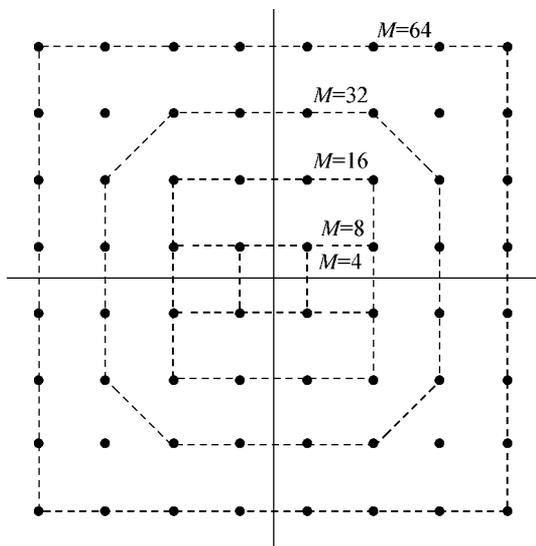


图 3.4 矩形 QAM 信号空间图

这与 PAM 的结果相同。在 $M=2^{2k_1}$ (k_1 是正整数) 的矩形星座特殊条件下,即 $M=4, 16, 64, 256, \dots$ 且在两个方向上的幅度为 $\pm 1, \pm 3, \dots, \pm(\sqrt{M-1})$,由式(3.2-37)可得

$$\mathcal{E}_{\text{avg}} = \frac{1}{M} \mathcal{E}_g \sum_{m=1}^{\sqrt{M}} \sum_{n=1}^{\sqrt{M}} (A_m^2 + A_n^2) = \frac{\mathcal{E}_g}{2M} \frac{2M(M-1)}{3} = \frac{M-1}{3} \mathcal{E}_g \quad (3.2-40)$$

由此可得到

$$\mathcal{E}_{\text{bavg}} = \frac{M-1}{3 \log_2 M} \mathcal{E}_g \quad (3.2-41)$$

利用式(3.2-39)可得到

$$d_{\min} = \sqrt{\frac{6 \log_2 M}{M-1} \mathcal{E}_{\text{bavg}}} \quad (3.2-42)$$

QAM 信号波形的另一种表示形式为

$$s_m(t) = \text{Re}[r_m e^{j\theta_m} e^{j2\pi f_c t}] = r_m \cos(2\pi f_c t + \theta_m) \quad (3.2-43)$$

式中, $r_m = \sqrt{A_{mi}^2 + A_{mq}^2}$, $\theta_m = \arctan(A_{mq}/A_{mi})$ 。该表达式表明, QAM 信号波形可以看成幅度(r_m)和相位(θ_m)的联合调制。事实上,可以选择 M_1 个电平的 PAM 和 M_2 个相位的 PSK 的任意组合来构成一个 $M=M_1 M_2$ 的组合 PAM-PSK 信号星座图。如果 $M_1=2^n$ 及 $M_2=2^m$, 则组合 PAM-PSK 信号星座图产生这样的结果:以符号速率 $R/(m+n)$ 同步传输, 每个符号所包含的 $m+n=\log_2 M_1 M_2$ 个二进制数字。

组合 PAM-PSK 信号空间图的例子如图 3.5 所示,其中 $M=8$ 及 $M=16$ 。

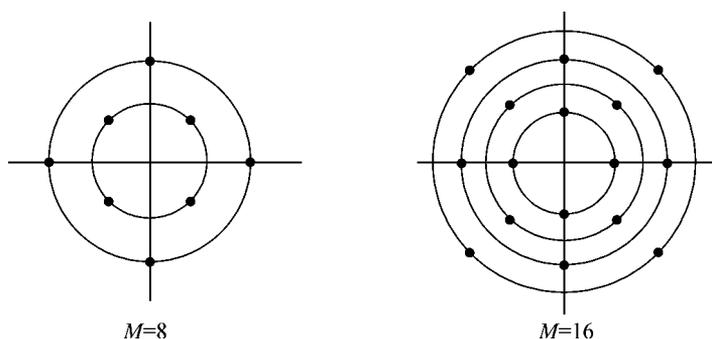


图 3.5 组合 PAM-PSK 星座图的例子

观察 PAM、PSK 和 QAM 的公式,注意到所有这些信号传输方式的通用形式为

$$s_m(t) = \text{Re}[A_m g(t) e^{j2\pi f_c t}], m=1, 2, \dots, M \quad (3.2-44)$$

式中, A_m 由信号传输方式确定, PAM 的 A_m 是实数, 一般等于 $\pm 1, \pm 3, \dots, \pm(M-1)$; M 元 PSK 的 A_m 是复数且等于 $e^{j\frac{2\pi}{M}(m-1)}$; QAM 的 A_m 也是复数, $A_m = A_{mi} + jA_{mq}$ 。从这个意义上讲, 可认为这三种信号传输方式属于同一种类型, PAM 和 PSK 可认为是 QAM 的特例。在 QAM 信号传输方式中, 幅度和相位都携带消息, 而 PAM 和 PSK 只是幅度或相位携带信息。还注意到, 在这些方式中信号空间的维度是相当低的 (PAM 为一维, PSK 和 QAM 为二维), 并且与星座的大小 M 无关。图 3.6 所示为这种一般类型信号传输方式调制器的结构, 其中 $\phi_1(t)$ 和 $\phi_2(t)$ 由式 (3.2-26) 和式 (3.2-27) 确定。注意, 调制器包含一个矢量映射器 (它将 M 个消息映射到 M 星座上), 其后是一个二维 (或 PAM 情况下为一维) 矢量到信号的映射器。

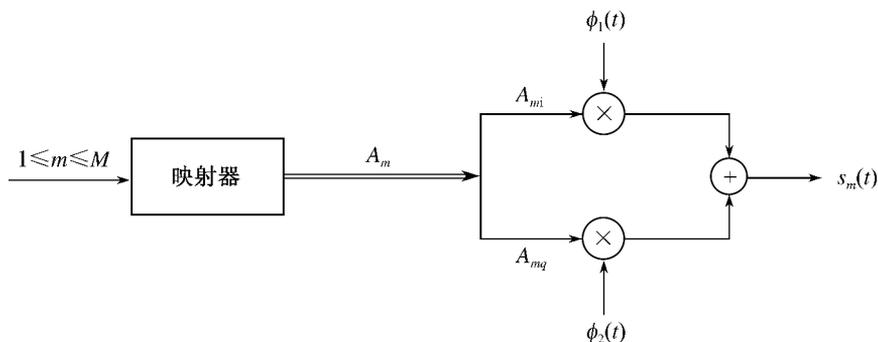


图 3.6 一般 QAM 调制器

3.2.4 正交信号调制

1. 频率调制 (FSK)

首先介绍正交信号传输, 正交信号定义为一个等能量的信号集 $s_m(t)$ ($1 \leq m \leq M$), 其中任意两个不同信号的内积为零, 即

$$\langle s_m(t), s_n(t) \rangle = 0, m \neq n \text{ 且 } 1 \leq m, n \leq M \quad (3.2-45)$$



由此定义可知

$$\langle s_m(t), s_n(t) \rangle = \begin{cases} \mathcal{E} & m=n \\ 0 & m \neq n \end{cases}, 1 \leq m, n \leq M \quad (3.2-46)$$

其中 \mathcal{E} 为信号能量。显然,信号之间是线性独立的。由

$$\phi_j(t) = \frac{s_j(t)}{\sqrt{\mathcal{E}}}, 1 \leq j \leq N \quad (3.2-47)$$

确定的标准正交集 $\{\phi_j(t), 1 \leq i \leq N\}$ 可作为 $\{s_m(t) (1 \leq m \leq M)\}$ 的正交基。信号的矢量表达式为

$$s_1 = (\sqrt{\mathcal{E}}, 0, 0, \dots, 0), s_2 = (0, \sqrt{\mathcal{E}}, 0, \dots, 0), \dots, s_M = (0, 0, \dots, 0, \sqrt{\mathcal{E}}) \quad (3.2-48)$$

从式(3.2-48)可以看出,当 $m \neq n$ 时,有

$$d_{mn} = \sqrt{2\mathcal{E}} \quad (3.2-49)$$

因此有

$$d_{\min} = \sqrt{2\mathcal{E}} \quad (3.2-50)$$

利用关系式

$$\mathcal{E}_b = \mathcal{E} / \log_2 M \quad (3.2-51)$$

最后可得到

$$d_{\min} = \sqrt{2\mathcal{E}_b \log_2 M} \quad (3.2-52)$$

本节我们考虑正交信号构成的一个特殊情况——频率不同的正交信号波形,它们可以表示为

$$\begin{aligned} s_m(t) &= \operatorname{Re}[s_{ml}(t) e^{j2\pi f_c t}], 1 \leq m \leq M, 0 \leq t \leq T \\ &= \sqrt{2\mathcal{E}/T} \cos(2\pi f_c t + 2\pi m \Delta f t) \end{aligned} \quad (3.2-53)$$

式中,

$$s_{ml}(t) = \sqrt{2\mathcal{E}/T} e^{j2\pi m \Delta f t}, 1 \leq m \leq M, 0 \leq t \leq T \quad (3.2-54)$$

引入系数 $\sqrt{2\mathcal{E}/T}$ 以保证每个信号的能量等于 \mathcal{E} 。这种形式的信号传输是将消息通过不同频率来传输的,称为频移键控(Frequency-Shift Keying, FSK)。注意 FSK 与 QAM 之间的主要差别(ASK 和 PSK 可认为是 QAM 的特殊情况):在 QAM 信号传输中,等效低通信号的形式为 $A_m g(t)$,其中 A_m 是复数。因此,两个不同 QAM 信号的两个等效低通信号之和仍然是 QAM 信号的等效低通信号的一般形式。在这个意义上,两个 QAM 信号之和是另一个 QAM 信号,因此称 ASK、PSK 和 QAM 为线性调制方式。另一方面,FSK 信号传输方式不满足这种性质,所以它属于非线性调制方式(Nonlinear Modulation Scheme)类型。

显然,对所有 $m \neq n$,正交信号集必须满足

$$\operatorname{Re} \left[\int_0^T s_{ml}(t) s_{nl}(t) dt \right] = 0 \quad (3.2-55)$$



对于由不同频率构成的信号集有

$$\begin{aligned}\langle s_{m_l}(t), s_{n_l}(t) \rangle &= \frac{2\mathcal{E}}{T} \int_0^T e^{j2\pi(m-n)\Delta f t} dt \\ &= \frac{2\mathcal{E} \sin[\pi T(m-n)\Delta f]}{\pi T(m-n)\Delta f} e^{j\pi T(m-n)\Delta f}\end{aligned}\quad (3.2-56)$$

以及

$$\begin{aligned}\operatorname{Re}[\langle s_{m_l}(t), s_{n_l}(t) \rangle] &= \frac{2\mathcal{E} \sin[\pi T(m-n)\Delta f]}{\pi T(m-n)\Delta f} \cos[\pi T(m-n)\Delta f] \\ &= \frac{2\mathcal{E} \sin[2\pi T(m-n)\Delta f]}{2\pi T(m-n)\Delta f} = 2\mathcal{E} \operatorname{sinc}[2\pi T(m-n)\Delta f]\end{aligned}\quad (3.2-57)$$

由式(3.2-57)可见,对所有 $m \neq n$, 当且仅当 $\operatorname{sinc}[2\pi T(m-n)\Delta f] = 0$ 时, $s_m(t)$ 与 $s_n(t)$ 是正交的。所以当 $\Delta f = \frac{k}{2T}$ (k 为正整数) 时, 此信号集才是正交的。 $\Delta f = \frac{1}{2T}$ 是保证 $\langle s_{m_l}(t), s_{n_l}(t) \rangle = 0$ 的最小频率间隔, 从而保证基带以及带通(频率调制)信号的正交性。

2. 正交码调制

M 进制的正交信号波形集也可以由形式为

$$\mathbf{c}_m = [c_{m1}, c_{m2}, \dots, c_{mN}], \quad m = 1, 2, \dots, M \quad (3.2-58)$$

的 M 个二进制码字集来生成, 式中, 对所有 m 和 j 均有 $c_{mj} = 0$ 或 1。码字集的每一个分量都可映射成基本的二进制 PSK 波形, 即

$$\begin{aligned}c_{mj} = 1 &\Rightarrow \sqrt{2\mathcal{E}_c T_c} \cos 2\pi f_c t, \quad 0 \leq t \leq T_c \\ c_{mj} = 0 &\Rightarrow -\sqrt{2\mathcal{E}_c T_c} \cos 2\pi f_c t, \quad 0 \leq t \leq T_c\end{aligned}\quad (3.2-59)$$

式中, $T_c = T/N$ 且 $\mathcal{E}_c = \mathcal{E}/N$ 。因此, M 个码字集 $\{c_m\}$ 可映射成一个 M 个波形集 $\{s_m(t)\}$, 该集合中的波形 $s_m(t)$ 可以表示成矢量形式, 即

$$\mathbf{s}_m = [s_{m1}, s_{m2}, \dots, s_{mN}], \quad m = 1, 2, \dots, M \quad (3.2-60)$$

式中, 对于所有 m 和 j , 有 $s_{mj} = \pm\sqrt{\mathcal{E}/N}$, N 称为码分组长度, 并且是 M 个波形的维数。

每个码字有 N 个码片, 对应有 2^N 个可能的波形存在, 它们可由 2^N 个可能的二进制码字生成。其中只有 M ($M < 2^N$) 种码字组合可以保证相互之间是正交的, 可以选择这 M 个信号波形子集来传输信息。 2^N 个可能的信号点对应于以原点为中心的超立方体的顶点。图 3.7 所示为由二进制码生成的信号的信号空间。此信号空间中每一个信号点对应的波形都具有能量 \mathcal{E} , 且任意相邻信号点相差 1 个码字, 因此, 相邻信号点对应的波形的互相关系数为

$$\rho = \frac{\mathcal{E}(1-2/N)}{\mathcal{E}} = \frac{N-2}{N} \quad (3.2-61)$$

以及相应的距离为

$$d_{\min} = \sqrt{2\mathcal{E}(1-\rho)} = \sqrt{4\mathcal{E}/N} \quad (3.2-62)$$



此处计算的距离是相邻信号点的距离,并不是正交码字集的最小距离,因为正交性的要求使得相邻信号点不一定存在于正交码字集中。因此,正交码字集的最小距离 $d_{\min} \geq d_{\min}$ 。

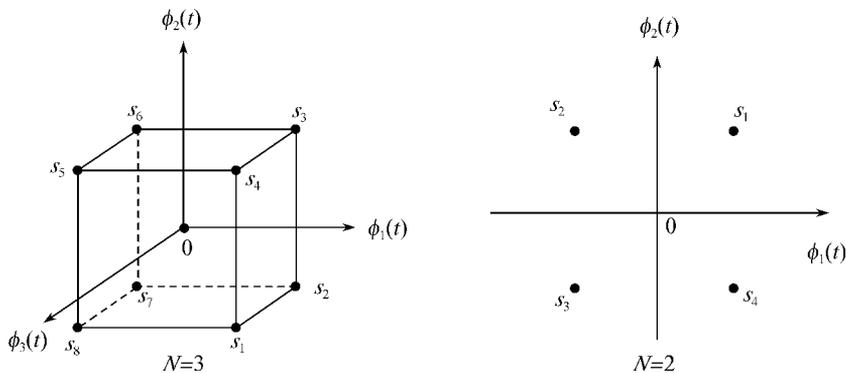


图 3.7 由二进制正交码生成的信号的信号空间图

3.3 有记忆调制

3.3.1 连续相位频率调制(CPFSK)

本节将研究连续相位频率调制(CPFSK),在这种调制中,信号相位的变化是连续的。这个约束条件要求信号波形不仅与当前发送符号有关,而且与之前符号有关,因而需要有记忆的相位或频率调制器。

由式(3.2-51)可见,常规 FSK 信号是由载波频移产生的,频移量 $m\Delta f (1 \leq m \leq M)$ 反映所要发送的数字信息。这种类型的 FSK 信号是无记忆的。从一个频率到另一个频率的切换的实现方法是使用 $M=2^k$ 个调谐到期望频率的振荡器,再从 M 个频率中选择一个频率,选择的依据是在信号间隔时间 $T=k/R$ 秒内要发送的特定的 k 比特符号。然而,在连续的信号传输时间间隔中,这种从一个频点的振荡器输出到另一个频点的突发式切换,会造成在信号主要频段之外有比较大的谱旁瓣(Side Lobe),因此,用这种方式传输信号需要较宽的频带,频带利用率低。为了避免使用具有较大谱旁瓣的信号,携带信息的信号频率调制通常采用单一的载波,其相位是连续变化的,所得到的频率调制信号是相位连续的,因此称为连续相位 FSK(Continuous-Phase Frequency-Shift Keying, CPFSK)。这种类型的 FSK 是有记忆的,因为载波相位是连续的。为了表示 CPFSK 信号,我们从 PAM 信号

$$d(t) = \sum_n I_n g(t - nT) \quad (3.3-1)$$

开始,式中 $\{I_n\}$ 表示幅度序列,它是由信息序列 $\{a_n\}$ 中的 k 比特二进制数字组映射到幅度电平 $\pm 1, \pm 3, \dots, \pm(M-1)$ 得到的。而 $g(t)$ 是一个幅度为 $1/(2T)$ 且持续时间为 T 秒的矩形脉冲。信号 $d(t)$ 用来对载波进行频率调制,从而等效低通波形 $\nu(t)$ 可表示为

$$\nu(t) = \sqrt{2\mathcal{E}/T} e^{j[4\pi T f_d \int_{-\infty}^t d(\tau) d\tau + \phi_0]} \quad (3.3-2)$$



式中, f_d 是峰值频率偏移(Peak Frequency Deviation), ϕ_0 是载波的初始相位。对应于式(3.3-2)的载波调制信号可以表示为

$$s(t) = \sqrt{2\mathcal{E}/T} \cos[2\pi f_c t + \phi(t; I) + \phi_0] \quad (3.3-3)$$

式中, $\phi(t; I)$ 表示载波的时变相位, 定义为

$$\phi(t; I) = 4\pi T f_d \int_{-\infty}^t d(\tau) d\tau = 4\pi T f_d \int_{-\infty}^t \left[\sum_n I_n g(\tau - nT) \right] d\tau \quad (3.3-4)$$

注意, 虽然 $d(t)$ 不具有连续性, 但 $d(t)$ 的积分是连续的, 因此, 可得到一个连续相位信号。在 $nT \leq t \leq (n+1)T$ 间隔内的载波相位由式(3.3-4)的积分确定。因此,

$$\phi(t; I) = 2\pi f_d T \sum_{k=-\infty}^{n-1} I_k + 4\pi f_d T q(t - nT) I_n = \theta_n + 2\pi h I_n q(t - nT) \quad (3.3-5)$$

式中, h , θ_n 及 $q(t)$ 的定义为

$$h = 2f_d T \quad (3.3-6)$$

$$\theta_n = \pi h \sum_{k=-\infty}^{n-1} I_k \quad (3.3-7)$$

$$q(t) = \begin{cases} 0 & t < 0 \\ t/2T & 0 \leq t \leq T \\ 1/2 & t > T \end{cases} \quad (3.3-8)$$

可以看到, θ_n 表示直到 $(n-1)T$ 时的所有符号的累积(记忆)值, 参数 h 称为调制指数(Modulation Index)。

3.3.2 连续相位调制(CPM)

当采用式(3.3-5)的形式表示时, CPFSK 变成一般类型连续相位调制(Continuous-Phase Modulation, CPM)信号的一个特例, CPM 的载波相位是

$$\phi(t; I) = 2\pi \sum_{k=-\infty}^n I_k h_k q(t - kT), nT \leq t \leq (n+1)T \quad (3.3-9)$$

式中, $\{I_k\}$ 是由符号表 $\pm 1, \pm 3, \dots, \pm(M-1)$ 中选出的 M 元信息符号序列, $\{h_k\}$ 是调制指数序列, $q(t)$ 是某个脉冲积分波形。当对于所有的 k 都有 $h_k = h$ 时, 调制指数对所有符号都是固定的, 此时称为单模 CPM。当多种不同调制指数随着符号而变化时, 该信号称为多重 h CPM。在这种情况下, $\{h_k\}$ 以循环方式在调制指数集中变化。波形 $q(t)$ 一般可以表示成某个脉冲 $g(t)$ 的积分, 即

$$q(t) = \int_0^t g(\tau) d\tau \quad (3.3-10)$$

如果对于 $t > T$ 有 $g(t) = 0$, 则 CPM 信号称为全响应 CPM(Full-Response CPM)。如果对于 $t > 0$ 有 $g(t) \neq 0$, 则已调信号称为部分响应 CPM(Partial-Response CPM)。图 3.8 说明了 $g(t)$ 的几种脉冲形状及其相应的 $q(t)$ 。显而易见, 通过选择不同的脉冲形状 $g(t)$ 以及改变调制指数和符号数目 M , 就可以产生不同 CPM 信号。注意, CPM 信号的记忆性是通过相位的连续性引入的。



表 3.1 列出了三种常用的 CPM 脉冲形状。LREC 表示持续时间为 LT 的矩形脉冲, 这里 L 为整数, 图 3.8(a) 所示为 $L=1$ 的 CPFSK 脉冲, 图 3.8(c) 所示为 $L=2$ 的矩形脉冲, LRC 表示持续时间为 LT 的升余弦脉冲, 图 3.8(b) 和图 3.8(d) 分别为 $L=1$ 和 $L=2$ 的升余弦脉冲。表 3.1 中的第 3 种脉冲称为高斯最小移频键控 (Gaussian Minimum Shift Keying, GMSK) 脉冲, B 表示高斯脉冲的 -3 dB 带宽。图 3.8(e) 所示为时间带宽乘积 BT 范围为 0.1 到 1 的一组 GMSK 脉冲。可以看出, 当脉冲带宽减小时脉冲持续时间增加。在实际应用中, 通常将脉冲持续时间截短到某特定的固定长度。例如欧洲在全球移动通信系统 (GSM) 中采用了 $BT=0.3$ 的 GMSK 调制方式。从图 3.8(e) 我们可以看到, 当 $BT=0.3$ 时, GMSK 脉冲可在 $|t|=1.5T$ 处截短, 而使 $t > 1.5T$ 的误差比较小。

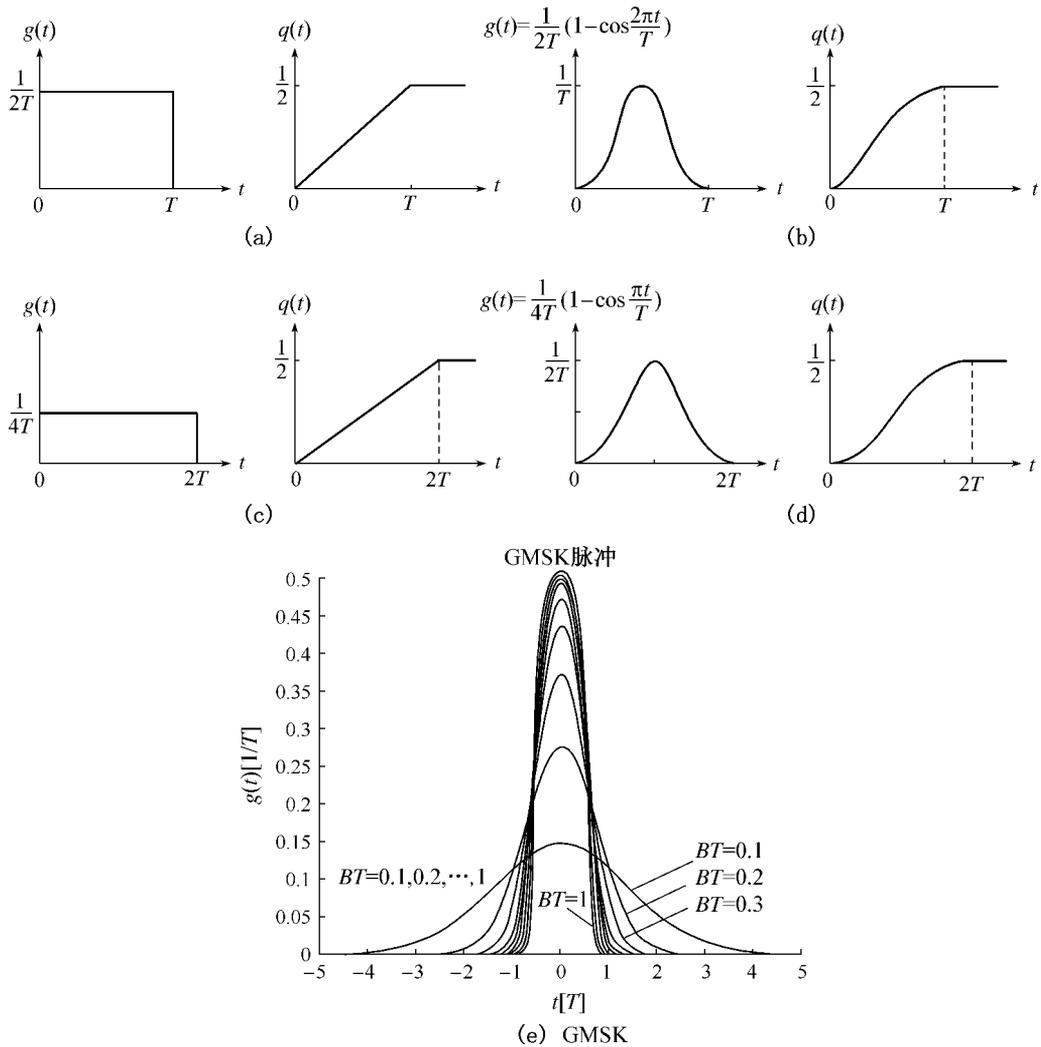


图 3.8 $g(t)$ 的几种脉冲形状和相应的 $q(t)$



表 3.1 三种常用的 CPM 脉冲形状

LREC	$g(t) = \begin{cases} \frac{1}{2LT} & 0 \leq t \leq LT \\ 0 & \text{其他} \end{cases}$
LRC	$g(t) = \begin{cases} \frac{1}{2LT} \left(1 - \cos \frac{2\pi t}{LT}\right) & 0 \leq t \leq LT \\ 0 & \text{其他} \end{cases}$
GMSK	$g(t) = \frac{Q\left(2\pi B\left(t - \frac{T}{2}\right)\right) - Q\left(2\pi B\left(t + \frac{T}{2}\right)\right)}{\sqrt{\ln 2}}$

画出由信息序列 $\{I_k\}$ 所有可能值所生成的一组相位轨迹 $\phi(t; I)$ 是很有用的, 例如, 在具有二进制符号 $I_k = \pm 1$ 的 CPFSK 情况下, 在 $t=0$ 起始的一组相位轨迹如图 3.9 所示。作为比较, 图 3.10 画出了 4 元 CPFSK 的相位轨迹。

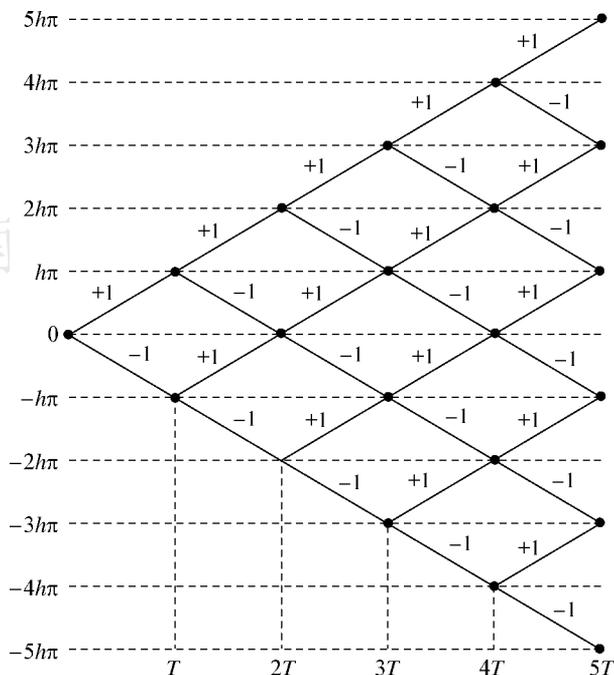


图 3.9 二进制 CPFSK 的相位轨迹

这些相位图称为相位树(Phase Tree), 可以看到, CPFSK 的相位树是分段线性的, 这是因为脉冲 $g(t)$ 是矩形的。较平滑的相位轨迹和相位树可以通过使用不包含跃变的脉冲来获得, 例如, 使用升余弦脉冲。图 3.11 所示为使用长度为 $3T$ 升余弦脉冲的部分响应 CPM 的相位轨迹, 是由序列 $(1, -1, -1, -1, 1, 1, -1, 1)$ 产生的。为了便于比较, 图 3.11 也示出了二进制 CPFSK 的相位轨迹。

在这些图中, 相位树随时间而增长, 但载波相位仅仅在 $0 \sim 2\pi$ (或等价于 $-\pi \sim \pi$) 范围内是唯一的。当相位轨迹以模 2π 画出时, 即在 $(-\pi, \pi)$ 范围内, 则相位树折叠到一个称为相

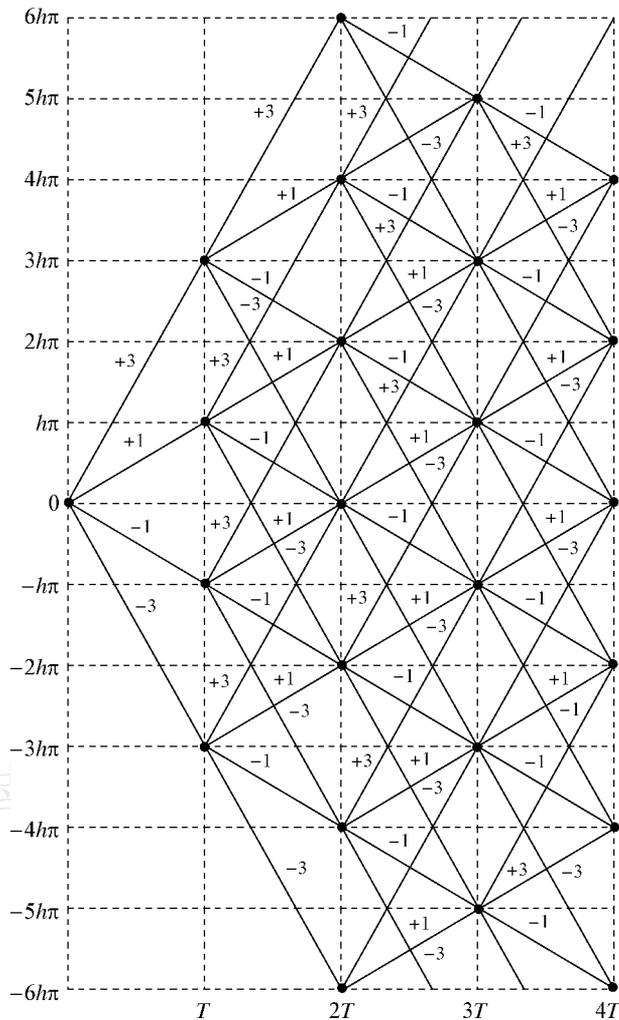


图 3.10 4元 CPFSK 相位轨迹

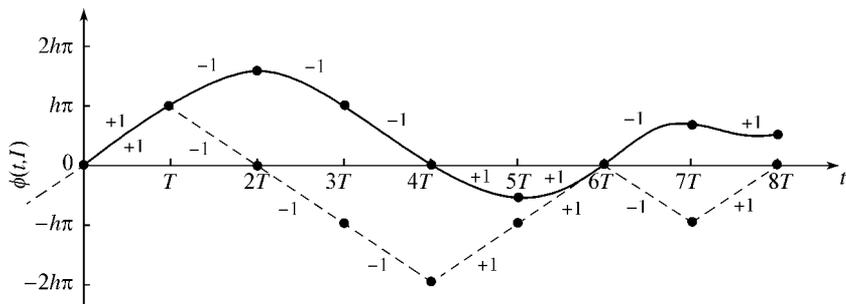


图 3.11 二进制 CPFSK 的相位轨迹(虚线)和使用长度为 $3T$ 升余弦脉冲的二进制部分响应 CPM 的相位轨迹(实线)

位网格(Phase Trellis)的结构中。为了全面地考察相位网格图,画出了两个正交分量 $x_i(t; I) = \cos \phi(t; I)$ 和 $x_q(t; I) = \sin \phi(t; I)$ 作为时间的函数,从而作出一条三维曲线,该



曲线的两正交分量 x_i 和 x_q 出现在单位半径的圆柱面上。例如,图 3.12 所示为相位网格或相位圆柱,它是由二进制调制、调制指数 $h=1/2$ 以及长度为 $3T$ 的升余弦脉冲得到的。

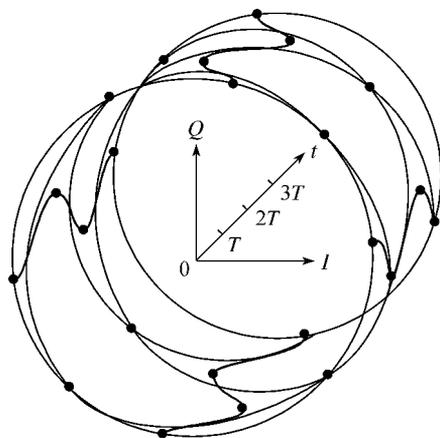


图 3.12 具有 $h=1/2$ 及长度为 $3T$ 升余弦脉冲的二进制 CPM 的相位圆柱

较简单的相位轨迹表示法可以通过仅显示在 $t=nT$ 时刻的信号相位终值来获得。在这种情况下,限制 CPM 信号的调制指数为有理数。特别地,假定 $h=m/p$,其中 m 与 p 是互质整数,从而在 $t=nT$ 时且 m 为偶数时的全响应 CPM 信号具有终值相位状态(Terminal Phase State)。

当 m 为偶数时:

$$\Theta_s = \left\{ 0, \frac{\pi m}{p}, \frac{2\pi m}{p}, \dots, \frac{(p-1)\pi m}{p} \right\} \quad (3.3-11)$$

当 m 为奇数时:

$$\Theta_s = \left\{ 0, \frac{\pi m}{p}, \frac{2\pi m}{p}, \dots, \frac{(2p-1)\pi m}{p} \right\} \quad (3.3-12)$$

式中, Θ_s 表示终值相位状态。因此,当 m 为偶数时,有 p 个终值相位状态;当 m 为奇数时,有 $2p$ 个终值相位状态。另一方面,当脉冲形状延伸 L 个符号间隔(部分响应 CPM)时,终值相位状态数可增至最大值 S_t ,这里

$$S_t = \begin{cases} pM^{L-1} & m \text{ 为偶数} \\ 2pM^{L-1} & m \text{ 为奇数} \end{cases} \quad (3.3-13)$$

式中, M 为符号表大小。例如,具有 $h=1/2$ 的二进制 CPFSK 信号(全响应、矩形脉冲)具有 $S_t=4$ 个终值相位状态。图 3.13 说明了该信号的状态网格(State Trellis)。应强调指出,从一个状态到另一个状态的相位转移并不是真正的相位轨迹,它们表示在 $t=nT$ 时刻终值相位状态的相位转移。

状态网格的另一种表示法是状态图(State Diagram),它也说明了在 $t=nT$ 时刻的状态转移。这是 CPM 信号特征一种更紧凑、简洁的表示法。这种状态图仅显示可能的终值相位状态及其相位转移,时间并没有作为变量直接出现。例如,具有 $h=1/2$ 的 CPFSK 信号的状态图如图 3.14 所示。

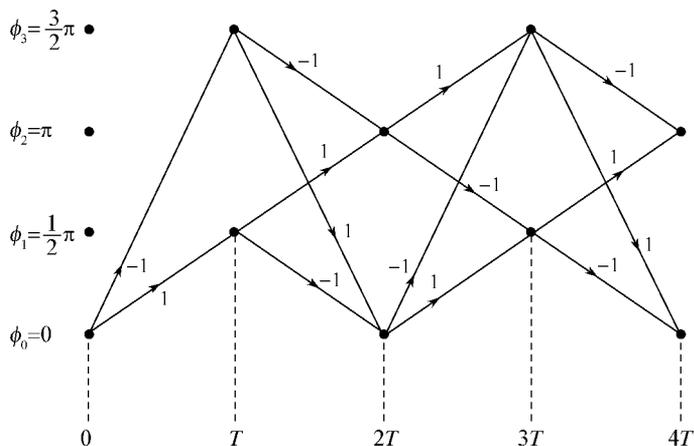


图 3.13 具有 $h=1/2$ 的二进制 CPFSK 的状态网格

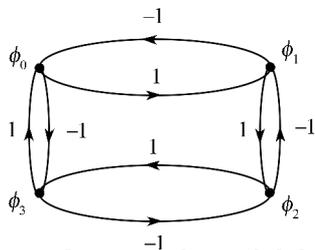


图 3.14 具有 $h=1/2$ 的 CPFSK 的状态图

3.3.3 最小频移键控 (MSK)

最小频移键控(MSK)是 2FSK 的改进。2FSK 信号相邻码元波形的相位可能不连续,在通过带通信道后会使得信号波形包络产生较大起伏。2FSK 信号的两种码元波形也不一定严格正交,若两种码元波形互相正交,其误码性能将更好。MSK 信号是一种包络恒定、相位连续、带宽最小且严格正交的 2FSK 信号。

MSK 是二进制 CPFSK 和 CPM 的一个特例,其调制指数 $h=1/2$ 。在 $nT \leq t \leq (n+1)T$ 间隔中的载波相位是

$$\begin{aligned} \phi(t; I) &= \frac{1}{2} \pi \sum_{k=-\infty}^{n-1} I_k + \pi I_n q(t - nT) \\ &= \theta_n + \frac{1}{2} \pi I_n \left(\frac{t - nT}{T} \right), nT \leq t \leq (n+1)T \end{aligned} \tag{3.3-14}$$

已调载波信号是

$$\begin{aligned} s(t) &= A \cos \left[2\pi f_c t + \theta_n + \frac{1}{2} \pi I_n \left(\frac{t - nT}{T} \right) \right] \\ &= A \cos \left[2\pi \left(f_c + \frac{1}{4T} I_n \right) t - \frac{1}{2} n \pi I_n + \theta_n \right], nT \leq t \leq (n+1)T \end{aligned} \tag{3.3-15}$$



式(3.3-15)表明:二进制 CPFSK 信号可以表示成在 $nT \leq t \leq (n+1)T$ 间隔中具有两个频率之一的正弦波。如果定义这些频率为

$$f_1 = f_c - \frac{1}{4T}, f_2 = f_c + \frac{1}{4T} \quad (3.3-16)$$

那么由式(3.3-15)确定的二进制 CPFSK 信号可以写成

$$s_i(t) = A \cos \left[2\pi f_i t + \theta_n + \frac{1}{2} n\pi (-1)^{i-1} \right], i=1,2 \quad (3.3-17)$$

它表示频率间隔为 $\Delta f = f_2 - f_1 = 1/(2T)$ 的 FSK 信号。由公式(3.2-55)可知,为了确保在长度为 T 的信号传输间隔上信号 $s_1(t)$ 和 $s_2(t)$ 的正交性,最小频率间隔 $\Delta f = 1/2T$ 是必要的。这就解释了为什么具有 $h = 1/2$ 的二进制 CPFSK 称为最小移频键控。在第 n 个信号传输间隔结束时,信号的相位值被设计为与第 $n+1$ 个间隔开始时的相位值一致,从而确保了相邻间隔之间的相位连续性。

3.3.4 高斯最小频移键控(GMSK)

作为 CPFSK 的一个特例,MSK 的相位是连续的,但其频率有跳变。如果平滑这些频率跳变,则频谱将更紧凑。可以在进行 MSK 调制前将矩形信号脉冲先通过一个高斯型的低通滤波器。这样的体制称为高斯最小频移键控(Gaussian MSK, GMSK)。此高斯型低通滤波器的频率特性表示式为

$$H(f) = \exp \left[- \left(\frac{\ln 2}{2} \right) (f/B)^2 \right] \quad (3.3-18)$$

式中, B 为滤波器的 3 dB 带宽。

将式(3.3-18)作逆傅里叶变换,得到此滤波器的冲激响应 $h(t)$ 为

$$h(t) = \frac{\sqrt{\pi}}{\alpha} \exp \left[- \left(\frac{\pi}{\alpha} t \right)^2 \right] \quad (3.3-19)$$

式中, $\alpha = \sqrt{\frac{\ln 2}{2}} \frac{1}{B}$ 。由于 $h(t)$ 具有高斯函数特性,故称为高斯型滤波器。

在 GSM 制的蜂窝网中采用 $BT = 0.3$ 的 GMSK 调制,即滤波器的 3 dB 带宽 B 等于码元速率的 0.3 倍,这是为了得到更大的用户容量,因为在那里对带外辐射的要求非常严格。GMSK 体制的缺点是有码间串扰。 BT 值越小,码间串扰越大。

3.4 扩频调制

扩频调制通过将信号的频谱扩展到远超原始带宽的范围,以显著提升抗干扰能力和通信安全性,有效抵御窄带干扰或其他有意干扰。

扩频技术主要分为直接序列扩频(Direct Sequence Spread Spectrum, DSSS)和跳频扩频(Frequency Hopping Spread Spectrum, FHSS)两种实现方式。直接序列扩频是将待传



输的数据序列与速率远高于数据速率的伪随机(PN)序列进行逐比特相乘,从而将信号能量“摊薄”到更宽的频谱上。在接收端,利用与发送端完全相同的伪随机序列对接收信号进行解扩,恢复出原始窄带数据信号,同时由于干扰未能与PN序列同步,相当于噪声被展宽、功率密度下降,从而实现抗干扰。跳频扩频则是通过发送端和接收端共享的跳频图案快速切换载波频率,使信号瞬时停留在每个频道的的时间极短,从而“躲避”固定频点的干扰。上述两种扩频方式均可有效抑制窄带干扰,并支持多址通信,如码分多址(Code Division Multiple Access,CDMA)。扩频调制技术在无线通信、卫星通信和军事通信等领域得到广泛应用,既保障了通信链路的可靠性,又提高了系统的安全性和抗截获能力。

3.4.1 直接序列扩频(DSSS)

1. 香农定理

在通信系统中,用于描述在带宽受限且存在噪声的信道上信息传输速率上限的香农-哈特利(Shannon-Hartley)定理表述如下:

$$C=W\log_2(1+S/N) \quad (3.4-1)$$

式中, C 为通信系统的信道容量(bit/s), W 为通信系统的信道带宽, S 为通信系统的发送信号的信号功率, N 为通信系统中信道的噪声功率。

当信道容量 C 一定时,增大带宽 W 可在降低信噪比 S/N 要求的同时保持相同的容量,这正是扩频通信“以牺牲带宽换取抗干扰能力”的理论基础。香农定理两边同除以带宽 W 可得:

$$C/W=1.44\ln(1+S/N) \quad (3.4-2)$$

当 $S/N \ll 1$ 时,对公式(3.4-2)求极限可得:

$$C/W=1.44 \times S/N \quad (3.4-3)$$

结合上式可以看出,在系统当前给定信噪比的前提下,可以通过用牺牲带宽的手段来保证较高的抗干扰能力。

2. 直接扩频通信模型

直接序列扩频(DSSS)通过在发送端用高速率的伪随机扩频码对原始窄带信号进行“摊薄”扩展,将其频谱展宽;在接收端再用同一扩频码对接收信号进行解扩,恢复出原始数据信号。图3.15为BPSK直接序列扩频系统的基本框图。

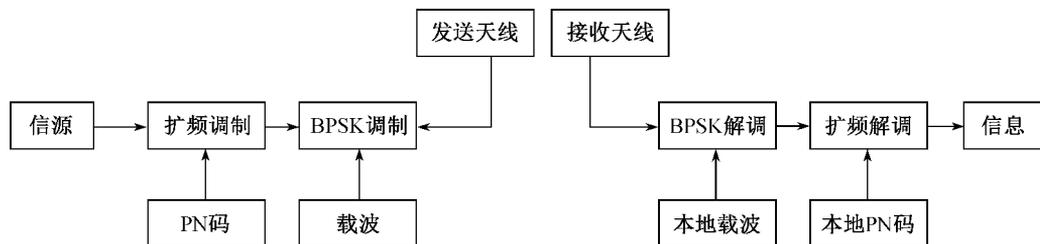


图 3.15 BPSK 直接扩频通信模型框图



扩频调制将码率远高于数据信号的扩频码序列 $c(t)$ 与输入的数字信号 $a(t)$ 逐比特相乘(或模二叠加),以实现频谱扩展。

设扩频码序列 $c(t)$ 的传输速率为 R_c ,扩频码的扩频倍数为 N ,输入的原始信息信号 $a(t)$ 的传输速率为 R ,那么 $R_c/R \gg 1$ 。令原始信息信号的码元宽度 T 为:

$$T=1/R \quad (3.4-4)$$

则输入信号可表示为:

$$a(t) = \sum a_n g_a(t - nT) \quad (3.4-5)$$

其中, a_n 为信息码, $g_a(t)$ 为门函数。

扩频码序列 $c(t)$ 可预先生成并在收发两端完全同步,表示为:

$$c(t) = \sum c_n g_c(t - nT_c) \quad (3.4-6)$$

其中, c_n 为扩频码码元, $T_c = 1/R_c$ 。

将此扩频码与码率较低的原始信号 $a(t)$ 相乘,即可得到频谱扩展后的信号:

$$d(t) = a(t)c(t) = \sum d_n g_c(t - nT_c) \quad (3.4-7)$$

其中, d_n 是 $d(t)$ 在第 n 个扩频码码片周期内的幅度值。

对扩展后的序列进行载波调制可得到:

$$s(t) = d(t)\cos(2\pi f_c t) \quad (3.4-8)$$

其中, f_c 为载波频率。调制后的信号被发往接收端,经过信道后,接收端的接收信号可表示为:

$$r(t) = s(t) + n(t) \quad (3.4-9)$$

其中, $n(t)$ 为噪声和干扰。为将信号频带恢复至原始宽度,利用同一扩频码序列与接收信号相乘进行解扩,然后采用本地载波解调后可得到:

$$\begin{aligned} & r(t)c(t)\cos(2\pi f_c t) \\ &= d(t)\cos(2\pi f_c t)c(t)\cos(2\pi f_c t) + n(t)c(t)\cos(2\pi f_c t) \quad (3.4-10) \\ &= a(t)c^2(t)\cos^2(2\pi f_c t) + n(t)c(t)\cos(2\pi f_c t) \\ &= a(t)\cos^2(2\pi f_c t) + n(t)c(t)\cos(2\pi f_c t) \end{aligned}$$

一般扩频码信号 $c(t)$ 采用双极性信号,故 $c^2(t)$ 为 1。因此由上式第一项可以看出, $d(t)$ 已恢复出原始信号 $a(t)$ 。而噪声与干扰在与 $c(t)$ 相乘后,其功率谱被展宽,在原始数据信号带宽内的功率密度大幅下降,从而提高了信噪比。

综上所述,直接序列扩频通信系统在发送端用高码率扩频码 $c(t)$ 扩展原始信号 $a(t)$ 的频谱;在接收端用相同扩频码解扩,先恢复出扩频后的基带信号 $d(t)$,继而还原成窄带信号 $a(t)$ 。扩频后信号带宽通常是原带宽的 N 倍,其中 N 为扩频倍数,对应功率谱密度下降到原来的 $1/N$,而不影响数据信号的透明传输。

3.4.2 跳频扩频(FHSS)

频率跳变扩频系统尽管和直接序列扩频系统一样,采用很宽的频带来传输窄带的信息



信号,但由于其工作原理和直接序列系统完全不同,因而其传输信号产生的方法、频谱结构等也与直接序列系统完全不同。

1. 频率跳变系统的物理概念

频率跳变就是用扩频码序列构成跳频指令来控制频率合成器输出信号的频率,使传输信号的载波频率按照某种特定算法发生的重复变化。在常规通信系统中,2FSK 调制方式只有 2 个频率,分别代表信息信号的“1”与“0”。而频率跳变系统则要求提供几百个、几千个甚至几万个离散的频率供随机选取。传输信号载波的频率由扩频码序列构成的跳频指令来随机选取。

在接收端,为了对输入信号进行解扩,需要有与发送端相同的跳频指令去控制本地的频率合成器,使其输出的频率跳变信号能在混频器中与接收到的频率跳变信号变换出一个固定的中频信号来,经中频放大器放大及带通滤波器滤波后,得到一个受信息信号调制的中频信号,然后送到信息解调器恢复出原来的信息信号。接收机的跳频指令由本地参考扩频码序列发生器与指令译码器等构成。频率跳变系统的简化原理方框图见图 3.16。

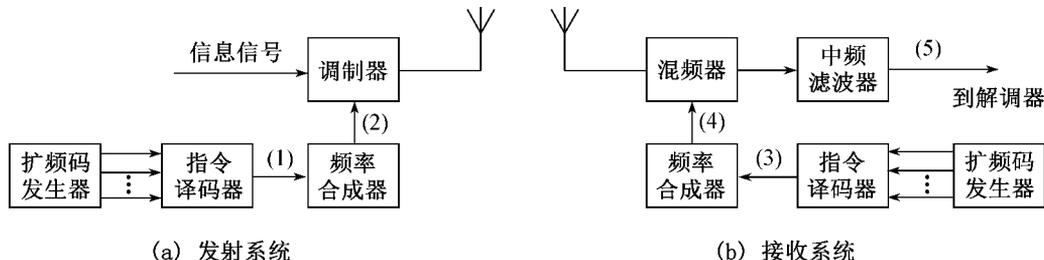


图 3.16 频率跳变系统的原理方框图

频率跳变系统的工作过程大致为:发送端扩频码序列(通常用其状态序列)发生器输出的跳频指令(跳频图案),去控制频率合成器输出信号的频率,使其输出信号的频率随机地跳跃变化,如图 3.17(a)所示。频率合成器输出的信号就是发射机的本振信号,被信息信号调制后作为频率跳变信号输出。

从时间-频率的对应关系来看,频率跳变信号的载波频率是由时频矩阵组成的,每个频率的持续时间为 T_h ,并按照跳频指令的规定在时频矩阵内跳变,如图 3.17(b)所示。图 3.18(a)为送入发送端频率合成器的跳频指令波形示意图;图 3.18(b)为发送端频率合成器输出的载波波形示意图;图 3.18(c)为送入接收端频率合成器的跳频指令波形示意图,该跳频指令由接收端的扩频码发生器产生;图 3.18(d)为接收端频率合成器输出的参考载波波形示意

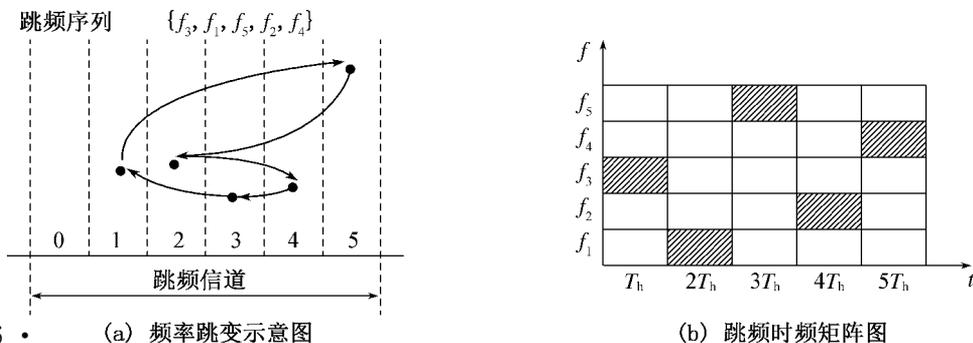


图 3.17 跳频时频矩阵图

图;图 3.18(e)为解跳后送入解调器的中频波形示意图。图中没有考虑信息信号,即跳变的载波未受信息信号的调制。

扩频码的传输速率为 R_c ,为了和信息码的传输速率 R 区别,扩频码的传输速率通常称为码片速率或切普(chirp)速率。在频率跳变系统中,码片速率决定了频率跳变的快慢,也就是系统的频率跳变速率(Frequency-Hopping Rate)。扩频码的码元宽度 T_c 称为码片,包括两部分内容,一是频率的转换时间 T_t ,另一是在该频点的持续时间,称为驻留时间(Dwell Time) T_h ,即 $T_c = T_t + T_h$ 。在频率慢跳变系统中,频率的转换时间可忽略;在频率快跳变系统中,频率转换时间的长短将成为系统成败的关键因素。在进行频率跳变系统性能的理论分析时,通常将 T_t 的影响忽略,近似认为 $T_c = T_h$ 。

从上述频率跳变系统的物理概念可以看出,频率跳变系统也占用了比信息带宽要宽得多的传输频带。从某一时间间隔看,跳频系统只是在单一射频载波上通信。但从总的通信时间上看,跳频信号用占据宽的射频频带来换取强的抗干扰能力,其扩频处理增益等于系统可选用的频率数 N (各频道之间的频谱不重叠)。例如一个具有可供选用的频率数为 $N=1000$ 的频率跳变系统,它的扩频处理增益 $G_p=30$ dB。任何外来干扰信号只有在与有用信号的载波频率相同,且在有用信号的载波持续时间(驻留时间) T_h 内才起作用。而有用信号载波的频率受扩频码的控制,当频率跳变后,干扰信号就不再起作用了。可以说频率跳变信号的特点就是在一个很宽的频带范围内采用“躲避”式的方法来抵抗干扰信号,所以也把频率跳变系统称为“躲避”式系统。

显然,频率跳变系统对扩频码发生器的要求没有直接序列系统中要求得那么高,同步要求也没有直接序列系统中要求得那样严格,但对频率合成器的要求比较苛刻,既要求频率合成器输出的频谱纯度高,又要求可供选用的频率数 N 足够大、频率转换时间短(跳变速率高)、频率锁定时间短和跳频图案(跳频指令)比较复杂等。

2. 频率跳变信号的频谱结构

这一节将分析频率跳变扩频系统跳频器输出信号的频谱。频率跳变系统频率合成器输出信号的时域表达式如下:

$$a(t) = A \sum_{k=1}^N \cos(2\pi f_k t + \varphi_k) g_{T_c}(t - kT_c) * \sum_{m=-\infty}^{\infty} \delta(t - mNT_c) \quad (3.4-11)$$

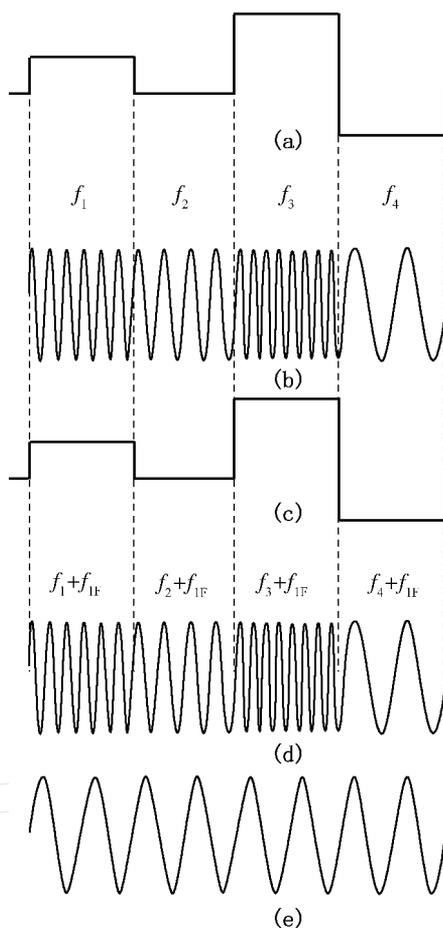


图 3.18 频率跳变系统各点信号波形示意图



由函数卷积的定义与 $\delta(f)$ 函数的取样性质, 可以推导出 $a(t)$ 的功率谱密度函数 $S_a(f)$ 如下:

$$\begin{aligned} S_a(f) &= \frac{A^2}{4} \left[\sum_{k=1}^N [\delta(f - f_k) + \delta(f + f_k)] * S_{g_{T_c}}(f) \right] \left[\frac{1}{NT_c} \sum_{m=-\infty}^{\infty} \delta\left(f - \frac{m}{NT_c}\right) \right] \\ &= \frac{A^2}{4NT_c} \sum_{k=1}^N \int_{-\infty}^{\infty} [\delta(F - f_k) + \delta(F + f_k)] S_{g_{T_c}}(f - F) dF \sum_{m=-\infty}^{\infty} \delta\left(f - \frac{m}{NT_c}\right) \\ &= \frac{A^2}{4NT_c} \sum_{k=1}^N \int_{-\infty}^{\infty} [\delta(F - f_k) + \delta(F + f_k)] \sum_{m=-\infty}^{\infty} S_{g_{T_c}}(f - F) \delta\left(f - \frac{m}{NT_c}\right) dF \end{aligned} \quad (3.4-12)$$

$a(t)$ 的功率谱密度函数 $S_a(f)$ 的单边表达式如下:

$$S_a(f) = \frac{A^2}{2N} \sum_{k=1}^N \delta(f - f_k) \quad (3.4-13)$$

由式(3.4-13)给出的频率跳变扩频系统频率合成器输出信号的频谱, 如图 3.19 所示。图 3.19 中, $f_k = f_0 + C_k f_\Delta$, $C_k = 1, 2, \dots, k, \dots, N$ ($1 \leq k \leq N$)。由图 3.19 可以看出, 频率跳变系统频率合成器输出信号的频谱由 N 个冲激函数(单频正弦波)组成, 各冲激函数的强度均为 $\frac{A^2}{2N}$, 频率间隔为 f_Δ 。这个结果是预料之中的, 由于信号的总功率为 $\frac{A^2}{2}$, 因此每个跳频载波的功率值为 $\frac{A^2}{2N}$ 。

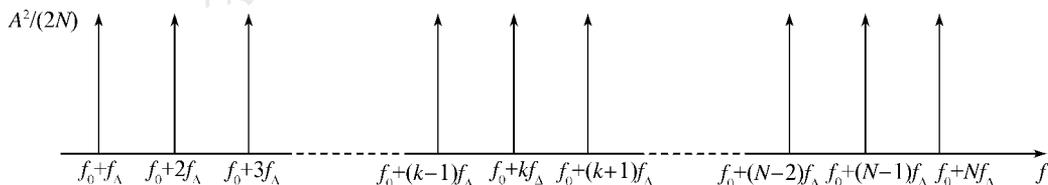


图 3.19 频率合成器输出信号的频谱

由图 3.19 或式(3.4-13)可以清楚地看出, 频率跳变扩频系统中频率合成器输出的理想频谱应为间隔相等的一系列线谱, 并且在每个频点均应具有相同的功率输出。

当信息信号 $d(t)$ 是等概取值 ± 1 的二进制数据信号, 并对载波进行 PAM 调制时, 发射机输出的信号为:

$$S_{\text{PSK}}(t) = Ad(t) \sum_{k=1}^N \cos(2\pi f_k t + \varphi_k) g_{T_c}(t - kT_c) * \sum_{m=-\infty}^{\infty} \delta(t - mNT_c) \quad (3.4-14)$$

假设信息信号 $d(t)$ 的传输速率为 R , 功率谱密度函数为 $S_d(f)$, 由式(3.4-13)和式(3.4-14)可以得出 $S_{\text{PSK}}(t)$ 的单边功率谱密度函数 $S_{\text{PSK}}(f)$, 即

$$S_{\text{PSK}}(f) = S_d(f) * S_a(f) = \frac{A^2}{2N} \sum_{k=1}^N S_d(f - f_k) \quad (3.4-15)$$

发射机输出信号的频谱示意图见图 3.20(a)和图 3.20(b)。图 3.20(a)为调制后各频



道间频谱重叠的情况;图 3.20(b)为调制后各频道间频谱不重叠的情况。

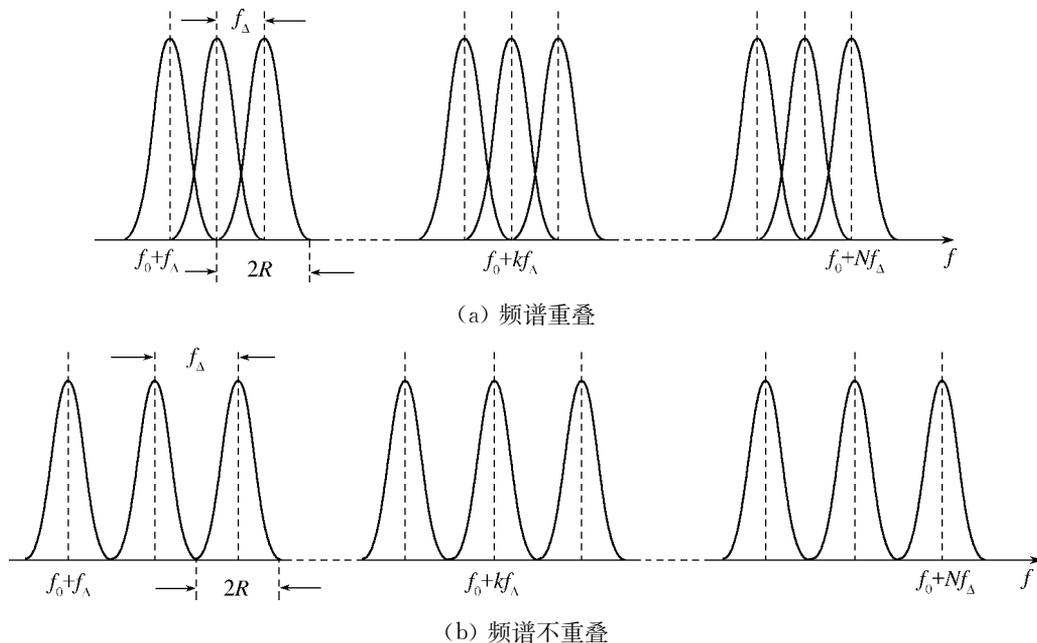


图 3.20 频率跳变系统 PSK 调制信号频谱示意图

当信息码的码速率为 R 时,为了尽可能减少邻道干扰,频率间隔 f_{Δ} 应选择为 R 。这样,中心频率为 f_k 的频道频谱的零值点就正好处于与其左右相邻的、中心频率为 $f_k \pm f_{\Delta} = f_k \pm R$ 的频道频谱的峰值处,构成所谓的频率正交关系,如图 3.20(a)所示。在这种情况下若可选取的跳频频率数为 N ,则跳频信号的带宽为 $B_{\text{RF}} = (N+1)R \approx NR$ 。

有时为了防止跳频频谱发生重叠,要求每个频道的带宽应不小于信息信号的带宽,如图 3.20(b)所示。例如,若允许射频带宽 $B_{\text{RF}} = 100 \text{ MHz}$,信息信号的速率是 1 kb/s ,为确保邻近频道不发生串扰和频谱重叠,每个频道应至少有 2 kHz 的带宽。频率跳变的间隔应不小于 2 kHz ,所以可得到 $100 \text{ MHz} / 2 \text{ kHz} = 50\,000$ 个跳频频道数。需要注意的是,在此情况下跳频信号所占据的带宽 $B_{\text{RF}} = 2NR$ 。

在频率跳变系统中,信息信号的调制也可采用 FSK(数字信号)或 FM(模拟信号)调制,以 FSK 调制为例,发射机输出信号的时域表达式为

$$S_{\text{FSK}}(t) = A \sum_{k=1}^N \cos\{2\pi[f_k + d(t)\Delta F]t + \varphi_k\} g_{T_c}(t - kT_c) * \sum_{m=-\infty}^{\infty} \delta(t - mNT_c) \quad (3.4-16)$$

式中, ΔF 为频偏或最大频移。

由式(3.4-16)求 $S_{\text{FSK}}(t)$ 的功率谱密度函数,有一定的困难,并且无助于理解系统的工作原理和工作过程。可根据系统的一般组成方式,按照系统的工作流程来推导 $S_{\text{FSK}}(t)$ 的功率谱密度函数。

频率跳变扩频系统发射信号的产生通常是通过信息信号对发信中频载波调制,然后通过中频载波的上变频过程来完成,其原理方框图参见图 3.21。

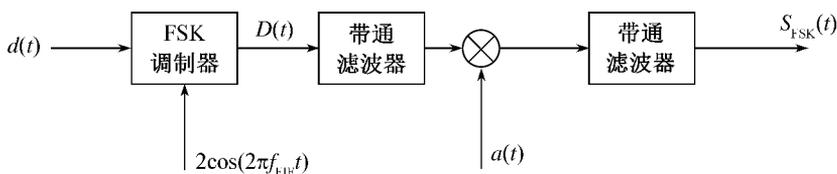


图 3.21 频率跳变扩频系统发射信号的产生

图 3.21 中,FSK 调制器后、混频器前的带通滤波器是用来滤除由于调制器的非线性所产生的带外分量并完成已调信号频谱的成形,若假设调制器的特性比较理想,在进行理论分析时,可忽略其影响;混频器后的带通滤波器是滤除混频后信号的差频分量。

由图 3.21 可看出,信息信号 $d(t)$ 被调制到发信中频 F_{TIF} 上,形成中心频率为 F_{TIF} 的带通信号 $D(t)$,即

$$D(t) = 2\cos\{2\pi[f_{\text{TIF}} + d(t)\Delta F]t + \varphi_0\} \quad (3.4-17)$$

设发射机的本振信号为

$$a(t) = A \sum_{k=1}^N \cos[2\pi(f_k - f_{\text{TIF}})t + \varphi'_k] g_{T_c}(t - kT_c) * \sum_{m=-\infty}^{\infty} \delta(t - mNT_c) \quad (3.4-18)$$

令 $\varphi_k = \varphi'_k + \varphi_0$,则信号 $D(t)$ 和 $a(t)$ 经过混频器后为

$$\begin{aligned} a(t)D(t) &= A \sum_{k=1}^N \cos\{2\pi[f_k + d(t)\Delta F]t + \varphi_k\} g_{T_c}(t - kT_c) * \sum_{m=-\infty}^{\infty} \delta(t - mNT_c) + \\ &A \sum_{k=1}^N \cos\{2\pi[f_k - 2f_{\text{TIF}} - d(t)\Delta F]t + \varphi'_k - \varphi_0\} g_{T_c}(t - kT_c) * \sum_{m=-\infty}^{\infty} \delta(t - mNT_c) \end{aligned} \quad (3.4-19)$$

经过带通滤波器滤除频率较低的差频分量 ($f_k - 2f_{\text{TIF}} - d(t)\Delta F$) 后,就得到了所需要的发射信号 $S_{\text{FSK}}(t)$,见式(3.4-16)。

设带通信号 $D(t)$ 的功率谱密度函数为

$$S_D(f) = S_{d,\text{FSK}}(f - f_{\text{TIF}}) + S_{d,\text{FSK}}(f + f_{\text{TIF}}) \quad (3.4-20)$$

则混频器输出信号的功率谱为

$$\begin{aligned} S_D(f) * S_a(f) &= \frac{A^2}{4N} \sum_{k=1}^N [S_{d,\text{FSK}}(f - f_k) + S_{d,\text{FSK}}(f + f_k)] \\ &+ \frac{A^2}{4N} \sum_{k=1}^N [S_{d,\text{FSK}}(f - f_k + 2f_{\text{TIF}}) + S_{d,\text{FSK}}(f + f_k - 2f_{\text{TIF}})] \end{aligned} \quad (3.4-21)$$

因此,带通滤波器输出信号 $S_{\text{FSK}}(t)$ 的功率谱密度函数 $S_{\text{FSK}}(f)$ 为

$$S_{\text{FSK}}(f) = \frac{A^2}{4N} \sum_{k=1}^N [S_{d,\text{FSK}}(f - f_k) + S_{d,\text{FSK}}(f + f_k)] \quad (3.4-22)$$

或者单边形式



$$S_{\text{FSK}}(f) = \frac{A^2}{2N} \sum_{k=1}^N S_{d,\text{FSK}}(f - f_k) \quad (3.4-23)$$

FSK 属于非线性调制,信息信号对载波进行 FSK 调制的频率跳变扩频信号频谱特性的研究也比较困难。但在一定条件下,如假设 FSK 信号的相位不连续,那么 FSK 信号就可以看作两个载波频率不同的振幅键控信号的叠加,这样可以根据式(3.4-23)近似给出频率跳变信号的功率谱密度函数示意图,参见图 3.22。

由式(3.4-15)和式(3.4-23)可以看出,在频率跳变系统中,尽管采用不同的调制方式,使得发射信号的频谱不同,但它们具有一个共同的特点:信号的频谱是由 N 个频谱结构完全相同的带通信号的频谱组成的,每个带通信号的中心频率分别为 f_1, f_2, \dots, f_N ,而每个带通信号的频谱结构取决于信息信号的调制方式。

假如频率合成器输出信号的频率固定不变为 f_0 , N 个带通信号的频谱叠加在一起。由于各带通信号的频谱结构完全相同,叠加后的频谱结构不会发生变化,仅是幅度增加了 N 倍。此时,频率跳变系统就退化为常规的通信系统了,其信号的频谱结构和常规系统的频谱结构也完全相同了,这从另一方面验证了上述结论的正确性。

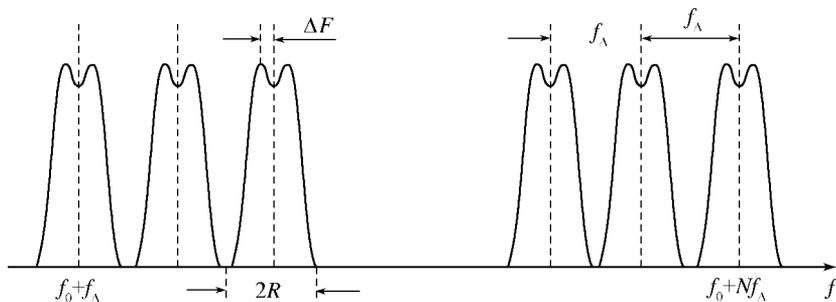


图 3.22 频率跳变系统信号频谱示意图(FSK 调制)

虽然上述结论是通过分析 PAM 和 FSK 调制方式得出的,但其具有普遍意义,对任何调制方式都适用,不同的调制方式仅改变每个带通信号的频谱结构。

3. 频率跳变信号的产生

频率跳变系统信号的产生,很少采用信息信号直接去调制频率跳变的载波,原因有几点:① 当系统工作的射频频率较高时,实际工程中在较高频率上实现信号的调制是有一定困难的。② 信息信号通过调制实现频谱搬移的过程中,不可避免地要产生一些带外分量。调制器根本无法抑制这些带外分量,通常是在调制器后插入带通滤波器来控制已调信号的频谱的。③ 调制器的参数或特性与工作频率有密切的关系,载波在如此大的范围内变化,已调信号在各个频道上的特性很难保证完全一致。

频率跳变系统信号的产生,通常采用在发射机的中频上进行信息信号的调制,再利用变频器进行上变频,将中频已调信号的频谱搬移到射频段。变频器的本振信号由频率合成器来提供,参见图 3.23。由于频率合成器输出信号的频率是跳变的,上变频器输出的带通信号的中心频率将随着频率合成器输出信号频率的变化而变化,从而完成了射频信号载波频率的跳变。

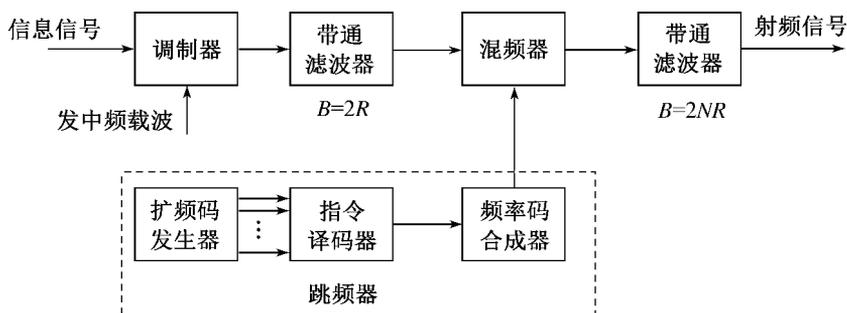


图 3.23 频率跳变系统信号产生原理方框图

由图 3.23 可看出,频率跳变扩频系统信号的产生和常规系统相比,区别仅在于参与变频的本振信号的频率是跳跃变化的,所以说跳频器是频率跳变扩频通信系统的核心器件。图 3.23 中,混频器前的带通滤波器是用来控制已调信号频谱的,滤除频谱搬移过程中产生的一些带外分量。混频器后的滤波器是用来滤除混频过程中产生的差频(或和频)分量,使输出信号中只含有和频(或差频)分量的。

3.5 多载波调制

南京大学出版社

3.5.1 正交频分复用 (OFDM)

1. 正交频分复用介绍

多载波调制通过将可用的信道带宽 W 划分为 N 条窄带子信道,每条子信道是带宽为 $\Delta f = W/N$ 的子带。在每一个子带中的信号可以独立地进行编码,并以相同的符号速率 $1/T$ 和最佳功率配置 $P(f)$ 来进行调制。当 Δf 充分窄时,子信道带宽小于多径信道的相干带宽,使得信道频率响应 $C(f)$ 在每个子带范围内可近似视为常数(即平坦衰落特性)。OFDM 将高速串行符号转换为并行的低速子数据流,每个子信道上符号持续时间是原来的 N 倍,大于最大多径时延,这样可以有效消除符号间干扰 (ISI, Inter Symbol Interference)。

图 3.24 展示了对信道带宽 W 的细分。

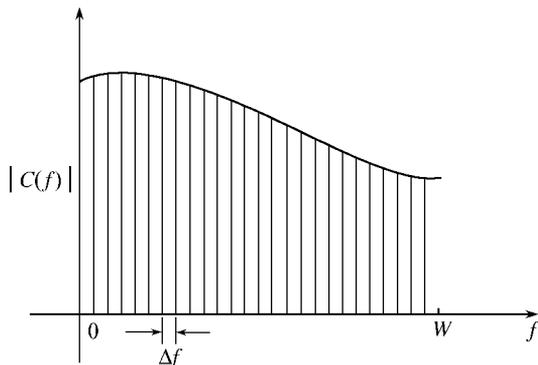


图 3.24 把信道带宽 W 细分为具有相同带宽 Δf 的子带



正交频分复用(Orthogonal Frequency Division Multiplexing, OFDM)是一种子载波互相混叠且互相正交的多载波调制技术。OFDM的基本原理是将发送的高速数据流分散到多个子载波上,使各子载波的信号速率大幅降低,从而提高抗多径和抗衰落的能力。对第 k 条子信道,关联的载波波形为

$$s_k(t) = \cos 2\pi f_k t, k=0, 1, \dots, N-1 \quad (3.5-1)$$

式中, f_k 是第 k 个子信道的中心频率。每个子信道上的符号速率 $1/T$ 都等于相邻子载波的频率间隔 Δf ,因此各子载波的已调信号频谱有 $1/2$ 重叠,但保持相互正交,并且这种正交性与子载波间的相对相位无关,即

$$\int_0^T \cos(2\pi f_k t + \phi_k) \cos(2\pi f_j t + \phi_j) dt = 0 \quad (3.5-2)$$

式中, $f_k - f_j = n/T (n=0, 1, \dots, N-1)$,与相位值 ϕ_k 和 ϕ_j 无关。这样就构建出了正交频分复用信号。换句话说,OFDM就是一类特殊的多载波调制,其各个子信道上的相应的子载波相互正交。在实际系统中,通过IFFT/FFT可高效地产生与检测上述正交子载波。OFDM已被广泛应用于数字音频广播、IEEE 802.11无线局域网、LTE/5G移动通信等标准中。

2. OFDM系统的调制和解调

在一个具有 N 个子信道的OFDM系统中,每一个子载波上的符号速率 $1/T$ 降低到 $1/(NT_s)$,其中 T_s 是单载波系统的符号间隔。这是相对于单载波系统而言的,该系统占据所有的带宽 W ,并且以同样的速率传输数据,因此OFDM的符号间隔 $T = NT_s$ 。只要选择 N 足够大,符号间隔 T 就可以远远大于信道时间色散的持续时间,所以符号间干扰可以通过选择 N 变得任意小。换句话说,每一个子信道看上去就好像拥有一个固定的频率响应 $C(f_k), k=0, 1, \dots, N-1$ 。

假定每一个子载波都采用 M 进制QAM,则第 k 个子载波上的信号可以表示为

$$\begin{aligned} u_k(t) &= \sqrt{\frac{2}{T}} A_{ki} \cos 2\pi f_k t - \sqrt{\frac{2}{T}} A_{kq} \sin 2\pi f_k t \\ &= \operatorname{Re} \left[\sqrt{\frac{2}{T}} A_k e^{j\theta_k} e^{j2\pi f_k t} \right] = \operatorname{Re} \left[\sqrt{\frac{2}{T}} X_k e^{j2\pi f_k t} \right] \end{aligned} \quad (3.5-3)$$

式中, $X_k = A_k e^{j\theta_k}$ 是第 k 个子载波上传输的QAM信号星座上的点, $A_k = \sqrt{A_{ki}^2 + A_{kq}^2}$,且 $\theta_k = \arctan(A_{kq}/A_{ki})$,每个调制符号的能量被吸收到 $\{X_k\}$ 中。

当子信道数 N 很大,每条子信道的带宽 Δf 足够窄时,可将其频率响应 $C(f_k)$ 近似为常数:

$$C(f_k) = C_k = |C_k| e^{j\phi_k} \quad (3.5-4)$$

所以第 k 个子信道上的接收信号就是

$$\begin{aligned} r_k(t) &= \sqrt{\frac{2}{T}} |C_k| A_{ki} \cos(2\pi f_k t + \phi_k) + \sqrt{\frac{2}{T}} |C_k| A_{kq} \sin(2\pi f_k t + \phi_k) + n_k(t) \\ &= \operatorname{Re} \left[\sqrt{\frac{2}{T}} C_k X_k e^{j2\pi f_k t} \right] + n_k(t) \end{aligned} \quad (3.5-5)$$



式中, $n_k(t)$ 代表了第 k 个子信道上的加性噪声。假定 $n_k(t)$ 是零均值高斯噪声, 并且在第 k 个子信道的带宽上是谱平坦的; 还假设信道参数 $|C_k|$ 和 ϕ_k 在接收机方都是已知的。这些参数通常可以这样来估计: 发送 $\cos(2\pi f_k t)$, 接收到的信号为 $|C_k| \cos(2\pi f_k t + \phi_k)$ 。

对于第 k 个子信道上的接收信号可以这样解调, 将 $r_k(t)$ 与两个基函数进行互相关, 假定接收机已知载波相位, 则两个基函数为

$$\psi_1(t) = \sqrt{2T} \cos(2\pi f_k t + \phi_k), 0 \leq t \leq T \quad (3.5-6)$$

$$\psi_2(t) = -\sqrt{2T} \sin(2\pi f_k t + \phi_k), 0 \leq t \leq T \quad (3.5-7)$$

在 $t=T$ 时刻对互相关器的输出进行抽样。这样, 得到的接收信号矢量为

$$\mathbf{y}_k = (|C_k| A_{ki} + \eta_{kr}, |C_k| A_{kq} + \eta_{ki}) \quad (3.5-8)$$

该矢量可以表示为复数

$$Y_k = |C_k| X_k + \eta_k \quad (3.5-9)$$

式中, $X_k = A_{ki} + jA_{kq}$, $\eta_k = \eta_{kr} + j\eta_{ki}$ 表示加性噪声。

假定已知信道增益 $|C_k|$, Y_k 除以 $|C_k|$, 我们得到

$$Y'_k = Y_k / |C_k| = X_k + \eta'_k \quad (3.5-10)$$

式中, $\eta'_k = \eta_k / |C_k|$ 。归一化变量 Y'_k 传输给检测器, 检测器计算出 Y'_k 与 QAM 信号星座图上各点之间的距离度量, 并选中最小距离的信号点。

根据以上讨论, 有一点很清楚, 就是用到两个互相关器或匹配滤波器来解调每个子信道上的接收信号, 所以, 如果 OFDM 信号含有 N 个子信道, 那么实现 OFDM 解调器就需要由 $2N$ 个互相关器或 $2N$ 个匹配滤波器组成的并行阵列。此外, OFDM 信号的产生过程还可以被看成用符号去激励 $2N$ 个并行滤波器组, 而符号来自 M 进制 QAM 信号星座图。

在发送机中用 $2N$ 个并行滤波器产生调制信号, 在接收机用 $2N$ 个并行滤波器解调接收信号, 这等价于计算离散傅里叶变换(DFT)及其逆变换。快速傅里叶变换(Fast Fourier Transform, FFT)是高效率地计算 DFT 的一种方法, 故当 N 很大时, 如 $N > 32$ 时, 我们可以采用 FFT 算法来实施调制和解调过程。下面将介绍一种实现 OFDM 系统的调制解调方案, 即采用 FFT 算法计算 DFT。

3. OFDM 系统的 FFT 算法实现

OFDM 系统的核心目标是将一组待传送的符号序列映射为一个复合的多载波信号, 进而通过无线信道传输。为了实现这一目标, 现代通信系统普遍采用基于数字信号处理的实现方案, 即在数字域内通过快速傅里叶变换(FFT)合成 OFDM 信号的采样序列, 然后通过数模转换(DAC)和射频调制完成发射, 如图 3.25 所示。

为了便于分析, 将多载波调制看成由 N 个独立的 QAM 信道所组成, 每个信道具有各自不同的 QAM 星座。我们用 $X_k (k=0, 1, \dots, N-1)$ 来表示相应于各子信道上的信息符号的复信号点, 于是 OFDM 信号波形可表示为

$$x(t) = \sum_{k=0}^{N-1} u_k(t) = \sum_{k=0}^{N-1} X_k e^{j2\pi f_k t} \quad (3.5-11)$$

其中 $f_k = f_c + k\Delta f$ 。由于 OFDM 信号的产生是首先在基带实现, 然后通过变频产生输出

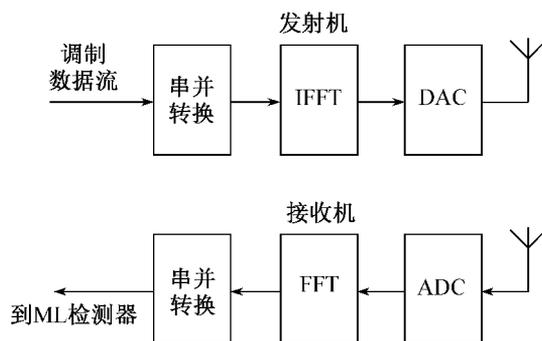


图 3.25 简化的 OFDM 收发机结构

信号,因此仅考虑基带处理时可令 $f_c=0$,则式(3.5-11)可简化为

$$x(t) = \sum_{k=0}^{N-1} X_k e^{j2\pi k \Delta f t} \quad (3.5-12)$$

子载波频率间隔为 $\Delta f = \frac{1}{T}$,于是式(3.5-12)又可化为

$$x(t) = \sum_{k=0}^{N-1} X_k e^{\frac{j2\pi k t}{T}} \quad (3.5-13)$$

我们考虑使用数字信号处理手段产生该信号,为此需要先获得信号的采样序列。设在 OFDM 符号的每个持续时间内采样 N 次,即以采样间隔 $\frac{T}{N}$ 对信号(3.5-13)进行采样,可以得到采样序列为

$$x[n] = x\left(\frac{nT}{N}\right) = \sum_{k=0}^{N-1} X_k e^{\frac{j2\pi k n}{N}} \quad (3.5-14)$$

让我们回顾离散傅里叶变换对(DFT)的表达式,即

$$Y[k] = \sum_{n=0}^{N-1} y[n] e^{-\frac{j2\pi k n}{N}} \quad (3.5-15)$$

$$y[n] = \frac{1}{N} \sum_{k=0}^{N-1} Y[k] e^{\frac{j2\pi k n}{N}} \quad (3.5-16)$$

可以发现,OFDM 信号的采样序列表达式与逆离散傅里叶变换(IDFT)的数学形式完全一致,只需将待传信息符号 X_k 作为频率采样信号 $Y[k]$,对其执行 IDFT 并乘以一个归一化因子,即可得到 OFDM 基带信号的时域样本序列 $x[n]$ 。将该序列分别取实部虚部,经过 DAC、低通滤波器后得到 OFDM 基带信号的同相与正交分量,经过 IQ 调制后即可获得可供传输的 OFDM 信号波形 $x(t)$,如图 3.26 所示。

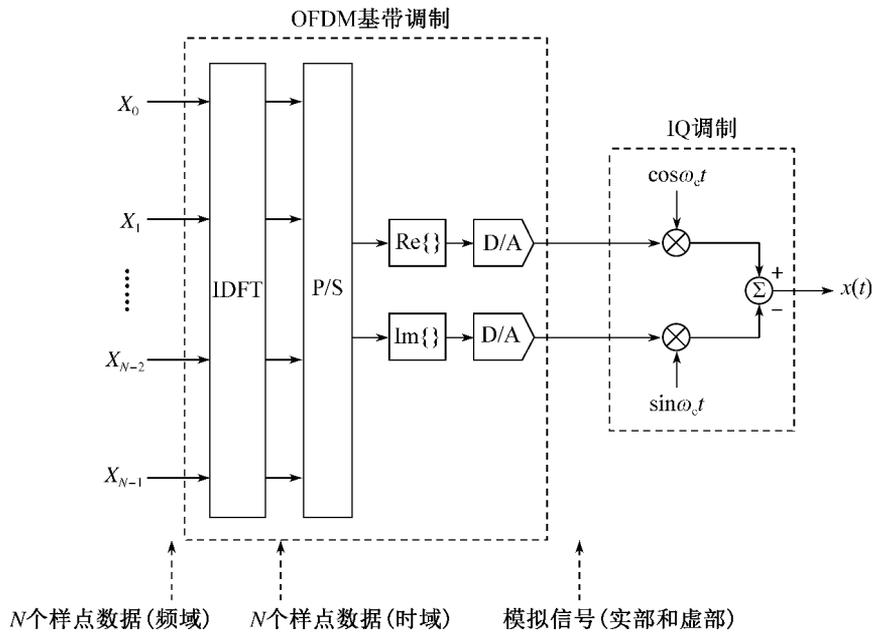


图 3.26 OFDM 调制实现原理

上述为在发射端利用 IDFT 产生 OFDM 信号的过程。相似地，在 OFDM 系统的接收端，我们可以用作 IDFT 逆过程的 DFT 来完成 OFDM 信号的解调。其具体过程如图 3.27 所示。

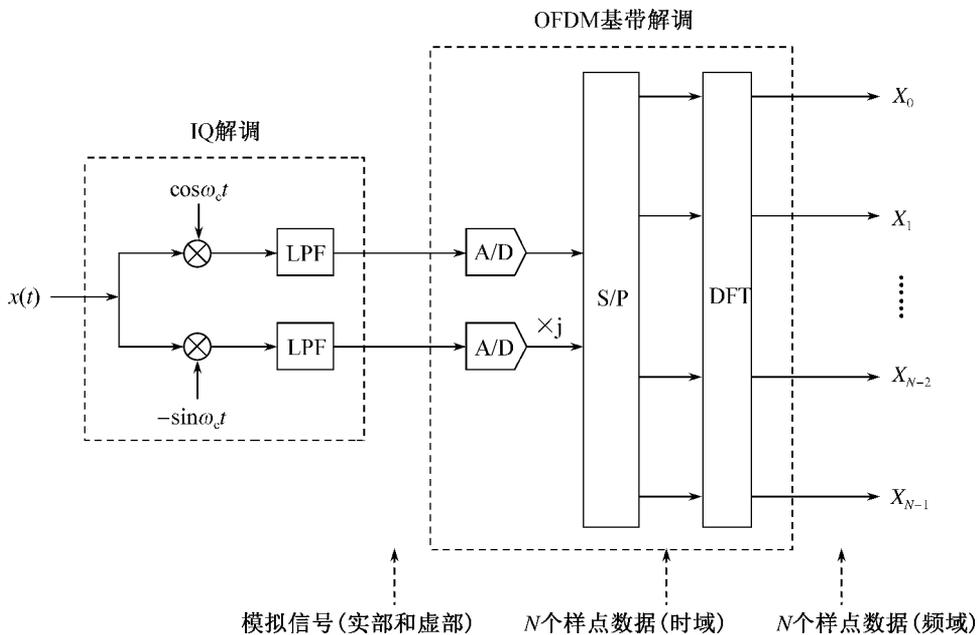


图 3.27 OFDM 解调实现原理

接收端接收到的 OFDM 信号首先经过 IQ 解调，分别提取出基带信号的同相与正交分

量,再通过 ADC 采样得到 OFDM 基带信号的复样本序列。若信道理想无噪声干扰,该序列应与发送端 IDFT 输出的序列一致。接下来,通过对该序列执行 DFT,即可恢复出调制符号序列 X_k 。

由此可见,若选择载波频率间隔 $\Delta f = \frac{1}{T}$,则 OFDM 信号不但保持各子载波相互正交,而且可以用离散傅里叶变换(DFT)来表示。为提升运算效率,实际系统中通常采用快速傅里叶变换(FFT)算法来实现 DFT。

因此,在 OFDM 系统中引入 FFT 技术实现频域符号的调制与解调,能够将 OFDM 信号的生成与恢复完全置于数字基带处理过程中,无需多个本振与滤波器的硬件支持,从而大幅度降低系统的实现复杂度,提高系统的可重构性与可扩展性。

在实际无线信道中,多径传播是一种普遍存在的现象。由于信号在不同路径上传播所经历的传播时延不同,接收端将接收到原始信号多个不同时延版本的叠加。这种时延拓展效应可能导致连续发送的符号在接收端发生重叠,进而引起符号间干扰(Inter-Symbol Interference, ISI)。在传统单载波系统中,由于每个符号持续时间较短,稍有时延扩展便可能导致严重的干扰。而在 OFDM 系统中,发送端首先将高速串行数据流并行划分到多个低速子信道中,使每个子载波上承载的数据速率大大降低,相应地,每个 OFDM 符号的持续时间被扩展为原始符号周期的 N 倍。这种低符号速率特性天然增强了对多径的容忍能力,使得同样的信道时延扩展相对于符号周期而言显得更短,从而显著减轻了 ISI 的影响。

尽管较长的符号周期增强了 OFDM 系统对时延拓展引发的符号间干扰的抵抗能力,但这可能仍无法完全消除干扰。为了进一步抑制码间串扰,OFDM 系统引入了保护间隔(Guard Interval, GI)机制。其基本思想是在两个连续 OFDM 符号之间插入一段冗余时间,只要该间隔长度大于信道的最大时延扩展 τ_{\max} ,则来自上一个符号的所有多径分量将在该间隔内衰减完毕,不会干扰当前符号,从而有效抑制 ISI。

设原始 OFDM 符号持续时间为 T ,则每个符号的实际传输周期扩展为

$$T' = T + T_{GI} \quad (3.5-17)$$

其中 T_{GI} 为保护间隔持续时间。只要保护间隔长度满足 $T_{GI} \geq \tau_{\max}$,即可保证无线信道中前一个符号的所有多径分量在当前符号有效部分开始前抵达并衰减完毕,从而避免对当前符号造成干扰,有效抑制 ISI。

既然只要保护间隔的长度足够,就能够在理论上避免多径信道下的符号间干扰,那么一个自然的问题便是:这段保护间隔应当以何种方式填充,才能在抑制干扰的同时不引入新的问题?

一种最朴素的保护间隔实现方式是在两个符号之间插入一段空白时间,即不传输任何有效信号,这种填充方式称为零符号填充(Zero Padding, ZP)。在这种情形下,OFDM 发射波形可描述为

$$x_{ZP}(t) = \sum_n \sum_{k=0}^{N-1} X_k^{(n)} e^{j2\pi f_k (t-nT')} \cdot \Pi_T(t-nT') \quad (3.5-18)$$

其中 $\Pi_T(t)$ 表示长度为 T 的矩形窗函数, $X_k^{(n)}$ 为第 n 个 OFDM 符号中第 k 个子载波的调制符号,子载波频率为 $f_k = f_0 + k\Delta f$,且 $\Delta f = 1/T$ 。



虽然满足 $T_{GI} \geq \tau_{\max}$ 的空白保护间隔可以有效防止符号间干扰,但这也带来了新的问题:由于信号在时间上不再连续,频域上会出现能量泄漏,破坏子载波之间的正交性,进而产生载波间干扰(Inter-Carrier Interference, ICI)。这是因为在接收端进行离散傅里叶变换(DFT)时,默认每个子载波在 FFT 窗口内完成整数个周期,但空白间隔导致子载波波形被截断,实际周期数偏离整数,结果不同子载波之间会互相干扰。

为了解决空白间隔带来的载波间干扰问题,OFDM 系统采用了循环前缀(Cyclic Prefix, CP)方案。循环前缀是将每个 OFDM 符号尾部的一段波形复制并插入到符号的起始部分,从而延长一个符号的持续时间,同时保持其在时间上的周期结构连续性。设 IFFT 输出的离散时域采样为

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{j\frac{2\pi kn}{N}}, n=0,1,\dots,N-1 \quad (3.5-19)$$

则循环前缀构造为

$$\tilde{x}[n] = x[n+N], n=-L_{CP}, \dots, -1 \quad (3.5-20)$$

最终加入保护间隔后的信号的离散时域采样为

$$x_{CP}[n] = \begin{cases} x[n+N] & -L_{CP} \leq n < 0 \\ x[n] & 0 \leq n < N \end{cases} \quad (3.5-21)$$

其中 L_{CP} 为循环前缀长度,要求能覆盖信道冲激响应长度 L_h 。循环前缀使得接收端观察窗口内的信号结构在时间上连续,子载波在积分区间内仍具有整数周期,因此保持正交性,有效抑制 ICI。另一方面,由于主符号与前缀属于同一个周期结构,其周期延拓也将原本的线性卷积转化为循环卷积,即在与信道冲击响应 $h[n]$ 的卷积运算中有

$$x_{CP}[n] * h[n] \Rightarrow x[n] \otimes h[n] \quad (3.5-22)$$

从而在频域变换后得到简单的乘法关系

$$Y[k] = X[k] \cdot H[k] \quad (3.5-23)$$

其中 $H[k]$ 为信道频率响应, $Y[k]$ 为接收符号频域样本。在接收端只要移除前缀并对主符号部分执行 N 点 DFT,即可得到 $Y[k]$ 。

此时各子载波上的信号彼此独立,接收端可以利用简单的单抽头均衡器进行符号估计,即有

$$\hat{X}[k] = \frac{Y[k]}{H[k]} \quad (3.5-24)$$

这种均衡器结构大幅度降低了接收端实现的复杂度,是 OFDM 系统的一项重要优势。

因此,循环前缀不仅具备了保护间隔对抗 ISI 的能力,还维持了频域正交性消除了 ICI,同时极大简化了信道均衡的实现。具有循环前缀保护间隔以及均衡的完整 OFDM 系统框图如图 3.28 所示。

尽管引入循环前缀有效抑制了干扰,但也带来了一定的频谱效率损失。例如,若循环前缀持续时长为 T_{CP} ,则 CP 占比为 $a = \frac{T_{CP}}{T}$,则系统效率变为

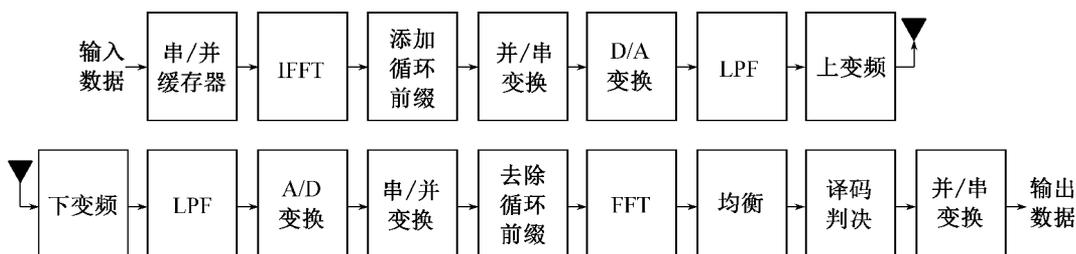


图 3.28 OFDM 系统调制解调过程

$$\eta = \frac{T}{T + T_{CP}} = \frac{1}{1 + \alpha} \quad (3.5-25)$$

为了在可靠性与效率之间取得平衡,现代 OFDM 系统还引入了信道自适应(Channel Adaptation)机制。系统根据实时的信道状态信息动态调整传输参数,如保护间隔长度、调制方式和编码速率等。在信道条件良好(高信噪比、低多径)时,可缩短循环前缀并采用高阶调制;而在信道劣化时,则延长保护间隔、降低调制阶数并加重纠错编码强度,以保证误码性能。

3.5.2 正交时频空间调制(OTFS)

1. 时频双选信道

在无线通信的现实环境中,信道衰落是难以避免的现象。通信信号的传播环境、用户的移动速度等都会对信道的衰落程度造成影响,而无线信道的衰落主要分为大尺度衰落和小尺度衰落。大尺度衰落是指信号在较大时间空间尺度上的变化,由距离、建筑物、山脉等影响导致的,小尺度衰落则是信号短期内的衰落,发生在较短的时间间隔和较小的空间间隔内,比如多径引起的快速变化,则是由于接收来自不同路径的信号到达的传输时间相位不同,叠加后出现相干叠加或相消叠加,从而导致信号的波动。此外,用户的移动也会使得接收信号出现小尺度衰落现象。具体为移动引起多普勒频移,从而改变接收信号频谱,影响信号相位幅度。移动速度越快,多普勒现象越明显,小尺度衰落现象更明显。

信道多径传输和用户移动是小尺度衰落的两个重要原因,因此,小尺度衰落一般使用信道的多径时延扩展和多普勒扩展描述。一般而言,相比于大尺度衰落,小尺度衰落对信号的影响更大,所以本节主要研究消除小尺度衰落的技术。

为更好地描述通信信道对信号的影响,在通信系统中,一般将信号在传输过程中经过的信道对不同频率成分响应的不同而导致的在接收端出现的不同频率的衰落现象称为信道的频率选择性;将信号在传播过程中经过的信道具有不同的瞬时响应导致信号在接收端经历时间上的快速变化的现象称为信道的时延选择性。根据频率选择性和时延选择性的定义可以分析出,造成频率选择性的主要原因是信道的多径效应,而造成时延选择性的主要原因是用户的移动。

因此,对于高速移动通信的场景,除了常见的信道多径传播现象之外,还由于用户的高速移动引起了信道的多普勒扩展现象,从而导致无线信道同时存在频率选择性衰落和时延选择性衰落。因此,在高速移动通信场景中,通信用户的发送信号会经过时频双选信道的作



用,接收信号会出现由于信道的时频双选性导致的衰落和干扰现象。

更为具体的,时频双选信道的冲击响应数学描述可以表述为:

$$h(t, \tau) = \sum_{i=1}^P h_i e^{j2\pi\nu_i t} \delta(\tau - \tau_i) \quad (3.5-26)$$

其中, P 表示该信道中总共有 P 条可解析子路径,并且第 i 条传播路径的信道增益、多普勒频移和时延分别为 $h_i, \nu_i, \tau_i, \delta(\tau)$ 表示冲击函数。假设接收端的信号采样周期为 T_s ,则离散等效信道模型为:

$$h(m, l) = \sum_{i=1}^P h_i e^{j2\pi\nu_i m T_s} \delta\left(l - \frac{\tau_i}{T_s}\right), m = [0, M-1], l = [0, L-1] \quad (3.5-27)$$

其中, M 为子载波数量, $L = \lceil \tau_{\max}/T_s \rceil$ 由最大时延确定。由上式可以看出,由于信道的时变性,导致获得时域信道估计参数增加且信道参数需要频繁更新估计值,所需估计参数数量多且接收时域信号变化较快,显著增加了接收端信号恢复的难度。上述对信道的描述与建模是从时间—频率域的角度出发,可以看出,在信道具有时频双选特性的情况下,时间—频率域的信道表示复杂而多变。在时间—频率域之外,还存在着许多其他变换域的信道表示。在不同变换域对信道不同的表示形式可以对信道的不同维度特性进行刻画。针对时频双选信道的情况,相比于情况复杂的时频域表征,研究学者发现,时延—多普勒域的信道表征更为简洁,其数学建模可以表示为:

$$h(\tau, \nu) = \sum_{i=1}^P h_i \delta(\tau - \tau_i) \delta(\nu - \nu_i) \quad (3.5-28)$$

其中, h_i, τ_i, ν_i 分别表示信道中第 i 条传播子路径的信道增益、对应的信道时延、信道多普勒频移。从上式可以看出,表示时延—多普勒域信道所需的参数远远小于时频域信道所需参数。另外,由于信道的时延—多普勒域信道特征是由信道的物理特性所确定且在一些典型的通信信道环境中,物理反射体数量往往比较少,因此,即使在高速移动的情况下,时延—多普勒域信道具有稀疏性和几乎保持不变的特性,利用这个特性能够显著降低高速移动场景下传统时频域通信系统由于信道时变性带来的性能下降。鉴于此,在高速移动通信场景下,利用时频双选信道的时延—多普勒域的特性,采用正交时频空(Orthogonal Time Frequency and Space, OTFS)调制技术有望成为解决用户通信质量下降问题的重要途径之一。在下一小节,将详细介绍基于时延—多普勒域的 OTFS 调制技术的基本原理。

2. OTFS 调制技术

OTFS 通信系统结构如图 3.29 所示。基于上述对高速移动场景下时延—多普勒域信道特征的分析,可以看出,OTFS 调制技术的主要核心思想是利用时延—多普勒域的信道的稳定性,因此,在 OTFS 系统中,调制信号被映射在时延—多普勒域,而非传统的时间—频率域。下面,首先介绍时延—多普勒域信号表示与时间—频率域信号表示之间的转换关系。

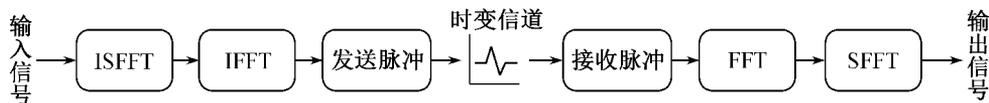


图 3.29 OTFS 系统结构示意图

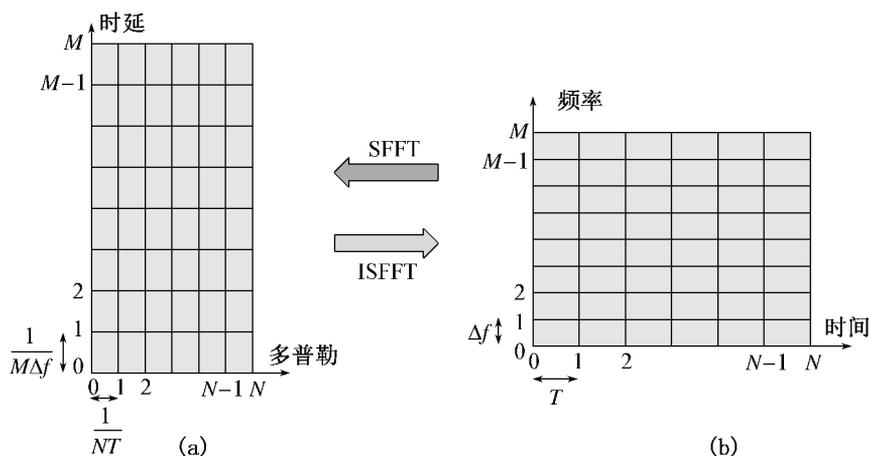


图 3.30 时延—多普勒平面和时间—频率平面示意图

在 OTFS 系统发送端,调制信号被映射在时延—多普勒域对应资源格后,通过逆辛有限傅里叶变换(Inverse Symplectic Finite Fourier Transform, ISFFT)在发送端将映射在时延—多普勒域的调制符号转变到时间—频率域,之后,为了对信号发射进入物理信道准备,通过逆傅里叶变换操作将时间—频率域的信号转变到时域信号,再通过发送脉冲承载时域信号发射后进入时变物理信道,这一步也可以被称为 Heisenberg 变换。相对应的,在接收端获得时域接收信号后,通过接收脉冲采样后得到接收时域离散信号,这一步也称为温格(Wigner)变换。接下来,在得到离散时域信号后,需要经过发送端变换的逆操作得到相应的在时延—多普勒域的信号表示。具体为接收端通过快速傅里叶变换将时域信号转变为时间—频率域信号,通过辛有限傅里叶变换(Symplectic Finite Fourier Transform, SFFT)将时间—频率域的信号转变为时延—多普勒域信号表示。

接下来,为了便于信号的数学表达与理解,定义如图 3.30 中所示的时延—多普勒域和时间—频率域网格以及其数学表达。具体地,定义时延—多普勒域平面表示为:

$$\Lambda^{\text{DD}} = \left\{ \left(\frac{l}{M\Delta f}, \frac{k}{NT} \right) \mid l=0, \dots, M-1, k=0, \dots, N-1 \right\} \quad (3.5-29)$$

在式(3.5-29)中, M, N 各表示一个通信帧中子载波数量和符号数量, $\Delta f, T$ 各表示子载波间隔和采样时间间隔, $1/M\Delta f, 1/NT$ 表示时延域和多普勒域的分辨率。上标 DD 表示时延—多普勒域。

与式(3.5-29)中相对应的时延—频率平面的数学表征可以定义为:

$$\Lambda^{\text{TF}} = \{ (m\Delta f, nT) \mid m=0, \dots, M-1, n=0, \dots, N-1 \} \quad (3.5-30)$$

在式(3.5-30)中,上标 TF 表示时间—频率域。如图 3.30 所示,时延—多普勒域信号表示和时间—频率域信号表示可以通过 ISFFT/SFFT 互相转换。OTFS 发送端利用 ISFFT 将时延—多普勒域信号 \mathbf{X}^{DD} 转换成时间—频率域,其表达式为:

$$\mathbf{X}^{\text{TF}}[m, n] = \frac{1}{NM} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} \mathbf{X}^{\text{DD}}[k, l] e^{j2\pi \left(\frac{kn}{N} - \frac{ml}{M} \right)} \quad (3.5-31)$$

在式(3.5-31)中, $\mathbf{X}^{\text{TF}} \in \mathbb{C}^{M \times N}$ 是信号的时间—频率域表示形式。接着,海森伯格变换



将时间—频率域的信号转成如下形式的时域信号：

$$s(t) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \mathbf{X}^{\text{TF}}[m, n] g_{\text{tx}}(t - nT) e^{j2\pi m \Delta f (t - nT)} \quad (3.5-32)$$

其中, $g_{\text{tx}}(t)$ 表示发送脉冲函数。以上为 OTFS 调制通信系统发送端信号处理过程。

由于发射信号 $s(t)$ 在收发端具有相对高速移动场景下的信道中传输, 故基带信道的脉冲响应是关于时延 τ 和多普勒频移参数 ν 的复杂响应。其中信道脉冲响应 $h(\tau, \nu)$ 可表示为

$$h(\tau, \nu) = \sum_{i=1}^P h_i \delta(\tau - \tau_i) \delta(\nu - \nu_i) \quad (3.5-33)$$

式中: P 表示路径数量; h_i, τ_i, ν_i 分别代表第 i 条路径相关的路径增益、时延和多普勒频移, 为简化说明, 这里忽略分数时延和多普勒频移, 只考虑整数情况, 则 τ_i, ν_i 取值分别为 $l/M\Delta f$ 和 k/NT , 其中 $l=0, \dots, M-1, k=0, \dots, N-1$ 。则传输信号 $r(t)$ 如下式所示:

$$r(t) = \iint h(\tau, \nu) s(t - \tau) e^{j2\pi \nu (t - \tau)} d\tau d\nu \quad (3.5-34)$$

在接收端, 为了将接收的时域信号转换到时延—多普勒域, 需要对时域信号进行式(3.5-31)中 ISFFT 和式(3.5-32)中海森伯格变换的逆操作, 具体为 SFFT 操作和温格变换, 这两个变换的数学表达为:

$$Y(t, f) = \int g_{\text{rx}}^*(t' - t) r(t') e^{-j2\pi f(t-t')} dt' \quad (3.5-35)$$

$$Y^{\text{TF}} = Y(t, f) \Big|_{t=nT, f=m\Delta f} \quad (3.5-36)$$

$$Y^{\text{DD}}[k, l] = \frac{1}{NM} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} Y^{\text{TF}}[m, n] e^{-j2\pi(\frac{nk}{N} - \frac{ml}{M})} \quad (3.5-37)$$

其中, $r(t)$ 表示 OTFS 系统接收端接收的时域信号, $g_{\text{rx}}^*(t)$ 表示接收滤波器的冲激响应, $Y(t, f)$ 和 Y^{TF} 分别表示时间—频率域接收连续信号和离散采样信号, 式(3.5-35)也称为温格变换。得到时间—频率域的离散信号后, 采用如式(3.5-37)中所示的 SFFT 操作可以得到时延—多普勒域接收信号。

通过上述公式(3.5-35)至(3.5-37)所示的 SFFT 操作和温格变换, 接收端的时域信号被成功转换到时延—多普勒域, 得到离散信号 $Y^{\text{DD}}[k, l]$ 。这一过程完成了 OTFS 解调的关键步骤, 为后续在时延—多普勒域进行信号检测和干扰消除提供了基础。

与在时频域进行操作的 OFDM 技术相比, OTFS 调制技术的根本优势在于其信号处理域的革新。OFDM 技术依赖于时频域的正交性子载波, 其在静态或低速信道下表现良好; 然而, 在高速移动场景下, 显著的多普勒扩展会严重破坏子载波间的正交性, 引发严重的载波间干扰 (ICI), 同时多径效应也会导致符号间干扰 (ISI), 这使得系统性能在信道快速变化时急剧恶化, 并需要复杂的干扰消除与信道均衡算法。OTFS 则通过将信息符号首先映射到时延—多普勒域, 再利用逆辛有限傅里叶变换 (ISFFT) 将整个符号帧变换至时频域进行传输, 从而从根本上改变了这一局面。其核心机理在于物理信道在时延—多普勒域的响应是稀疏且缓慢变化的, OTFS 通过这种域变换, 将时频域中快速时变的弥散信道转换成了一个在时延—多普勒域中准静态的、稀疏的二维卷积信道。这一转换使得每个信息符号的能



量得以扩展到整个时频帧,从而让 OTFS 能够获取近乎全时频域的分集增益,极大地增强了在深衰落条件下的传输可靠性。

本节详细介绍了 OTFS 调制技术应对高速移动通信挑战的核心原理。其关键在于将信息符号映射并处理在时延—多普勒域,充分利用了该域信道响应的稀疏性和时不变性。这种设计使其能够有效抵抗由高速移动和多径传播共同导致的复杂时频双选衰落效应,显著降低了传统时频域系统所需的信道估计复杂度,并缓解了高速场景下的符号间干扰和载波间干扰。尽管如公式(3.5-38)所示,多径效应仍然会在接收端引入符号间干扰,但由于信道在时延—多普勒域的稀疏性,有效的干扰消除和低复杂度检测算法成为可能。OTFS 技术展现出的这种对高速移动信道的鲁棒性,使其成为未来高速移动通信系统(如 6G、车联网、高速铁路通信)中提升链路可靠性和频谱效率的关键候选技术之一,具有广阔的应用前景和研究价值。

3.6 小结

本章全面梳理了现代通信中各类调制技术的原理与特点。首先,无记忆调制部分涵盖了幅度调制、相位调制、相位与幅度联合调制、频率调制等,详细解析了这些方法的信号特性和传输效率。随后,在有记忆调制部分着重介绍了连续相位频率调制、连续相位调制、最小频移键控(MSK)和高斯最小频移键控(GMSK),突出其在抗噪性能和带宽利用方面的优势。扩频调制部分展示了直接序列扩频与跳频扩频两种技术,说明了其在提高抗干扰能力和保密性方面的重要作用。而在多载波调制中,重点探讨了正交频分复用(OFDM)和正交时频空间调制(OTFS),阐述了它们在提升对抗多径时延扩展和多普勒扩展性能中的作用。通过对上述内容的系统性总结,本章为读者提供了数字调制理论与实际应用的全面认知基础。

3.7 思考题

3-1 已知 4PSK 系统的传输速率为 2 400 b/s,试问:

(a) 4PSK 信号的谱零点带宽和频带利用率($\text{b}/(\text{s} \cdot \text{Hz})$);

(b) 若对基带信号采用 $\alpha=0.4$ 余弦滚降滤波预处理,再进行 4PSK 调制,这时占用的信道带宽和频带利用率为多大?

(c) 若传输带宽不变,而比特率加倍,则调制方式应作何改变?

3-2 $\pi/4$ -QPSK 可以看成偏移 $\pi/4$ 的两个 QPSK 系统。

(a) 画出 $\pi/4$ -QPSK 信号的信号空间图;

(b) 以相应的格雷编码数据比特表示各信号点。

3-3 设有一个 MSK 信号,其码元速率为 1 000 波特,分别用频率 f_1 和 f_0 表示码元“1”和“0”。若 $f_1=1\ 250\ \text{Hz}$,试确定 f_0 的值,并画出三个码元“101”的波形。

3-4 在下列情况下,试求状态网格图中终值相位的状态数目。



- (a) 全响应二进制 CPFSK, 且 $h=2/3$ 或 $3/4$;
 (b) 部分响应($L=3$)二进制 CPFSK, 且 $h=2/3$ 或 $3/4$ 。

3-5 设二进制信息为“0101”, 采用 2FSK 系统传输的码元速率为 1 200Baud, 已调信号的载频分别为 4 800 Hz(对应“1”码)和 2 400 Hz(对应“0”码)。

- (a) 若采用包络检波方式进行解调, 试画出各点时间波形;
 (b) 若采用相干方式进行解调, 试画出各点时间波形。

3-6 计算周期平稳随机过程

$$v(t) = \sum_{n=-\infty}^{\infty} I_n g(t - nT)$$

的功率谱密度的一种方法是在周期 T 上对自相关函数 $R_v(t+\tau, t)$ 求平均, 再对平均自相关函数进行傅里叶变换。另一种方法是添加一个在 $(0, T)$ 上均匀分布的随机变量 Δ , 将周期平稳过程 $v(t)$ 变为平稳过程 $v_\Delta(t)$:

$$v_\Delta(t) = \sum_{n=-\infty}^{\infty} I_n g(t - nT - \Delta)$$

再把 $v(t)$ 的功率谱密度定义为平稳过程 $v_\Delta(t)$ 的自相关函数的傅里叶变换。试求 $v_\Delta(t)$ 自相关函数及其傅里叶变换。

3-7 试证明 16-QAM 可以表示为两个四相恒包络信号的叠加, 并在相加前对每一分量放大, 即

$$s(t) = G(A_n \cos 2\pi f_c t + B_n \sin 2\pi f_c t) + (C_n \cos 2\pi f_c t + D_n \sin 2\pi f_c t)$$

式中, A_n 、 B_n 、 C_n 和 D_n 为统计独立的二进制序列, 其元素取值为 ± 1 ; G 为放大器的增益。试进一步证明该信号等价于

$$s(t) = I_n \cos 2\pi f_c t + Q_n \sin 2\pi f_c t$$

并用 A_n 、 B_n 、 C_n 和 D_n 来确定 I_n 和 Q_n 。

参考文献

- [1] Proakis, J. G., & Salehi, M. Digital Communications (5th ed.) [M]. McGraw-Hill Science/Engineering/Math, 2007.
 [2] 樊昌信, 曹丽娜. 通信原理(第 7 版) [M]. 北京: 国防工业出版社, 2012.
 [3] 罗鹏飞, 张文明. 随机信号分析与处理 [M]. 北京: 清华大学出版社, 2021.
 [4] 钱玲, 谷亚林, 王海青. 信号与系统(第 5 版) [M]. 北京: 电子工业出版社, 2017.
 [5] William F. Gabriel, Special issue on Adaptive Processing Antenna Systems [J], IEEE Transactions on Antennas and Propagation, 1986, 34(3).

第4章 AWGN 信道最佳接收机

4.1 引言

第3章介绍了各种类型的数字调制方法,即将数字信息序列通过发送端调制器映射为信号波形。调制后的信号需要进一步通过信道传输后到达接收机,进行接收解调操作。第1章已说明,通信信道存在各类损耗,包括噪声、衰落、失真、干扰等,而不同通信信道的特征决定了其主要的损耗类型和影响性能的关键因素。其中,加性噪声是信号在各类信道内传输面临的普遍问题。因此,本章将在加性高斯白噪声(Additive White Gaussian Noise, AWGN)信道表征的传输环境下,研究最佳接收机的理论设计和性能特征,并分析各类数字调制系统的可靠性。

4.2 最佳检测理论

4.2.1 矢量 AWGN 信道

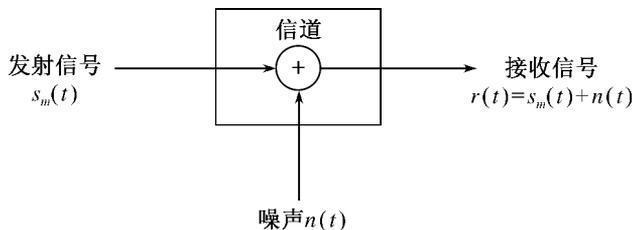


图 4.1 AWGN 信道模型

加性高斯白噪声信道模型用于表示高斯白噪声过程对发送信号的影响,图 4.1 给出了该模型的示意图,其数学描述为

$$r(t) = s_m(t) + n(t) \quad (4.2-1)$$

其中, $s_m(t)$ 是发送信号,当采用 M 进制调制时, $s_m(t)$ 是 M 个可能信号之一,即 $s_m(t) \in s_1(t), s_2(t), \dots, s_M(t)$; $n(t)$ 是高斯白噪声过程的样本波形,其均值为 0,功率谱密度为 $N_0/2$; $r(t)$ 是接收信号。

第3章已说明,利用标准正交基 $\{\phi_j(t), 1 \leq j \leq N\}$,每一个信号 $s_m(t)$ 可以用矢量



$s_m \in \mathbb{R}^N$ 表示。因此,将发射信号的矢量表达式记为 $\{s_m, 1 \leq m \leq M\}$, 由定义有

$$s_m(t) = \sum_{j=1}^N s_{mj} \phi_j(t), \quad s_{mj} = \langle s_m(t), \phi_j(t) \rangle \quad (4.2-2)$$

由于噪声过程 $n(t)$ 不能以基函数 $\{\phi_j(t)\}_{j=1}^N$ 全部展开,我们将其分解为两个分量:一个分量是噪声中能以 $\{\phi_j(t)\}_{j=1}^N$ 展开的部分,记为 $n_1(t)$;另一个分量是噪声中不能以基函数展开的部分,记为 $n_2(t)$ 。由上述定义有

$$n_1(t) = \sum_{j=1}^N n_j \phi_j(t), \quad n_j = \langle n_1(t), \phi_j(t) \rangle \quad (4.2-3)$$

$$n_2(t) = n(t) - n_1(t) \quad (4.2-4)$$

根据式(4.2-3)和式(4.2-4),式(4.2-1)可表示为

$$r(t) = \sum_{j=1}^N (s_{mj} + n_j) \phi_j(t) + n_2(t) \quad (4.2-5)$$

定义变量

$$r_j = s_{mj} + n_j \quad (4.2-6)$$

则有

$$\begin{aligned} r_j &= \langle s_m(t), \phi_j(t) \rangle + \langle n(t), \phi_j(t) \rangle = \langle s_m(t) + n(t), \phi_j(t) \rangle \\ &= \langle r(t), \phi_j(t) \rangle \end{aligned} \quad (4.2-7)$$

故式(4.2-5)可写为

$$r(t) = \sum_{j=1}^N r_j \phi_j(t) + n_2(t), \quad r_j = \langle r(t), \phi_j(t) \rangle \quad (4.2-8)$$

任何标准正交基都可以用于零均值白高斯过程的展开式,展开式的系数(即 n_j)是独立同分布(IID)、均值为0、方差为 $N_0/2$ 的高斯随机变量。证明如下:

n_j 定义为

$$n_j = \int_{-\infty}^{\infty} n(t) \phi_j(t) dt \quad (4.2-9)$$

它是高斯随机过程 $n(t)$ 的积分变换,因此它符合高斯分布,其均值为

$$E[n_j] = E \left[\int_{-\infty}^{\infty} n(t) \phi_j(t) dt \right] = \int_{-\infty}^{\infty} E[n(t)] \phi_j(t) dt = 0 \quad (4.2-10)$$

式中,由于 $n(t)$ 均值为零,故最后的等式成立。

可以求得 n_i 与 n_j 的协方差为

$$\begin{aligned} \text{COV}[n_i, n_j] &= E[n_i n_j] - E[n_i] E[n_j] \\ &= E \left[\int_{-\infty}^{\infty} n(t) \phi_i(t) dt \int_{-\infty}^{\infty} n(s) \phi_j(s) ds \right] \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E[n(t) n(s)] \phi_i(t) \phi_j(s) dt ds \\ &= \frac{N_0}{2} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \delta(t-s) \phi_i(t) dt \right] \phi_j(s) ds \end{aligned}$$



$$\begin{aligned}
 &= \frac{N_0}{2} \left[\int_{-\infty}^{\infty} \phi_i(s) \phi_j(s) ds \right] \\
 &= \begin{cases} N_0/2 & i=j \\ 0 & i \neq j \end{cases} \quad (4.2-11)
 \end{aligned}$$

式中,利用如下事实: n_i 和 n_j 均值为零; $n(t)$ 自相关函数为 $\frac{N_0}{2} \delta(\tau)$; 最后一步利用了 $\{\phi_j(t)\}$ 的标准正交性。式(4.2-11)说明,当 $i \neq j$ 时, n_i 和 n_j 不相关,又因为它们符合高斯分布,故它们相互独立。这也说明每一个 n_j 的方差均为 $N_0/2$ 。

接下来研究 $n_2(t)$ 的性质:由于 n_j 是联合高斯随机变量, $n_1(t)$ 是高斯过程,则 $n_2(t) = n(t) - n_1(t)$ 是两个联合高斯过程的线性组合,从而 $n_2(t)$ 也是高斯过程。在任何给定 t 时刻,有

$$\begin{aligned}
 \text{COV}[n_j n_2(t)] &= E[n_j n_2(t)] = E[n_j n(t)] - E[n_j n_1(t)] \\
 &= E[n(t) \int_{-\infty}^{\infty} n(s) \phi_j(s) ds] - E[n_j \sum_{i=1}^N n_i \phi_i(t)] \\
 &= \frac{N_0}{2} \int_{-\infty}^{\infty} \delta(t-s) \phi_j(s) ds - \frac{N_0}{2} \phi_j(t) \\
 &= \frac{N_0}{2} \phi_j(t) - \frac{N_0}{2} \phi_j(t) \\
 &= 0
 \end{aligned} \quad (4.2-12)$$

式中,利用这样的事实: $E[n_j n_i] = 0$ (除 $i=j$ 外); 当 $i=j$ 时, $E[n_j n_i] = N_0/2$ 。

式(4.2-12)表明, $n_2(t)$ 与所有的 n_j 都不相关。并且,由于它们服从联合高斯分布, $n_2(t)$ 与所有的 n_j 都独立,因此 $n_2(t)$ 与 $n_1(t)$ 独立。

因为 $n_2(t)$ 独立于 $s_m(t)$ 和 $n_1(t)$, 可以推断式(4.2-8)中 $r(t)$ 的两个分量,即 $\sum_j r_j \phi_j(t)$ 和 $n_2(t)$ 是独立的。由于第一个分量是携带发送信号的唯一分量,且第二个分量独立于第一个分量,第二个分量不能提供有关发送信号的任何信息。因此, $n_2(t)$ 对最佳检测而言是无关信息,即可以忽略 $n_2(t)$ 而不损失检测器的最佳性。

由上述讨论可得,式(4.2-1)表征的 AWGN 信道模型可等效视为 N 维矢量信道模型, N 为信号矢量的维度,如下式所示

$$\mathbf{r} = \mathbf{s}_m + \mathbf{n}, \quad 1 \leq m \leq M \quad (4.2-13)$$

其中 $\mathbf{r} = (r_1, r_2, \dots, r_N)$, $\mathbf{s}_m = (s_{m1}, s_{m2}, \dots, s_{mN})$, $\mathbf{n} = (n_1, n_2, \dots, n_N)$ 。

4.2.2 最佳检测准则

首先考虑一般的矢量信道模型,模型示意图如图 4.2 所示。模型中,发射矢量 \mathbf{s}_m 是按先验概率 P_m 从信号集 $\{\mathbf{s}_m, 1 \leq m \leq M\}$ 中选取的,其通过信道传输。条件概率密度函数 $p(\mathbf{r}|\mathbf{s}_m)$ 给出了在发射矢量为 \mathbf{s}_m 的情况下,接收矢量 \mathbf{r} 的统计分布。

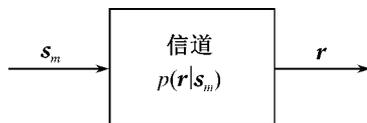


图 4.2 一般矢量信道模型



接收机根据接收矢量 \mathbf{r} 判断发送的是哪个消息(即 m 的值)。定义接收机检测器的判决函数为 $g(\mathbf{r})$, $\hat{m} = g(\mathbf{r})$ 为最终判决结果, 则当判决正确时有 $\hat{m} = m$ 。检测器通常将接收信号向量空间 \mathbb{R}^N 划分为 M 个域, 记为 D_1, D_2, \dots, D_M , 如果 $\mathbf{r} \in D_{m'}$, 则 $\hat{m} = g(\mathbf{r}) = m'$, 即检测器的判决结果为 m' 。域 $D_{m'} (1 \leq m' \leq M)$ 称为消息 m' 的判决域, 它是被检测器映射为消息 m' 的所有信道输出的集合。

1. 最小差错概率(Minimum probability of error rate)准则

当发射机发送 s_m 而接收 $\mathbf{r} \notin D_m$ 时, 检测器判决结果 $\hat{m} \neq m$, 发生差错。因此, 接收机的符号差错概率可表示为

$$P_e = \sum_{m=1}^M P_m P[\mathbf{r} \notin D_m | s_m] = \sum_{m=1}^M P_m P_{e|m} \quad (4.2-14)$$

式中, P_m 是发送的消息为 m 的概率, $P_{e|m}$ 表示发送消息 m 时的差错概率, 即

$$P_{e|m} = \sum_{\substack{1 \leq m' \leq M \\ m' \neq m}} \int_{D_{m'}} p(\mathbf{r} | s_m) d\mathbf{r} \quad (4.2-15)$$

将式(4.2-15)代入式(4.2-14), 得

$$P_e = \sum_{m=1}^M P_m \sum_{\substack{1 \leq m' \leq M \\ m' \neq m}} \int_{D_{m'}} p(\mathbf{r} | s_m) d\mathbf{r} \quad (4.2-16)$$

最小差错概率准则即使 P_e 最小的准则。

2. 最大后验概率(Maximum a posteriori probability, MAP)准则

最佳检测器应使差错概率最小, 或等价为使正确判决的概率最大。当发射机发送 s_m 而接收 $\mathbf{r} \in D_m$ 时, 检测器判决结果 $\hat{m} = m$, 即判决正确。因此, 正确判决的概率为

$$P_c = \sum_{m=1}^M P_m \int_{D_m} p(\mathbf{r} | s_m) d\mathbf{r} \quad (4.2-17)$$

由贝叶斯法则, 有 $P_m p(\mathbf{r} | s_m) = p(\mathbf{r}) P(s_m | \mathbf{r})$, 故正确判决概率也可写为

$$P_c = \sum_{m=1}^M \int_{D_m} p(\mathbf{r}) P(s_m | \mathbf{r}) d\mathbf{r} \quad (4.2-18)$$

式(4.2-18)中, 因为 $p(\mathbf{r}) \geq 0$, 若对每一个 \mathbf{r} 最大化 $P(s_m | \mathbf{r})$, 则可以最大化正确判决概率。这意味着接收机可以根据对 \mathbf{r} 的观察选取使 $P(s_m | \mathbf{r})$ 最大的消息 \hat{m} 判决, 实现最佳检测, 该准则称为最大后验概率准则。判决结果 \hat{m} 可表示为

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} P(s_m | \mathbf{r}) \quad (4.2-19)$$

由贝叶斯法则, 式(4.2-19)可改写为

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} \frac{P_m p(\mathbf{r} | s_m)}{p(\mathbf{r})} \quad (4.2-20)$$

因为 $p(\mathbf{r})$ 独立于 m 且对所有 m 都一样, 该式等价于

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} P_m p(\mathbf{r} | s_m) \quad (4.2-21)$$



式(4.2-21)说明,判决结果 \hat{m} 是由先验概率 P_m 和信道的概率描述 $p(\mathbf{r}|\mathbf{s}_m)$ 确定的,这两者是已知的。

当检测器采用最大后验概率准则时(称为 MAP 检测器),其最佳判决域 D_m 将使差错概率降到最低,最佳判决域可表示为

$$D_m = \{\mathbf{r} \in \mathbb{R}^N : P[\mathbf{r} \in D_m | \mathbf{r}] > P[\mathbf{r} \in D_{m'} | \mathbf{r}], 1 \leq m' \leq M \text{ 且 } m' \neq m\} \quad (4.2-22)$$

3. 极大似然(Maximum-likelihood, ML)准则

在发送信息先验概率相等的情况下,即当 $P_m = 1/M (1 \leq m \leq M)$ 时,式(4.2-21)表示的判决规则可简化为

$$\hat{m} = \underset{1 \leq m \leq M}{\operatorname{argmax}} p(\mathbf{r} | \mathbf{s}_m) \quad (4.2-23)$$

式中, $p(\mathbf{r} | \mathbf{s}_m)$ 称为消息 m 的似然函数。式(4.2-23)确定的检测准则称为极大似然准则。采用该准则的检测器称为 ML 检测器。

注意,如果发送消息的先验概率不相等,ML 检测器就不是最佳检测器。然而,ML 检测器是很常用的检测器,因为在大多数情况下接收机难以准确获取消息的先验概率,另一方面在发送消息前可以按照某随机序列进行信息置乱处理,近似满足等概率发送的要求。

4.2.3 矢量 AWGN 信道的最佳检测

式(4.2-13)已经给出了 AWGN 矢量信道的数学模型。式中,所有矢量均有 N 维;噪声分量 $n_j (1 \leq n_j \leq N)$ 是 IID 高斯随机变量且服从 $N(0, N_0/2)$ 分布。因此,噪声矢量 \mathbf{n} 的 PDF 为

$$p(\mathbf{n}) = \left(\frac{1}{\sqrt{\pi N_0}} \right)^N e^{-\frac{\sum_{j=1}^N n_j^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{\pi N_0}} \right)^N e^{-\frac{\|\mathbf{n}\|^2}{N_0}} \quad (4.2-24)$$

根据式(4.2-21)和式(4.2-24),可以将矢量 AWGN 信道下的 MAP 检测器表示为

$$\begin{aligned} \hat{m} &= \underset{1 \leq m \leq M}{\operatorname{argmax}} [P_m p(\mathbf{r} | \mathbf{s}_m)] = \underset{1 \leq m \leq M}{\operatorname{argmax}} P_m [p_n(\mathbf{r} - \mathbf{s}_m)] \\ &= \underset{1 \leq m \leq M}{\operatorname{argmax}} \left[P_m \left(\frac{1}{\sqrt{\pi N_0}} \right)^N e^{-\frac{\|\mathbf{r} - \mathbf{s}_m\|^2}{N_0}} \right] \stackrel{(a)}{=} \underset{1 \leq m \leq M}{\operatorname{argmax}} [P_m e^{-\frac{\|\mathbf{r} - \mathbf{s}_m\|^2}{N_0}}] \\ &\stackrel{(b)}{=} \underset{1 \leq m \leq M}{\operatorname{argmax}} \left[\ln P_m - \frac{\|\mathbf{r} - \mathbf{s}_m\|^2}{N_0} \right] \stackrel{(c)}{=} \underset{1 \leq m \leq M}{\operatorname{argmax}} \left[\frac{N_0}{2} \ln P_m - \frac{1}{2} \|\mathbf{r} - \mathbf{s}_m\|^2 \right] \\ &= \underset{1 \leq m \leq M}{\operatorname{argmax}} \left[\frac{N_0}{2} \ln P_m - \frac{1}{2} (\|\mathbf{r}\|^2 + \|\mathbf{s}_m\|^2 - 2\mathbf{r} \cdot \mathbf{s}_m) \right] \\ &\stackrel{(d)}{=} \underset{1 \leq m \leq M}{\operatorname{argmax}} \left[\frac{N_0}{2} \ln P_m - \frac{1}{2} \epsilon_m + \mathbf{r} \cdot \mathbf{s}_m \right] \stackrel{(e)}{=} \underset{1 \leq m \leq M}{\operatorname{argmax}} [\eta_m + \mathbf{r} \cdot \mathbf{s}_m] \quad (4.2-25) \end{aligned}$$

式中应用的简化公式如下:

- (a) $1/\sqrt{\pi N_0}$ 是正常数,可以舍去。
- (b) $\ln(\cdot)$ 是增函数。
- (c) $N_0/2$ 是正的,且与正数相乘并不影响 argmax 的结果。
- (d) $\|\mathbf{r}\|^2$ 可以舍去,因为它与 m 无关,且 $\|\mathbf{s}_m\|^2 = \epsilon_m$ 。



(e) 定义偏差项为

$$\eta_m = \frac{N_0}{2} \ln P_m - \frac{1}{2} \mathcal{E}_m \quad (4.2-26)$$

由式(4.2-25)可见, 矢量 AWGN 信道的 MAP 判决规则为

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} [\eta_m + \mathbf{r} \cdot \mathbf{s}_m] \quad (4.2-27)$$

对应的判决域为

$$D_m = \{\mathbf{r} \in \mathbb{R}^N : \mathbf{r} \cdot \mathbf{s}_m + \eta_m > \mathbf{r} \cdot \mathbf{s}_{m'} + \eta_{m'}, 1 \leq m' \leq M \text{ 且 } m' \neq m\} \quad (4.2-28)$$

由于实际接收到的信号均为实信号, 则有

$$\mathbf{r} \cdot \mathbf{s}_m = \int_{-\infty}^{\infty} r(t) s_m(t) dt \quad (4.2-29)$$

$$\mathcal{E}_m = \|\mathbf{s}_m\|^2 = \int_{-\infty}^{\infty} s_m^2(t) dt \quad (4.2-30)$$

式(4.2-27)等价于

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} \left[\frac{N_0}{2} \ln P_m + \int_{-\infty}^{\infty} r(t) s_m(t) dt - \frac{1}{2} \int_{-\infty}^{\infty} s_m^2(t) dt \right] \quad (4.2-31)$$

式(4.2-27)为 MAP 判决规则的向量表示形式, 式(4.2-31)为其波形表示形式。

当发送信号先验概率相等时, 由于 $P_m = 1/M$, 式(4.2-25)中的步骤(c)可写为

$$\begin{aligned} \hat{m} &= \operatorname{argmax}_{1 \leq m \leq M} \left[\frac{N_0}{2} \ln P_m - \frac{1}{2} \|\mathbf{r} - \mathbf{s}_m\|^2 \right] \\ &= \operatorname{argmax}_{1 \leq m \leq M} [-\|\mathbf{r} - \mathbf{s}_m\|^2] \\ &= \operatorname{argmin}_{1 \leq m \leq M} \|\mathbf{r} - \mathbf{s}_m\| \end{aligned} \quad (4.2-32)$$

对式(4.2-32)可以有相当直观的几何解释: 接收机接收 \mathbf{r} , 并以标准欧氏距离为参考, 在所有的 \mathbf{s}_m 中寻找与 \mathbf{r} 距离最近者, 这样的检测器被称为最近邻(Nearest-Neighbor)或最小距离(Minimum-Distance)检测器。这也说明, 当发送信号等概率时, MAP 和 ML 检测器均等价于最小距离检测器。

式(4.2-32)为 ML 判决规则的向量表示形式, 其波形表示形式可写为

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} \left[\int_{-\infty}^{\infty} r(t) s_m(t) dt - \frac{1}{2} \int_{-\infty}^{\infty} s_m^2(t) dt \right] \quad (4.2-33)$$

当信号等概且等能量时, 偏差项 $\eta_m = N_0 \ln P_m / 2 - \mathcal{E}_m / 2$ 独立于 m , 可以从式(4.2-27)中舍去。此时, 最佳判决规则可进一步简化为

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} \mathbf{r} \cdot \mathbf{s}_m \quad (4.2-34)$$

对应的判决域为

$$D_m = \{\mathbf{r} \in \mathbb{R}^N : \mathbf{r} \cdot \mathbf{s}_m > \mathbf{r} \cdot \mathbf{s}_{m'}, 1 \leq m' \leq M \text{ 且 } m' \neq m\} \quad (4.2-35)$$

为了方便表示, 引入三种度量。



(1) 距离度量(Distance Metric)

$$D(\mathbf{r}, \mathbf{s}_m) = \|\mathbf{r} - \mathbf{s}_m\|^2 = \int_{-\infty}^{\infty} [r(t) - s_m(t)]^2 dt \quad (4.2-36)$$

它表示 \mathbf{r} 和 \mathbf{s}_m 之间欧式距离的平方。

(2) 修正距离度量(Modified Distance Metric)

$$D'(\mathbf{r}, \mathbf{s}_m) = -2\mathbf{r} \cdot \mathbf{s}_m + \|\mathbf{s}_m\|^2 \quad (4.2-37)$$

当 $\|\mathbf{r}\|^2$ 与 m 无关而被舍去后,它等于距离度量。

(3) 相关度量(Correlation Metric)

$$C(\mathbf{r}, \mathbf{s}_m) = 2\mathbf{r} \cdot \mathbf{s}_m - \|\mathbf{s}_m\|^2 = 2 \int_{-\infty}^{\infty} r(t)s_m(t) dt - \int_{-\infty}^{\infty} s_m^2(t) dt \quad (4.2-38)$$

根据上述定义,MAP 检测规则可表示为

$$\begin{aligned} \hat{m} &= \operatorname{argmax}_{1 \leq m \leq M} [N_0 \ln P_m - D(\mathbf{r}, \mathbf{s}_m)] \\ &= \operatorname{argmax}_{1 \leq m \leq M} [N_0 \ln P_m + C(\mathbf{r}, \mathbf{s}_m)] \end{aligned} \quad (4.2-39)$$

ML 检测规则可表示为

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} C(\mathbf{r}, \mathbf{s}_m) \quad (4.2-40)$$

4.3 最佳接收机设计

4.3.1 相关接收机

假设 AWGN 信道下的最佳接收机采用式(4.2-27)表示的 MAP 判决规则,即

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} [\eta_m + \mathbf{r} \cdot \mathbf{s}_m], \quad \eta_m = \frac{N_0}{2} \ln P_m - \frac{1}{2} \mathcal{E}_m \quad (4.3-1)$$

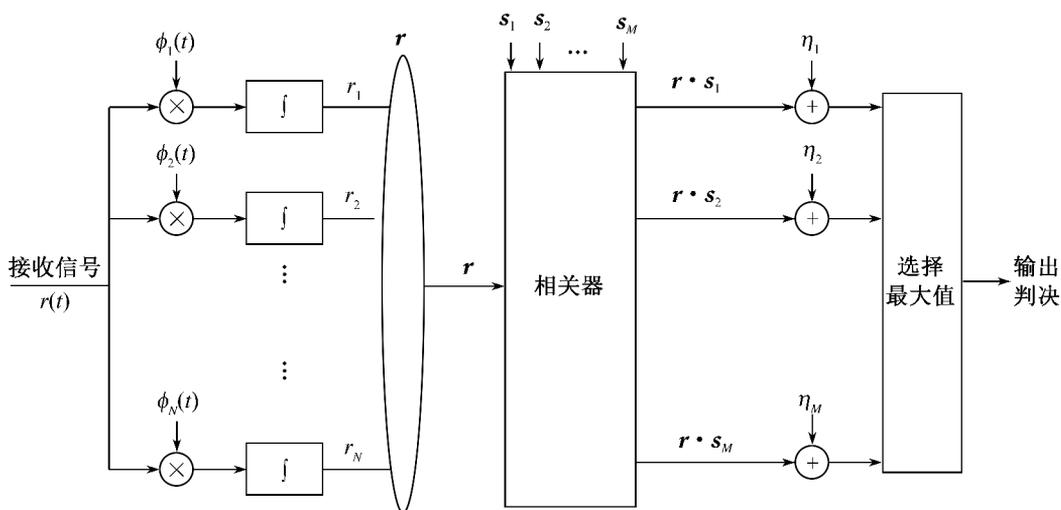
相应的判决过程如下:首先,接收机的输入为时域信号 $r(t)$,而非接收矢量 \mathbf{r} ,故第一步需要由 $r(t)$ 导出 \mathbf{r} ,可利用如下关系式

$$r_j = \int_{-\infty}^{\infty} r(t) \phi_j(t) dt \quad (4.3-2)$$

上式说明,利用每一个基函数 $\phi_j(t)$ 乘 $r(t)$ 后积分可得到 \mathbf{r} 的全部分量。之后,可以求得 \mathbf{r} 与每一个 \mathbf{s}_m 的内积,并加上对应的偏差项 η_m 。最后,比较并选择使计算值最大的 m 作为判决结果。因为接收信号与每一个 $\phi_j(t)$ 相关,这种最佳接收机的实现方法称为相关接收机(Correlation Receiver)。

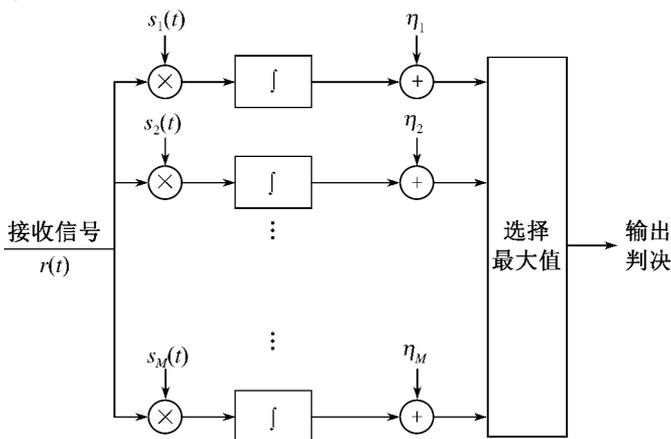
具有 N 个相关器的相关接收机结构如图 4.3 所示。注意, η_m 和 \mathcal{E}_m 独立于接收信号 $r(t)$,因此它们可以只计算一次并存储在存储器中,便于后续读取。图中需要计算的部分是求解 $\mathbf{r} \cdot \mathbf{s}_m$ 的相关器。

同样的,AWGN 信道最佳接收机也可以采用式(4.2-31)表示的 MAP 判决规则,即

图 4.3 具有 N 个相关器的相关接收机结构

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} \left[\eta_m + \int_{-\infty}^{\infty} r(t) s_m(t) dt \right], \eta_m = \frac{N_0}{2} \ln P_m - \frac{1}{2} \mathcal{E}_m \quad (4.3-3)$$

通过对 $r(t)$ 和 $s_m(t)$ 作相关运算,可以直接求得 $\langle r(t), s_m(t) \rangle$ 。随后,每一计算值加上对应的偏差项并比较大小,最后得到判决结果 m 。图 4.4 给出了该实现方法的接收机结构,它是相关接收机的第二种形式。

图 4.4 具有 M 个相关器的相关接收机结构

注意,虽然第二种形式的相关接收机结构看起来比第一种简单,但因为大多数情况下基函数个数少于发射信号种类(即 $N < M$,实际上通常有 $N \ll M$,比如 MPSK、MQAM 信号等),故第一种相关接收机结构更为常见。

相关接收机需要 N 或 M 个相关器(乘法器+积分器)实现。下面阐述最佳接收机的另一种实现。



4.3.2 匹配滤波器接收机

1. 接收机结构

在两种相关接收机的实现方法中,需要对信号进行相关运算,即

$$r_x = \int_{-\infty}^{+\infty} r(t)x(t)dt \quad (4.3-4)$$

其中, $x(t)$ 是 $\phi_j(t)$ 或 $s_m(t)$ 。

定义 $h(t) = x(T-t)$, 其中 T 为任意值, 具有冲激响应 $h(t)$ 的滤波器称为匹配于 $x(t)$ 的滤波器或匹配滤波器。如果该滤波器的输入为 $r(t)$, 则其输出 $y(t)$ 为 $r(t)$ 与 $h(t)$ 的卷积, 即

$$\begin{aligned} y(t) &= r(t) * h(t) = \int_{-\infty}^{\infty} r(\tau)h(t-\tau)d\tau \\ &= \int_{-\infty}^{\infty} r(\tau)x(T-t+\tau)d\tau \end{aligned} \quad (4.3-5)$$

由式(4.3-5), 可得

$$r_x = y(T) = \int_{-\infty}^{\infty} r(\tau)x(\tau)d\tau \quad (4.3-6)$$

上式说明, 相关器的输出 r_x 可以通过匹配滤波器在 $t=T$ 时刻抽样得到。注意, 抽样时刻必须准确地设置为 $t=T$, 其中 T 在匹配滤波器的设计中取任意值。只要满足这个条件, T 的选择是不相关的; 然而从实际情况来看, T 的选择原则是必须使滤波器符合因果关系, 即当 $t < 0$ 时 $h(t) = 0$, 这就要对 T 的可能取值加以限制。

将图 4.3 中与基函数相关器替换为匹配滤波器, 其中匹配滤波器与每一个基函数匹配, 可得到最佳接收机的匹配滤波器实现结构, 如图 4.5 所示。

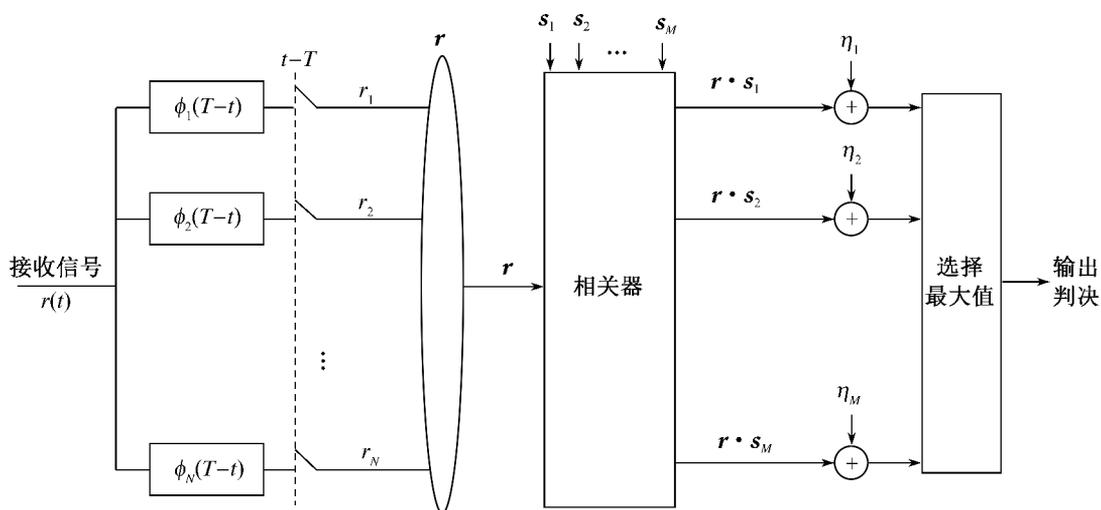


图 4.5 具有 N 个匹配滤波器的接收机实现结构



另一种最佳接收器的匹配滤波实现方法类似于图 4.4, 其中 M 个匹配滤波器分别匹配于发射信号 $\{s_m(t), 1 \leq m \leq M\}$, 具体结构如图 4.6 所示。

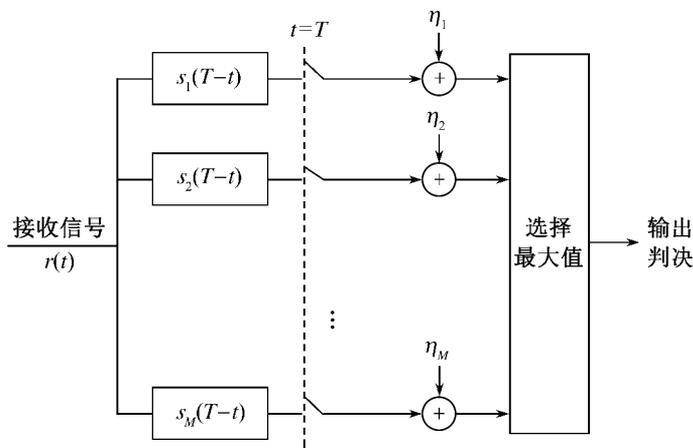


图 4.6 具有 M 个匹配滤波器的接收机实现结构

2. 频域解释

假设发送信号是 $s(t)$, 匹配于信号 $s(t)$ 的匹配滤波器可以从频域角度进行解释。由 $h(t) = s(T-t)$, 利用基本傅里叶变换性质, 可得该关系式的傅里叶变换为

$$H(f) = S^*(f) e^{-j2\pi fT} \quad (4.3-7)$$

式(4.3-7)说明, 匹配滤波器的频率响应是发送信号频谱的共轭乘以相位因子 $e^{-j2\pi fT}$, 该相位因子表示抽样延时 T 。换言之, 匹配滤波器的幅度响应与发送信号幅度谱相同, 即 $|H(f)| = |S(f)|$ 。另一方面, $|H(f)|$ 是 $|S(f)|$ 相移 $-2\pi fT$ 的结果。

假如信号只受到 AWGN 的影响, 其经过匹配滤波器后的输出信噪比(信号均方值和噪声方差的比)将在抽样时刻达到最大, 下面给出证明:

假设 $r(t) = s(t) + n(t)$ 通过一个冲激响应为 $h(t)$ 、频率响应为 $H(f)$ 的滤波器, 对应输出 $y(t) = y_s(t) + v(t)$ 在 T 时刻抽样, 输出由信号部分 $y_s(t)$ 和噪声部分 $v(t)$ 组成, 前者的傅里叶变换为 $H(f)S(f)$, 后者的功率谱密度为 $\frac{N_0}{2} |H(f)|^2$ 。 T 时刻信号部分的抽样值为

$$y_s(T) = \int_{-\infty}^{\infty} H(f)S(f)e^{j2\pi fT} df \quad (4.3-8)$$

零均值高斯分量 $v(T)$ 的方差为

$$\text{VAR}[v(T)] = \frac{N_0}{2} \int_{-\infty}^{\infty} |H(f)|^2 df = \frac{N_0}{2} \mathcal{E}_h \quad (4.3-9)$$

式中, \mathcal{E}_h 为 $h(t)$ 的能量。定义滤波器 $H(f)$ 输出端的信噪比为

$$\text{SNR}_o = \frac{y_s^2(T)}{\text{VAR}[v(T)]} \quad (4.3-10)$$



由柯西-施瓦兹不等式,有

$$\begin{aligned} y_s(T) &= \int_{-\infty}^{\infty} H(f)S(f)e^{j2\pi fT} df \\ &\leq \int_{-\infty}^{\infty} |H(f)|^2 df \cdot \int_{-\infty}^{\infty} |S(f)e^{j2\pi fT}|^2 df \\ &= \mathcal{E}_h \mathcal{E}_s \end{aligned} \quad (4.3-11)$$

\mathcal{E}_s 为 $s(t)$ 的能量, 当且仅当存在复常数 α 使 $H(f) = \alpha S^*(f)e^{-j2\pi fT}$ 时, 不等式变为等式。在式(4.3-10)中利用式(4.3-11)可得

$$\text{SNR}_o \leq \frac{\mathcal{E}_s \mathcal{E}_h}{\frac{N_0}{2} \mathcal{E}_h} = \frac{2\mathcal{E}_s}{N_0} \quad (4.3-12)$$

这表明要使滤波器 $H(f)$ 的输出信噪比最大化, 必须满足关系式 $H(f) = S^*(f)e^{-j2\pi fT}$, 即它是一个匹配滤波器。同时, 上式说明最大可能输出信噪比为 $2\mathcal{E}_s/N_0$ 。

4.4 典型调制差错性能分析

4.4.1 二进制信号

1. 双极性信号

二进制幅度调制信号(PAM)和二进制相位调制信号(PSK)均为双极性信号, 即 $s_1(t) = s(t)$, $s_2(t) = -s(t)$ 。由于信号是一维的, 对应的矢量表示只包含一个分量, 可写为标量的形式, 即 $s_1 = \sqrt{\mathcal{E}_s}$, $s_2 = -\sqrt{\mathcal{E}_s}$ 。其中, $\mathcal{E}_s = \mathcal{E}_b$, 即信号能量等于比特能量。

假设消息 1 和消息 2 的概率分别为 p 和 $1-p$ 。由式(4.2-28)可得判决域 D_1 为

$$\begin{aligned} D_1 &= \left\{ r : r \sqrt{\mathcal{E}_b} + \frac{N_0}{2} \ln p - \frac{1}{2} \mathcal{E}_b > -r \sqrt{\mathcal{E}_b} + \frac{N_0}{2} \ln(1-p) - \frac{1}{2} \mathcal{E}_b \right\} \\ &= \left\{ r : r > \frac{N_0}{4\sqrt{\mathcal{E}_b}} \ln \frac{1-p}{p} \right\} = \{ r : r > r_{\text{th}} \} \end{aligned} \quad (4.4-1)$$

其中, 门限 r_{th} 定义为

$$r_{\text{th}} = \frac{N_0}{4\sqrt{\mathcal{E}_b}} \ln \frac{1-p}{p} \quad (4.4-2)$$

类似的可得判决域 D_2 为

$$D_2 = \{ r : r < r_{\text{th}} \} \quad (4.4-3)$$

星座图和判决域如图 4.7 所示。

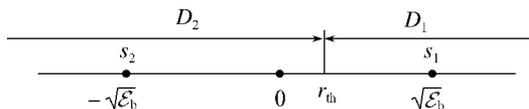


图 4.7 双极性信号的星座图与判决域



利用式(4.2-16),可以推导出该系统的符号差错概率为

$$\begin{aligned}
 P_e &= \sum_{m=1}^2 P_m \sum_{\substack{1 \leq m' \leq 2 \\ m' \neq m}} \int_{D'_m} p(\mathbf{r} | \mathbf{s}_m) d\mathbf{r} \\
 &= p \int_{D_2} p(r | s = \sqrt{\mathcal{E}_b}) dr + (1-p) \int_{D_1} p(r | s = -\sqrt{\mathcal{E}_b}) dr \\
 &= p \int_{-\infty}^{r_{th}} p(r | s = \sqrt{\mathcal{E}_b}) dr + (1-p) \int_{r_{th}}^{\infty} p(r | s = -\sqrt{\mathcal{E}_b}) dr \\
 &= pP[N(\sqrt{\mathcal{E}_b}, N_0/2) < r_{th}] + (1-p)P[N(-\sqrt{\mathcal{E}_b}, N_0/2) > r_{th}] \\
 &= pQ\left(\frac{\sqrt{\mathcal{E}_b} - r_{th}}{\sqrt{N_0/2}}\right) + (1-p)Q\left(\frac{r_{th} + \sqrt{\mathcal{E}_b}}{\sqrt{N_0/2}}\right)
 \end{aligned} \tag{4.4-4}$$

式中, Q 函数是标准正态分布的拖尾概率, Q 函数定义为:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{u^2}{2}\right) du \tag{4.4-5}$$

当发送消息等概时($p=1/2$ 时), $r_{th}=0$, 判决规则可简化为最小距离规则。此时符号差错概率可简化为

$$P_e = Q\left(\sqrt{\frac{2\mathcal{E}_b}{N_0}}\right) \tag{4.4-6}$$

注意到信号之间的距离 d_{12} 满足

$$d_{12}^2 = \int_{-\infty}^{+\infty} |s_1(t) - s_2(t)|^2 dt = 4\mathcal{E}_b \tag{4.4-7}$$

则当消息等概时, 式(4.4-6)也可表示为

$$P_e = Q\left(\sqrt{\frac{d_{12}^2}{2N_0}}\right) \tag{4.4-8}$$

上式说明, 信号的距离越大(即发射信号的功率越大), 检测的符号差错概率就越小。

2. 二进制正交信号

对于二进制正交信号 $s_1(t)$ 、 $s_2(t)$, 有

$$\int_{-\infty}^{\infty} s_1(t)s_2(t) dt = 0 \tag{4.4-9}$$

由于信号是二进制的, 比特能量等于符号能量, 即

$$\mathcal{E}_b = \mathcal{E} = \int_{-\infty}^{\infty} s_1^2(t) dt = \int_{-\infty}^{\infty} s_2^2(t) dt \tag{4.4-10}$$

选择 $\phi_j(t) = s_j(t)/\sqrt{\mathcal{E}_b}$ ($j=1, 2$) 作为基函数, 则信号集的矢量形式表示为

$$\mathbf{s}_1 = (\sqrt{\mathcal{E}_b}, 0), \mathbf{s}_2 = (0, \sqrt{\mathcal{E}_b}) \tag{4.4-11}$$

易知接收信号 $\mathbf{r} = \mathbf{s}_m + \mathbf{n}$, 其中 $\mathbf{r} = (r_1, r_2)$, $\mathbf{n} = (n_1, n_2)$ 。所以有 $\mathbf{n} = \mathbf{r} - \mathbf{s}_m$ 。假设发送信号等概, 由式(4.2-35)可得判决域 D_1 为



$$D_1 = \{\mathbf{r} : r s_1 > r s_2\} = \{\mathbf{r} : r_1 \sqrt{\mathcal{E}_b} > r_2 \sqrt{\mathcal{E}_b}\} \quad (4.4-12)$$

$$= \{\mathbf{r} : r_2 - r_1 > 0\} = \{\mathbf{n} : n_2 - n_1 > \sqrt{\mathcal{E}_b}\}$$

类似可得判决域 D_2 为

$$D_2 = \{\mathbf{n} : n_2 - n_1 < -\sqrt{\mathcal{E}_b}\} \quad (4.4-13)$$

由 $n_1 \sim N(0, \frac{N_0}{2})$, $n_2 \sim N(0, \frac{N_0}{2})$, 有 $n_2 - n_1 \sim N(0, N_0)$ 。利用式(4.4-4), 可以推导出该系统的符号差错概率为

$$P_e = \sum_{m=1}^2 P_m \sum_{\substack{1 \leq m' \leq 2 \\ m' \neq m}} \int_{D_m} p(\mathbf{r} | \mathbf{s}_m) d\mathbf{r}$$

$$= \frac{1}{2} \int_{D_2} p(\mathbf{r} | \mathbf{s} = (\sqrt{\mathcal{E}_b}, 0)) d\mathbf{r} + \frac{1}{2} \int_{D_1} p(\mathbf{r} | \mathbf{s} = (0, \sqrt{\mathcal{E}_b})) d\mathbf{r} \quad (4.4-14)$$

$$= \frac{1}{2} P[N(0, N_0) < -\sqrt{\mathcal{E}_b}] + \frac{1}{2} P[N(0, N_0) > \sqrt{\mathcal{E}_b}]$$

$$= Q\left(\sqrt{\frac{\mathcal{E}_b}{N_0}}\right)$$

注意到信号之间的距离 d_{12} 满足

$$d_{12}^2 = \int_{-\infty}^{+\infty} [s_1(t) - s_2(t)]^2 dt = 2\mathcal{E}_b \quad (4.4-15)$$

则式(4.4-14)也可表示为

$$P_e = Q\left(\sqrt{\frac{d_{12}^2}{2N_0}}\right) \quad (4.4-16)$$

与双极性信号相比, 在相同差错概率条件下, 双极性信号的比特信噪比要求是二进制正交信号的 $\frac{1}{2}$, 即二者相差 3 dB, 原因是双极性信号星座点之间的距离更大。图 4.8 给出了二进制 PAM 信号(即双极性信号)和二进制正交信号的差错概率曲线。

3. 一般的二进制信号

对于一般的二进制信号 $s_1(t)$, $s_2(t)$, 假设信号等概且等能量, 首先定义相关系数为

$$\rho_r = \frac{1}{\mathcal{E}_b} \int_0^T s_1(t) s_2(t) dt \quad (4.4-17)$$

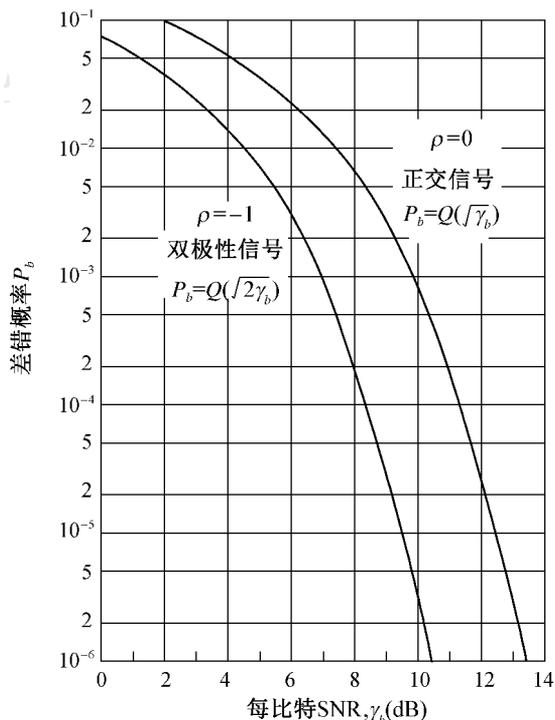


图 4.8 双极性信号与二进制正交信号的差错概率曲线



容易看出,对双极性信号而言, $\rho_r = -1$;对二进制正交信号而言, $\rho_r = 0$ 。

由于 $s_1(t)$ 与 $s_2(t)$ 不一定正交,无法直接求得它们的向量表示,故下面从波形相关角度分析。对于接收信号 $r(t) = s_m(t) + n(t)$, $m = 1, 2$, 可以采用如图 4.9 所示的接收机形式解调:

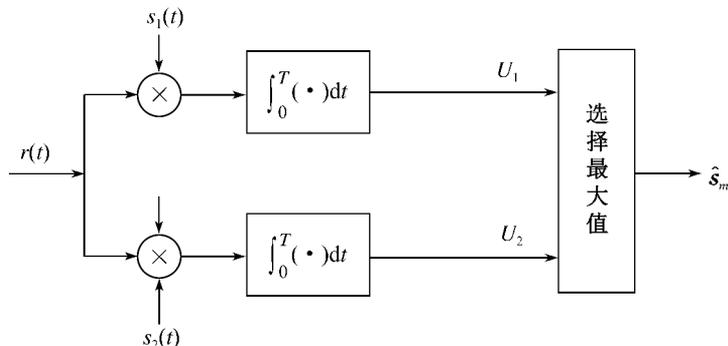


图 4.9 一般二进制信号的接收机结构

假设发送信号 $s_1(t)$, 则两路相关器的输出为

$$\begin{aligned} U_1 &= \int_0^T r(t)s_1(t)dt = \mathcal{E}_b + n_1, n_1 = \int_0^T n(t)s_1(t)dt \\ U_2 &= \int_0^T r(t)s_2(t)dt = \rho_r \mathcal{E}_b + n_2, n_2 = \int_0^T n(t)s_2(t)dt \end{aligned} \quad (4.4-18)$$

且 n_1, n_2 满足

$$\begin{aligned} E[n_i^2] &= \int_0^T \int_0^T E[n(t_1)n(t_2)]s_i(t_1)s_i(t_2)dt_1dt_2 \\ &= \frac{N_0}{2} \int_0^T \int_0^T \delta(t_1 - t_2)s_i(t_1)s_i(t_2)dt_1dt_2 \\ &= \frac{N_0}{2} \int_0^T s_i^2(t)dt = \frac{N_0}{2} \mathcal{E}_b \end{aligned} \quad (4.4-19)$$

$$\begin{aligned} E[n_1n_2] &= \int_0^T \int_0^T E[n(t_1)n(t_2)]s_1(t_1)s_2(t_2)dt_1dt_2 \\ &= \frac{N_0}{2} \int_0^T \int_0^T \delta(t_1 - t_2)s_1(t_1)s_2(t_2)dt_1dt_2 \\ &= \frac{N_0}{2} \int_0^T s_1(t)s_2(t)dt = \frac{N_0}{2} \rho_r \mathcal{E}_b \end{aligned} \quad (4.4-20)$$

当 $U_1 < U_2$ 时发生判决错误,即

$$\begin{aligned} P[\text{error} | s_1(t)] &= P[U_1 < U_2 | s_1(t)] \\ &= P[n_2 - n_1 > \mathcal{E}_b(1 - \rho_r) | s_1(t)] \end{aligned} \quad (4.4-21)$$

记 $n = n_2 - n_1$, 则 n 满足高斯分布,且其均值和方差为

$$E[n] = E[n_2 - n_1] = 0 \quad (4.4-22)$$



$$\begin{aligned}
 \sigma_n^2 &= E[n^2] = E[(n_2 - n_1)^2] \\
 &= E[n_1^2] + E[n_2^2] - 2E[n_1 n_2] \\
 &= N_0 \mathcal{E}_b (1 - \rho_r)
 \end{aligned} \tag{4.4-23}$$

则式(4.4-21)可写为

$$\begin{aligned}
 P[\text{error} | s_1(t)] &= P[n > \mathcal{E}_b (1 - \rho_r)] \\
 &= Q \left[\sqrt{\frac{[\mathcal{E}_b (1 - \rho_r)]^2}{\sigma_n^2}} \right] \\
 &= Q \left[\sqrt{\frac{\mathcal{E}_b (1 - \rho_r)}{N_0}} \right]
 \end{aligned} \tag{4.4-24}$$

由于对称性,发送 $s_1(t)$ 、 $s_2(t)$ 时的判决错误概率相等,即

$$P[\text{error} | s_1(t)] = P[\text{error} | s_2(t)] \tag{4.4-25}$$

根据信号距离的定义,有

$$\begin{aligned}
 d_{12}^2 &= \int_{-\infty}^{\infty} [s_1(t) - s_2(t)]^2 dt \\
 &= \int_{-\infty}^{\infty} s_1^2(t) + s_2^2(t) - 2s_1(t)s_2(t) dt \\
 &= 2\mathcal{E}_b (1 - \rho_r)
 \end{aligned} \tag{4.4-26}$$

则由式(4.4-24)、式(4.4-25)、式(4.4-26),可求得一般二进制信号的差错概率为

$$\begin{aligned}
 P_b &= \frac{1}{2} P[\text{error} | s_1(t)] + \frac{1}{2} P[\text{error} | s_2(t)] \\
 &= P[\text{error} | s_1(t)] \\
 &= Q \left[\sqrt{\frac{d_{12}^2}{2N_0}} \right]
 \end{aligned} \tag{4.4-27}$$

由上式可以看出,当 $\rho_r = -1$ (即为双极性信号时),信号距离最大,此时系统差错概率最小。

4.4.2 最大似然检测差错概率的一致边界

式(4.2-14)给出了二进制信号符号差错概率的求解公式。然而,一些调制方式的星座判决域不规则,导致无法对式中的积分进行闭式计算。这种情况下,通常求解差错概率的上边界。在最大似然检测中,一致边界是最简单、应用最广泛的边界。下面给出一致边界的推导过程。

在 ML 检测中一般判决域 $D_{m'}$ 可表示为

$$D_{m'} = \{\mathbf{r} \in \mathbb{R}^N : p(\mathbf{r} | \mathbf{s}_{m'}) > p(\mathbf{r} | \mathbf{s}_k), 1 \leq k \leq M \text{ 且 } k \neq m'\} \tag{4.4-28}$$

假设存在一个等概二进制系统,两个信号分别为 \mathbf{s}_m 和 $\mathbf{s}_{m'}$,定义其判决域 $D_{mm'}$ 为

$$D_{mm'} = \{p(\mathbf{r} | \mathbf{s}_{m'}) > p(\mathbf{r} | \mathbf{s}_m)\} \tag{4.4-29}$$

根据 $D_{m'}$ 和 $D_{mm'}$ 的定义,有

$$D_{m'} \subseteq D_{mm'} \tag{4.4-30}$$



因此

$$\int_{D_{m'}} p(\mathbf{r} | \mathbf{s}_m) d\mathbf{r} \leq \int_{D_{mm'}} p(\mathbf{r} | \mathbf{s}_m) d\mathbf{r} \quad (4.4-31)$$

上式右边是等概二进制系统发送 \mathbf{s}_m 时的差错概率。

定义成对差错概率(Pairwise Error Probability)为

$$P_{m \rightarrow m'} = \int_{D_{mm'}} p(\mathbf{r} | \mathbf{s}_m) d\mathbf{r} \quad (4.4-32)$$

由式(4.2-15),有

$$\begin{aligned} P_{e|m} &= \sum_{\substack{1 \leq m' \leq M \\ m' \neq m}} \int_{D_{m'}} p(\mathbf{r} | \mathbf{s}_m) d\mathbf{r} \\ &\leq \sum_{\substack{1 \leq m' \leq M \\ m' \neq m}} \int_{D_{mm'}} p(\mathbf{r} | \mathbf{s}_m) d\mathbf{r} = \sum_{\substack{1 \leq m' \leq M \\ m' \neq m}} P_{m \rightarrow m'} \end{aligned} \quad (4.4-33)$$

由式(4.2-16),有

$$P_e \leq \frac{1}{M} \sum_{m=1}^M \sum_{\substack{1 \leq m' \leq M \\ m' \neq m}} \int_{D_{mm'}} p(\mathbf{r} | \mathbf{s}_m) d\mathbf{r} = \frac{1}{M} \sum_{m=1}^M \sum_{\substack{1 \leq m' \leq M \\ m' \neq m}} P_{m \rightarrow m'} \quad (4.4-34)$$

式(4.4-34)为一般通信信道的一致边界。

对于 AWGN 信道,由式(4.4-8)可知成对差错概率为

$$P_{m \rightarrow m'} = P_b = Q\left(\sqrt{\frac{d_{mm'}^2}{2N_0}}\right) \quad (4.4-35)$$

利用该式,则式(4.4-34)可写为

$$P_e \leq \frac{1}{M} \sum_{m=1}^M \sum_{\substack{1 \leq m' \leq M \\ m' \neq m}} Q\left(\sqrt{\frac{d_{mm'}^2}{2N_0}}\right) \quad (4.4-36)$$

利用 Q 函数的上边界

$$Q(x) \leq \frac{1}{2} e^{-\frac{x^2}{2}} \quad (4.4-37)$$

可得

$$P_e \leq \frac{1}{2M} \sum_{m=1}^M \sum_{\substack{1 \leq m' \leq M \\ m' \neq m}} e^{-\frac{d_{mm'}^2}{4N_0}} \quad (4.4-38)$$

式(4.4-38)为 AWGN 信道一致边界的一般形式。

如果知道星座图的距离结构,则可进一步简化上述边界。定义星座图的距离枚举函数 $T(X)$ 为

$$T(X) = \sum_{\substack{d_{mm'} = \|\mathbf{s}_m - \mathbf{s}_{m'}\| \\ 1 \leq m, m' \leq M \\ m \neq m'}} X^{d_{mm'}^2} = \sum_{\text{不同的 } d} a_d X^{d^2} \quad (4.4-39)$$

式中, a_d 表示有序对 (m, m') 的数目, $m \neq m'$ 且 $\|\mathbf{s}_m - \mathbf{s}_{m'}\| = d$ 。利用该函数,式(4.4-38)可



写为

$$P_e \leq \frac{1}{2M} T(X) \Big|_{X=e^{-\frac{1}{4N_0}}} \quad (4.4-40)$$

定义星座图的最小距离为

$$d_{\min} = \min_{\substack{1 \leq m, m' \leq M \\ m \neq m'}} \|s_m - s_{m'}\| \quad (4.4-41)$$

因为 Q 函数是减函数,有

$$Q\left(\sqrt{\frac{d_{mm'}^2}{2N_0}}\right) \leq Q\left(\sqrt{\frac{d_{\min}^2}{2N_0}}\right) \quad (4.4-42)$$

代入式(4.4-36)可得

$$P_e \leq (M-1)Q\left(\sqrt{\frac{d_{\min}^2}{2N_0}}\right) \quad (4.4-43)$$

式(4.4-43)是以 Q 函数和 d_{\min} 表示一致边界的较稀疏的形式,利用 Q 函数的指数边界可得到一致边界的简单形式:

$$P_e \leq \frac{M-1}{2} e^{-\frac{d_{\min}^2}{4N_0}} \quad (4.4-44)$$

显然,一致边界表明星座图的最小距离对通信系统的性能有重要的影响。一个好的星座图应当这样设计:在功率和带宽约束范围内,它能提供最大可能的最小距离,即星座图上的信号点分离最大。

4.4.3 M 进制幅度调制

多进制 PAM 信号的星座图如图 4.10 所示。在该星座图中,任意两点之间的最小距离为 d_{\min} ,由下式确定

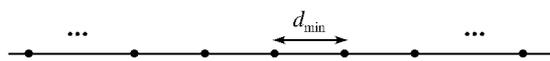


图 4.10 多进制 PAM 星座图

$$d_{\min} = \sqrt{\frac{12 \log_2 M}{M^2 - 1}} \mathcal{E}_{\text{avg}} \quad (4.4-45)$$

星座点位于 $\left\{ \pm \frac{1}{2} d_{\min}, \pm \frac{3}{2} d_{\min}, \dots, \pm \frac{M-1}{2} d_{\min} \right\}$ 。

PAM 信号的星座图中有两种类型的点: $M-2$ 个内点和 2 个外点。如果一个内点对应的信号 s_m 被发送,当接收矢量与发送信号距离大于 $d_{\min}/2$ (即 $|\mathbf{r} - s_m| > d_{\min}/2$,也可写作 $|\mathbf{n}| > d_{\min}/2$) 时,检测发生错误。将内点的差错概率记为 P_{ei} ,则有

$$P_{ei} = P\left[|\mathbf{n}| > \frac{d_{\min}}{2}\right] = 2Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right) \quad (4.4-46)$$

对于外点,由于噪声只可能在一个方向上引起检测错误,故外点的差错概率是内点的一



半。将外点的差错概率记为 P_{eo} , 则有

$$P_{eo} = \frac{1}{2} P_{ei} = Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right) \quad (4.4-47)$$

在发送消息等概时, 符号差错概率为

$$\begin{aligned} P_e &= \frac{1}{M} \sum_{m=1}^M P[\text{error} | s_m(t)] \\ &= \frac{1}{M} \left[2(M-2)Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right) + 2Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right) \right] \\ &= \frac{2(M-1)}{M} Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right) \end{aligned} \quad (4.4-48)$$

将式(4.4-45)的 d_{\min} 代入, 得到

$$\begin{aligned} P_e &= 2\left(1 - \frac{1}{M}\right) Q\left(\sqrt{\frac{6\log_2 M}{M^2 - 1}} \frac{\mathcal{E}_{\text{avg}}}{N_0}\right) \\ &\stackrel{M \text{较大时}}{\approx} 2Q\left(\sqrt{\frac{6\log_2 M}{M^2 - 1}} \frac{\mathcal{E}_{\text{avg}}}{N_0}\right) \end{aligned} \quad (4.4-49)$$

注意到式(4.4-49)中, 比特信噪比 $\mathcal{E}_{\text{avg}}/N_0$ 前有一个系数 $6\log_2 M/(M^2-1)$, 且该系数随 M 增加而趋于 0。这意味着, 当 M 增加时, Q 函数内部的值将减小, 则 Q 函数将增大, 即符号差错概率会增大。要想维持符号差错概率不变, 就必须提高比特信噪比。当 M 较大时, M 每加一倍(等效为每个符号增加 1 比特), 要保持差错概率不变, 大约要将比特信噪比提高到 4 倍, 即增加 6 dB。换言之, 为保持符号差错概率不变, 每个符号增加 1 bit 需要功率增加 6 dB。图 4.11 所示为基带 PAM (ASK) 信号在不同 M 值的符号差错概率曲线。可以看出, M 的增大将使接收机检测性能恶化, 且当 M 较大时, M 与 $2M$ 对应的曲线之间相差约 6 dB。虽然使用多进制信号会使接收机检测性能恶化, 但每个符号承载的比特数变多, 带宽效率得到提高。

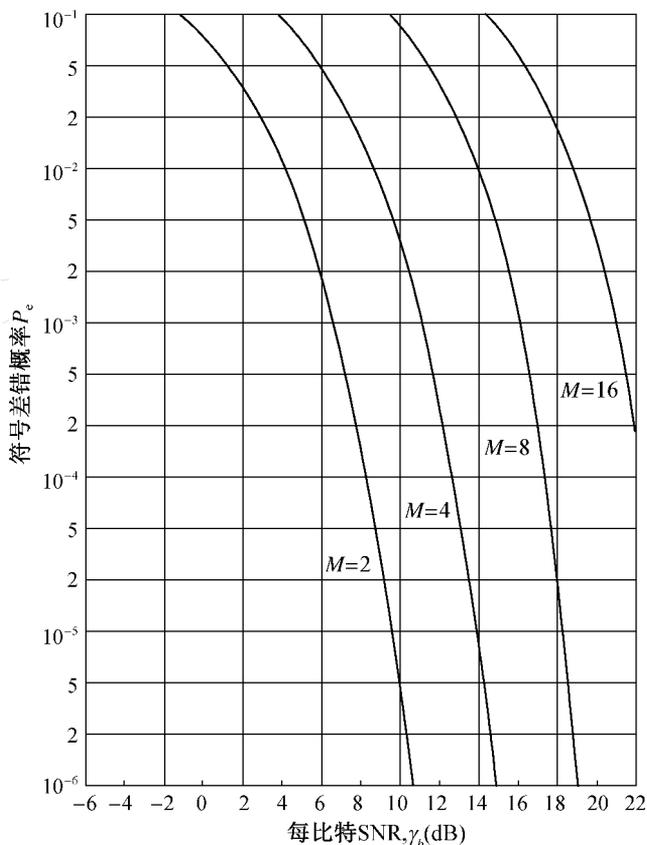


图 4.11 PAM 信号的符号差错概率曲线



4.4.4 M 进制相位调制

M 进制的 PSK 信号星座图如图 4.12 所示。假设发送消息是等概的,判决域基于最小距离检测规则。由于星座图的对称性,该系统的符号差错概率等于发送 $s_1 = (\sqrt{\mathcal{E}}, 0)$ 时的差错概率。对应的接收矢量 \mathbf{r} 为

$$\mathbf{r} = (r_1, r_2) = (\sqrt{\mathcal{E}} + n_1, n_2) \quad (4.4-50)$$

r_1 和 r_2 是方差 $\sigma^2 = N_0/2$ 、均值分别为 $\sqrt{\mathcal{E}}$ 和 0 的独立高斯随机变量。因此

$$p(r_1, r_2) = \frac{1}{\pi N_0} e^{-\frac{(r_1 - \sqrt{\mathcal{E}})^2 + r_2^2}{N_0}} \quad (4.4-51)$$

为了方便后续求解,采用极坐标来描述判决域 D_1 ,引入 (r_1, r_2) 后极坐标变换为

$$v = \sqrt{r_1^2 + r_2^2}, \Theta = \arctan(r_2/r_1) \quad (4.4-52)$$

据此,可导出 v 和 Θ 的联合 PDF 为

$$p_{v,\Theta}(v, \theta) = \frac{v}{\pi N_0} e^{-\frac{v^2 + \mathcal{E} - 2\sqrt{\mathcal{E}}v\cos\theta}{N_0}} \quad (4.4-53)$$

由于最终判决结果只取决于角度 Θ ,和幅度 v 无关,因此对联合概率密度关于 v 进行积分,导出 Θ 的边际 PDF 为

$$\begin{aligned} p_{\Theta}(\theta) &= \int_0^{\infty} p_{v,\Theta}(v, \theta) dv \\ &= \frac{1}{2\pi} e^{-\gamma_s \sin^2 \theta} \int_0^{\infty} v e^{-\frac{(v - \sqrt{2\gamma_s} \cos \theta)^2}{2}} dv \end{aligned} \quad (4.4-54)$$

式中,符号 SNR 定义为

$$\gamma_s = \mathcal{E}/N_0 \quad (4.4-55)$$

判决域 D_1 可描述为 $D_1 = \{\theta: -\pi/M < \theta \leq \pi/M\}$,所以消息差错概率为

$$P_e = 1 - \int_{-\pi/M}^{\pi/M} p_{\Theta}(\theta) d\theta \quad (4.4-56)$$

一般情况下, $p_{\Theta}(\theta)$ 的积分不能化简成简单的形式,除 $M=2$ 和 $M=4$ 外,必须采用数值计算。

当 $M=2$ 时(即采用二进制相位调制时), $s_1(t)$ 和 $s_2(t)$ 是双极性信号,因此比特差错概率为

$$P_b = Q\left(\sqrt{\frac{2\mathcal{E}_b}{N_0}}\right) \quad (4.4-57)$$

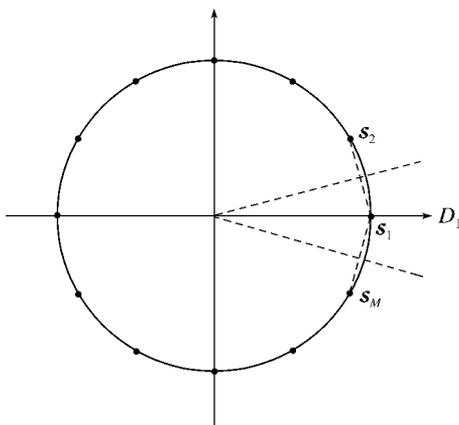


图 4.12 PSK 星座图



当 $M=4$ 时,可以看作两个相位正交的二进制相位调制。因为在两个正交载波之间没有串扰,所以比特差错概率与式(4.4-57)相同。符号判决正确当且仅当(同向/正交)两支路的 2 比特信息均判决正确,其概率可写为

$$P_c = (1 - P_b)^2 = \left[1 - Q\left(\sqrt{\frac{2\mathcal{E}_b}{N_0}}\right) \right]^2 \quad (4.4-58)$$

因此, $M=4$ 的符号差错概率为

$$P_c = 1 - P_b = 2Q\left(\sqrt{\frac{2\mathcal{E}_b}{N_0}}\right) \left[1 - \frac{1}{2}Q\left(\sqrt{\frac{2\mathcal{E}_b}{N_0}}\right) \right] \quad (4.4-59)$$

在 M 值和 SNR 值较大的情况下,差错概率主要是利用 $p_\Theta(\theta)$ 的近似表达式求得的。对于 $\mathcal{E}/N_0 \gg 1$ 且 $|\theta| \leq \pi/2$, $p_\Theta(\theta)$ 可近似为

$$p_\Theta(\theta) \approx \sqrt{\frac{\gamma_s}{\pi}} \cos\theta e^{-\gamma_s \sin^2\theta} \quad (4.4-60)$$

根据式(4.4-56),将变量 θ 变换为 $u = \sqrt{\gamma_s} \sin\theta$, 可求得

$$\begin{aligned} P_c &\approx 1 - \int_{-\pi/M}^{\pi/M} \sqrt{\frac{\gamma_s}{\pi}} \cos\theta e^{-\gamma_s \sin^2\theta} d\theta \\ &\approx \frac{2}{\sqrt{\pi}} \int_{\sqrt{2\gamma_s} \sin(\pi/M)}^{\infty} e^{-u^2} du \\ &= 2Q\left[\sqrt{2\gamma_s} \sin\left(\frac{\pi}{M}\right)\right] \\ &= 2Q\left[\sqrt{(2\log_2 M) \sin^2\left(\frac{\pi}{M}\right) \frac{\mathcal{E}_b}{N_0}}\right] \end{aligned} \quad (4.4-61)$$

注意,对于所有的 M , SNR 足够大时,该式对差错概率的近似程度都是比较好的。

对于 M 值较大的情况,利用近似式 $\sin\frac{\pi}{M} \approx \frac{\pi}{M}$ 可求得 P_c 的另外一种近似式,即

$$P_c \approx 2Q\left(\sqrt{\frac{2\pi^2 \log_2 M \mathcal{E}_b}{M^2 N_0}}\right) \quad M \text{ 较大时} \quad (4.4-62)$$

上式表明, M 较大时,为保持误码率不变, M 每增加一倍,比特信噪比就要增加到原来的 4 倍,即增加 6 dB。

PSK 信号的比特错误概率推导相当复杂,因为这种推导与 k 比特符号映射到相应信号相位有关。当映射采用格雷码时,相应于临近相位的两个 k 比特符号仅相差 1 比特。由于噪声引起的大多数可能的判决错误的情况是错误地选择了与正确相位相邻的相位,所以 k 比特中仅包含一个比特差错。因此, M 进制 PSK 的比特错误概率可近似为

$$P_b \approx \frac{1}{k} P_c \quad (4.4-63)$$

M 进制 PSK 信号的符号差错概率曲线如图 4.13 所示。可以看到,随着进制数 M 的增大,检测性能变差。当 M 较大时, M 与 $2M$ 对应的曲线之间相差约 6 dB。

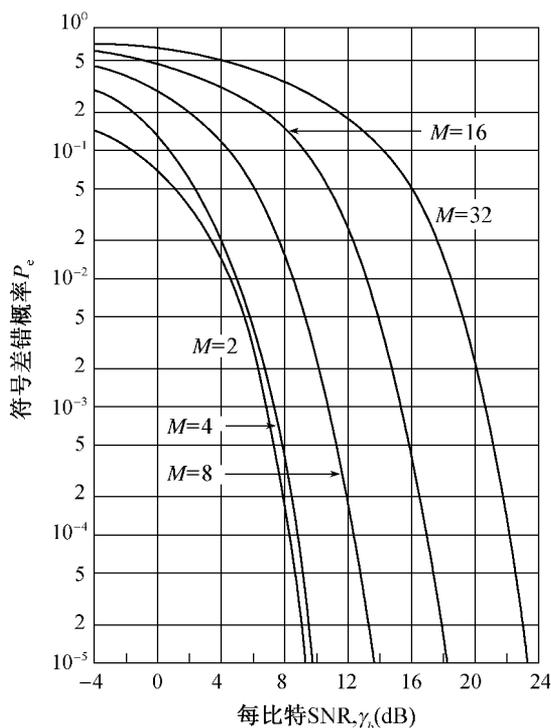


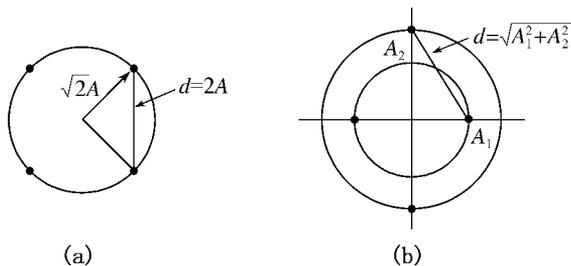
图 4.13 PSK 信号的符号差错概率曲线

4.4.5 M 进制相位幅度联合调制

QAM 信号采用两个正交载波进行调制。为了求 QAM 的差错概率,首先需要确定信号的星座图形式。图 4.14 给出了 $M=4$ 时的两种 QAM 信号星座图:图(a)是一个四相位调制信号;图(b)是具有两个幅度、四个相位的信号。假设信号等概,且信号点之间的最小距离相等,为 $d_{\min}=2A$ 。可以根据下式计算信号的平均功率

$$\mathcal{E}_{\text{avg}} = \frac{1}{M} \sum_{m=1}^M (A_{mc}^2 + A_{ms}^2) = \frac{A}{M} \sum_{m=1}^M (a_{mc}^2 + a_{ms}^2) \quad (4.4-64)$$

式中, (A_{mc}, A_{ms}) 是星座点的坐标, (a_{mc}, a_{ms}) 是由 A 归一化后的星座点坐标。

图 4.14 $M=4$ 时的两种 QAM 信号星座图

对于图(a)所示信号,平均功率为



$$\mathcal{E}_{\text{avg}} = \frac{1}{4} [4(A^2 + A^2)] = 2A^2 \quad (4.4-65)$$

对于图(b)所示信号,将星座点置于半径为 A 和 $\sqrt{3}A$ 的圆周上,满足 $d_{\min} = 2A$,则平均功率为

$$\mathcal{E}_{\text{avg}} = \frac{1}{4} [2(\sqrt{3}A)^2 + 2A^2] = 2A^2 \quad (4.4-66)$$

由上式可知两种星座图的信号具有相等的平均功率。换言之,这两种信号为达到相同的差错概率所需的平均功率相等,即两者的差错性能相同。

图 4.15 给出了 $M=8$ 时的四种 QAM 信号星座图。图中所有信号点之间的最小距离均为 $2A$,且坐标点已经由 A 归一化。可以由式(4.4-64)计算出各星座图的平均功率:对图 4.15(a)有 $\mathcal{E}_{\text{avg}} = 6A^2$;对图 4.15(b)有 $\mathcal{E}_{\text{avg}} = 6.83A^2$;对图 4.15(c)有 $\mathcal{E}_{\text{avg}} = 6A^2$;对图(d)有 $\mathcal{E}_{\text{avg}} = 4.73A^2$ 。图 4.15(d)被称为最佳 8-QAM 信号星座图,因为它在给定信号间最小距离的条件下所需功率最小。

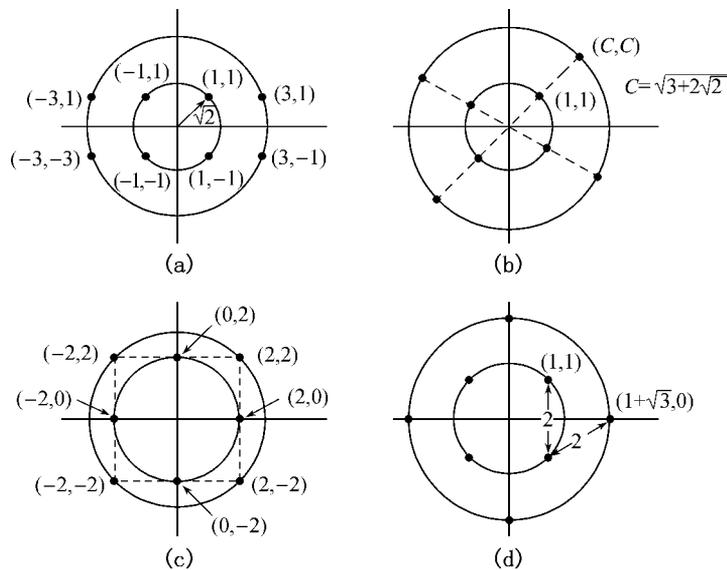


图 4.15 $M=8$ 时的四种 QAM 信号星座图

由上述讨论可见,不同的星座图在最小距离一定的情况下,平均功率和峰值功率存在不同。在各类星座图中,矩形 QAM 信号星座图可以通过两个相位正交载波上施加两个 PAM 信号来产生,具有容易产生及解调的优点。虽然当 $M \geq 16$ 时,该星座图并非最佳 M -QAM 信号星座图,但是对于给定的最小距离要求,该星座图的平均功率要求仅稍大于最佳 M -QAM 信号星座图的功率要求。因此,矩形 M -QAM 信号在实际中应用最多。

对于矩形星座图,当 k 是偶数时,该星座图可以看作是两个在同相和正交方向上的 \sqrt{M} 维的 PAM 星座图。如果在两个 PAM 信号中,存在任意一个噪声分量(n_1 或 n_2) 足够大使得相应信号的判决出错,则整个 QAM 符号的判决发生差错。因此该 QAM 系统的正确检测概率是两个 PAM 系统正确检测概率的乘积,即

$$P_{c,M-QAM} = P_{c,\sqrt{M}-QAM}^2 = (1 - P_{e,\sqrt{M}-PAM})^2 \quad (4.4-67)$$

可以计算出对应的符号差错概率为

$$P_{c,M-QAM} = 1 - (1 - P_{e,\sqrt{M}-PAM})^2 = 2P_{e,\sqrt{M}-PAM} \left(1 - \frac{1}{2}P_{e,\sqrt{M}-PAM}\right) \quad (4.4-68)$$

式(4.4-49)给出了多进制 PAM 的符号差错概率,因此有

$$P_{e,\sqrt{M}-PAM} = 2 \left(1 - \frac{1}{\sqrt{M}}\right) Q \left(\sqrt{\frac{3 \log_2 M \mathcal{E}_{\text{avg}}}{M-1 N_0}} \right) \quad (4.4-69)$$

第3章(3.2-42)式给出了该星座图的最小距离为

$$d_{\min} = \sqrt{\frac{6 \log_2 M}{M-1} \mathcal{E}_{\text{avg}}} \quad (4.4-70)$$

将式(4.4-69)、式(4.4-70)代入式(4.4-68)可得

$$\begin{aligned} P_{c,M-QAM} &= 4 \left(1 - \frac{1}{\sqrt{M}}\right) Q \left(\sqrt{\frac{3 \log_2 M \mathcal{E}_{\text{avg}}}{M-1 N_0}} \right) \times \left[1 - \left(1 - \frac{1}{\sqrt{M}}\right) Q \left(\sqrt{\frac{3 \log_2 M \mathcal{E}_{\text{avg}}}{M-1 N_0}} \right) \right] \\ &\leq 4Q \left(\sqrt{\frac{3 \log_2 M \mathcal{E}_{\text{avg}}}{M-1 N_0}} \right) \end{aligned} \quad (4.4-71)$$

图 4.16 给出了 M -QAM 的符号差错概率曲线。结合式(4.4-71)可以看到,在信噪比一定的情况下,随着 M 增大,符号差错概率变大,原因是 M 增大将导致星座点之间距离变小。比较发现,对该 QAM 信号而言, M 每增大一倍,为保持符号差错概率不变,比特信噪

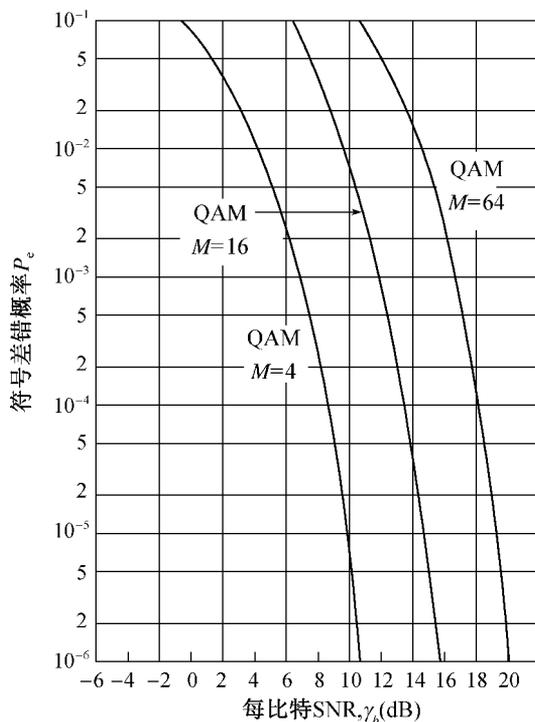


图 4.16 QAM 信号符号差错概率曲线



比仅需增加 3 dB, 而 PAM 和 PSK 信号的比特信噪比需要增加 6 dB。所以, QAM 的功率效率比 PAM 和 PSK 高。然而, PSK 的优点是其恒包络特性。

对于非矩形 QAM 信号星座图, 可利用一致边界作为差错概率的上边界:

$$P_e \leq (M-1)Q\left(\sqrt{\frac{d_{\min}^2}{2N_0}}\right) \quad (4.4-72)$$

式中, d_{\min} 是星座图的最小欧氏距离。当 M 值较大时, 该边界也许是比较松弛的。在这样的情况下, 可以用 N_{\min} 取代 $M-1$ 的办法来近似 P_e , 这里 N_{\min} 是与任何星座点距离为 d_{\min} 的最多相邻点的数目。

将 M 进制的 PSK 和 QAM 的性能进行比较, 因为这两种信号都是二维的。式(4.4-61)给出了 M 进制 PSK 符号差错概率的近似表达式, 即

$$P_{e,M\text{-PSK}} \approx 2Q\left[\sqrt{(2\log_2 M)\sin^2\left(\frac{\pi}{M}\right)\frac{\mathcal{E}_b}{N_0}}\right] \quad (4.4-73)$$

对于 M -QAM 的符号差错概率, 采用式(4.4-71), 即

$$P_{e,M\text{-QAM}} = 4\left(1 - \frac{1}{\sqrt{M}}\right)Q\left(\sqrt{\frac{3\log_2 M \mathcal{E}_{\text{bavg}}}{M-1 N_0}}\right) \times \left[1 - \left(1 - \frac{1}{\sqrt{M}}\right)Q\left(\sqrt{\frac{3\log_2 M \mathcal{E}_{\text{bavg}}}{M-1 N_0}}\right)\right] \quad (4.4-74)$$

因为符号差错概率主要受 Q 函数影响, M 较大时, Q 函数前面的系数差异可以忽略。所以为简化比较过程, 可以只比较 Q 函数内部的自变量, 这两个自变量的比值为

$$\mathcal{R}_M = \frac{\frac{3}{M-1}}{2\sin^2\left(\frac{\pi}{M}\right)} \quad (4.4-75)$$

随着 M 增大, 比值变大。 $M=4$ 时, $\mathcal{R}_M=1$, 说明 4-PSK 与 4-QAM 在符号信噪比相同时有相似的性能。当 $M>4$ 时, $\mathcal{R}_M>1$, 说明 M -QAM 的性能比 M -PSK 好。表 4.1 给出了 M 取不同值时, M -QAM 相对于 M -PSK 的 SNR 的改善量。例如, 32-QAM 相对于 32-PSK 有 7 dB 左右的改善量。 M -QAM 相对于 M -PSK 的改善可以从另一角度理解, M -PSK 的星座点均匀分布在圆周上, 仅利用相位信息, 信号幅值恒定; 而 M -QAM 信号利用幅度与相位信息使星座点间距离更大, 因而有着更好的性能。注意, 由于 M -PSK 为恒幅信号, 有不受幅度衰落影响的优点。

表 4.1 M -QAM 相对 M -PSK 的 SNR 改善量

M	$10\log\mathcal{R}_M$
8	1.65
16	4.20
32	7.02
64	9.95



4.4.6 M 进制正交信号调制

对于等能量的多进制正交信号,可以用矢量表示为

$$\begin{aligned} \mathbf{s}_1 &= (\sqrt{\mathcal{E}}, 0, \dots, 0) \\ \mathbf{s}_2 &= (0, \sqrt{\mathcal{E}}, \dots, 0) \\ &\dots \\ \mathbf{s}_M &= (0, 0, \dots, \sqrt{\mathcal{E}}) \end{aligned} \quad (4.4-76)$$

对于等概、等能量的正交信号,最佳接收器根据接收矢量 \mathbf{r} 与 M 个发送信号矢量之间的最大互相关(Largest Cross-Correlation)来选择信号,即

$$\hat{m} = \operatorname{argmax}_{1 \leq m \leq M} \mathbf{r} \cdot \mathbf{s}_m \quad (4.4-77)$$

由于星座图的对称性,其中任意两个信号点之间的距离等于 $\sqrt{2\mathcal{E}}$,故差错概率与发送信号无关。所以,可假设发送信号 \mathbf{s}_1 来计算差错概率,此时接收信号为

$$\mathbf{r} = (\sqrt{\mathcal{E}} + n_1, n_2, n_3, \dots, n_M) \quad (4.4-78)$$

式中, \mathcal{E} 表示符号能量, $n_1, n_2, n_3, \dots, n_M$ 是均值为 0、方差 $\sigma_n^2 = \frac{1}{2} N_0$ 相互统计独立的高斯随机变量。定义随机变量 $R_m (1 \leq m \leq M)$ 为

$$R_m = \mathbf{r} \cdot \mathbf{s}_m \quad (4.4-79)$$

由此定义、式(4.4-76)、式(4.4-78),可知

$$\begin{aligned} R_1 &= \mathcal{E} + \sqrt{\mathcal{E}} n_1 \\ R_m &= \sqrt{\mathcal{E}} n_m, 2 \leq m \leq M \end{aligned} \quad (4.4-80)$$

由于假设发送 \mathbf{s}_1 , 当 $R_1 > R_m (2 \leq m \leq M)$ 时,检测器判决正确,正确判决的概率为

$$\begin{aligned} P_{c|s_1} &= P[R_1 > R_2, R_1 > R_3, \dots, R_1 > R_M | s_1] \\ &= P[\sqrt{\mathcal{E}} + n_1 > n_2, \sqrt{\mathcal{E}} + n_1 > n_3, \dots, \sqrt{\mathcal{E}} + n_1 > n_M | s_1] \end{aligned} \quad (4.4-81)$$

事件 $\sqrt{\mathcal{E}} + n_1 > n_2, \sqrt{\mathcal{E}} + n_1 > n_3, \dots, \sqrt{\mathcal{E}} + n_1 > n_M$ 不是独立的,这是由于其中都有随机变量 n_1 ,但特定条件的 n_1 能使这些事件独立,所以

$$\begin{aligned} P_{c|s_1} &= \int_{-\infty}^{\infty} P[n_2 < n + \sqrt{\mathcal{E}}, n_3 < n + \sqrt{\mathcal{E}}, \dots, n_M < n + \sqrt{\mathcal{E}} | s_1, n_1 = n] p_{n_1}(n) dn \\ &= \int_{-\infty}^{\infty} (P[n_2 < n + \sqrt{\mathcal{E}} | s_1, n_1 = n])^{M-1} p_{n_1}(n) dn \end{aligned} \quad (4.4-82)$$

式中,最后一步利用了 $n_m (m=2, 3, \dots, M)$ 是 IID 随机变量的条件,于是

$$P[n_2 < n + \sqrt{\mathcal{E}} | s_1, n_1 = n] = 1 - Q\left(\frac{n + \sqrt{\mathcal{E}}}{N_0/2}\right) \quad (4.4-83)$$

因此



$$P_{e|s_1} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} \left[1 - Q\left(\frac{n + \sqrt{\mathcal{E}}}{N_0/2}\right) \right]^{M-1} e^{-\frac{n^2}{N_0}} dn \quad (4.4-84)$$

则符号差错概率为

$$P_{e|s_1} = 1 - P_{c|s_1} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \{1 - [1 - Q(x)]^{M-1}\} e^{-\frac{(x - \sqrt{2\mathcal{E}/N_0})^2}{2}} dx \quad (4.4-85)$$

式中引入了一个新变量 $x = (n + \sqrt{\mathcal{E}}) / \sqrt{N_0/2}$ 。一般地,式(4.4-85)不能化简,差错概率可通过不同的 SNR 计算得到。

在正交信号传输中,由于星座图的对称性,当发送 s_1 时,接收任何消息 $m = 2, 3, \dots, M$ 的概率都是相等的,所以,对任何 $2 \leq m \leq M$,均有

$$P_{e|s_m} = P_{e|s_1} \quad (4.4-86)$$

于是所有 M 进制正交调制信号的平均误码率为

$$P_e = P_{e|s_m} \quad (4.4-87)$$

假设 s_1 对应于长度为 k 且第一分量为 0 的数据序列,该分量的差错概率就是第一分量为 1 的序列对应的 s_m 的检测概率。因为有 $2^k - 1$ 个这样的序列,故比特差错概率为

$$P_b = 2^{k-1} \frac{P_e}{2^k - 1} \approx \frac{1}{2} P_e \quad (4.4-88)$$

式中,当 $k \gg 1$ 时,最后一步近似成立。

图 4.17 给出了正交信号比特差错概率的曲线。该图表明,在保持比特差错概率不变的情况下,增加 M 可降低对比特信噪比的要求。

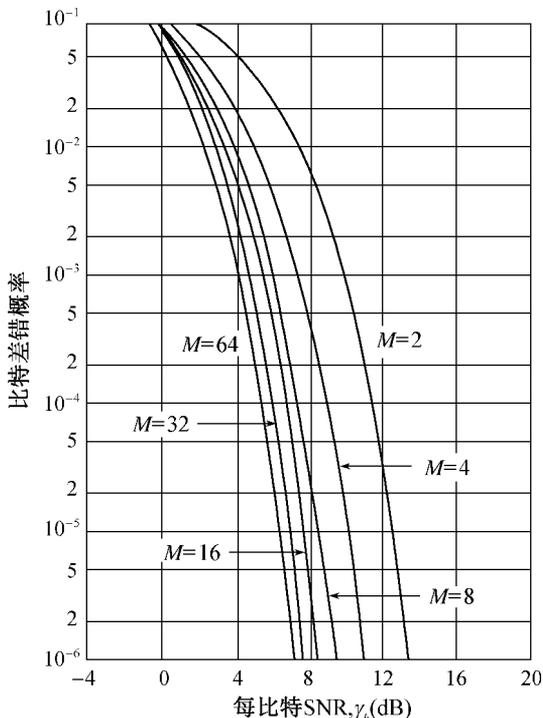


图 4.17 正交信号比特差错概率曲线



下面推导正交信号差错概率的一致边界。

式(4.4-44)给出的一致边界为

$$P_e \leq \frac{M-1}{2} e^{-\frac{d_{\min}^2}{4N_0}} \quad (4.4-89)$$

正交信号传输的 $d_{\min} = \sqrt{2\mathcal{E}}$, 所以

$$P_e \leq \frac{M-1}{2} e^{-\frac{\mathcal{E}}{2N_0}} < M e^{-\frac{\mathcal{E}}{2N_0}} \quad (4.4-90)$$

利用 $M=2^k$ 和 $\mathcal{E}_b = \mathcal{E}/k$ 可得

$$P_e < 2^k e^{-\frac{k\mathcal{E}_b}{2N_0}} = e^{-\frac{k}{2} \left(\frac{\mathcal{E}_b}{N_0} - 2\ln 2 \right)} \quad (4.4-91)$$

式(4.4-91)表明, 如果

$$\frac{\mathcal{E}_b}{N_0} > 2\ln 2 = 1.39 = 1.42 \text{ dB} \quad (4.4-92)$$

则当 $k \rightarrow \infty$ 时 $P_e \rightarrow 0$ 。换言之, 如果比特信噪比超过 1.42 dB, 可靠通信就是可能的。但对于可靠通信而言, 该条件只是充分条件而非必要条件。

式(4.4-91)表示的一致边界在低比特信噪比时不够紧密。若采用更严格的边界逼近方法, 可以推导出了更紧密的边界, 即

$$P_e \leq \begin{cases} e^{-\frac{k}{2} \left(\frac{\mathcal{E}_b}{N_0} - 2\ln 2 \right)} & \mathcal{E}_b/N_0 > 4\ln 2 \\ 2e^{-k \left(\sqrt{\mathcal{E}_b/N_0} - \sqrt{\ln 2} \right)^2} & \ln 2 \leq \mathcal{E}_b/N_0 \leq 4\ln 2 \end{cases} \quad (4.4-93)$$

可以证明, 可靠通信的充分必要条件为

$$\frac{\mathcal{E}_b}{N_0} > \ln 2 = 0.693 = -1.6 \text{ dB} \quad (4.4-94)$$

式(4.4-94)说明, 可靠通信必需的比特信噪比最小值为 -1.6 dB, 这一结论与香农在其奠基性论文 *A Mathematical Theory of Communication* 中使用随机编码进行理论推导的结果一致, 该值称为香农极限(Shannon Limit)。

4.5 数字调制带宽效率、功率效率和差错性能的权衡

4.4 节对各类数字调制信号的差错性能进行了初步分析比较。其中一种比较方式是在保持一定差错概率情况下, 比较所需的比特信噪比 $\gamma_b = \mathcal{E}_b/N_0$ 。为了更全面地衡量不同数字调制方式的性能, 本节将从带宽效率、功率效率和差错性能三方面对调制方式进行分析。

评价信号功率效率的准则是: 某调制方式达到某差错概率所需的比特信噪比。通常, 在差错概率固定的条件下比较各信号功率效率, 达到该差错概率所需的 γ_b 较低则系统功率效率较高。

定义带宽效率 r 为信号比特率与其带宽之比, 即



$$r = \frac{R}{W} \text{ (bit/s/Hz)} \quad (4.5-1)$$

r 大的系统是带宽效率高的系统,因为它能在每赫兹带宽中实现更高的比特率。

简要介绍一下带宽和维度的概念:抽样定理阐明,要重构带宽为 W 的信号,该信号的抽样速率至少为每秒 $2W$ 样值。换言之,该信号具有每秒 $2W$ 自由度(维度)。所以,带宽为 W 且持续时间为 T 的信号的维度是 $N=2WT$ 。

如果一个信号的信号集由 M 个信号组成,其每个信号的持续时间为 T_s (信号传输间隔),该信号集的带宽可近似为 W ,则该信号空间的维度 $N=2WT_s$ 。

利用关系式 $R_s=1/T_s$,得到

$$W = \frac{R_s N}{2} \quad (4.5-2)$$

由于 $R=R_s \log_2 M$,则有

$$W = \frac{RN}{2 \log_2 M} \quad (4.5-3)$$

则带宽效率可表示为

$$r = \frac{R}{W} = \frac{2 \log_2 M}{N} \quad (4.5-4)$$

式(4.5-4)以星座图的大小和星座图的维度表示信号的带宽效率。

4.4 节讨论的数字调制信号可以分为带限信号与功限信号。带限信号是指带宽被限制在一定范围内的信号,这类信号包括 PAM、PSK、QAM。对于一维调制方式(ASK 和 PAM),有 $N=1$, $r=2 \log_2 M$ 。对二维调制方式(如 M -QAM、 M -PSK),有 $N=2$, $r=\log_2 M$ 。显然可见,在 M -PAM、 M -PSK 和 M -QAM 信号传输方式中,带宽效率随 M 增大而增大。而如 4.4 节所述,随着 M 增大,这些系统为保持差错概率不变,需要更高的比特信噪比,即功率效率随 M 增大而降低。所以,这些系统的星座图大小可以在功率效率和带宽效率之间折中确定。这些系统适合于如下场合:带宽受限且要求比特率与带宽的比值 $r > 1$,有足够高的 SNR,支持增大 M 。电话信道和数字微波信道就属于这种带限信道的例子。

功限信号是指功率有限的信号。某些通信场景(如深空通信)希望在尽可能小的功率或者尽可能小的比特信噪比下实现可靠通信,而对信道带宽没有限制,因此采用功限信号。典型的功限信号就是正交调制信号(如 MFSK 信号)。对于 M 元正交信号传输, $N=M$,因此带宽效率为

$$r = \frac{2 \log_2 M}{M} \quad (4.5-5)$$

显然可见,正交调制信号的带宽效率随 M 增大而下降。而如 4.4 节所述,在保持差错概率不变情况下,正交信号可通过增大 M 来降低比特信噪比要求,即功率效率随 M 增大而增大。实际上,当 M 增大时该系统能够达到香农极限。所以,正交调制系统参数可以在功率效率和带宽效率之间折中确定。

接下来将从带宽效率、功率效率、差错性能三个方面讨论它们之间的权衡关系。图

4.18 给出了两类信号比特差错概率与比特信噪比的关系曲线,图 4.18(a)是正交信号(如 MFSK 信号),图 4.18(b)是多进制相位键控信号。沿着图中线条①所示方向移动,即在比特差错概率和比特信噪比(即功率效率)之间进行权衡。随着信噪比增大,两类信号的比特差错概率均下降。沿着图中线条②所示方向移动,即在比特差错概率和带宽效率之间进行权衡。例如图 4.18(a)中,从 c 点到 d 点, M 变大,差错性能提升,但带宽效率下降;图 4.18(b)中,从 c 点到 d 点, M 变小,差错性能提升,但带宽效率下降。沿着图中线条③所示方向移动,即差错概率不变时,在带宽效率和比特信噪比(即功率效率)之间进行权衡。例如图 4.18(a)中,从 e 点到 f 点, M 变小,带宽效率提高,功率效率降低;图 4.18(b)中,从 e 点到 f 点, M 变大,带宽效率提高,功率效率降低。

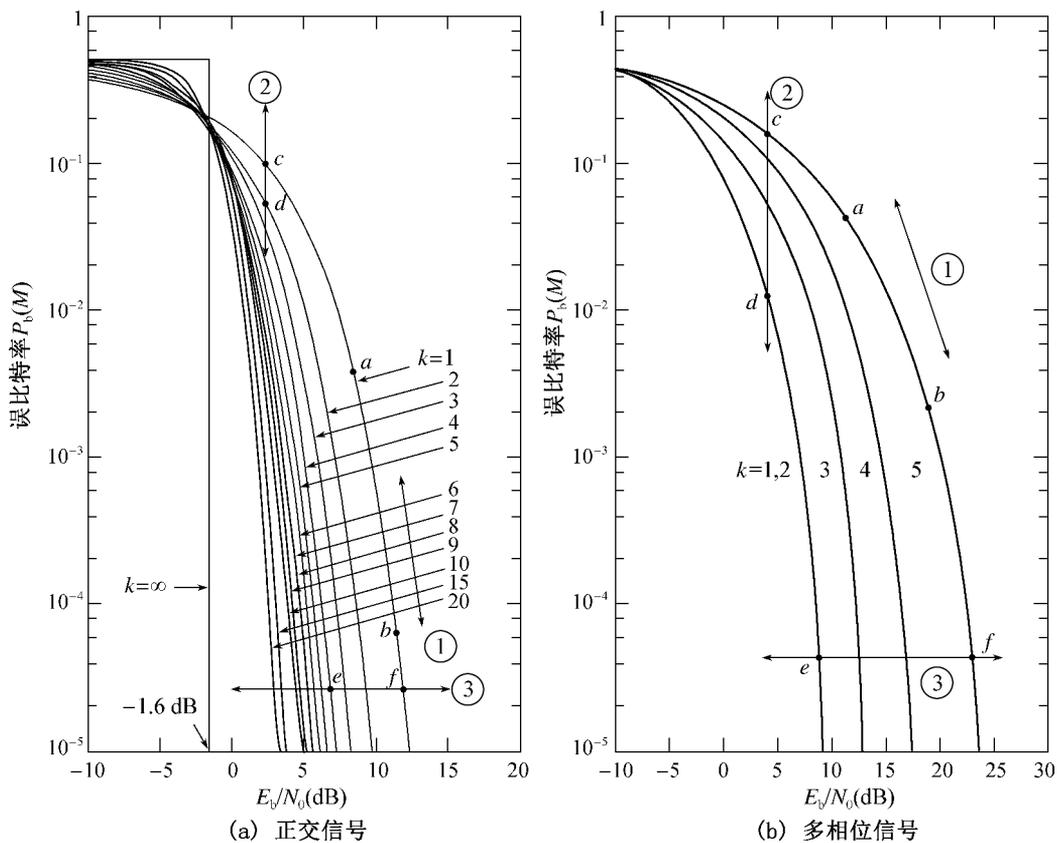


图 4.18 比特差错概率与比特信噪比的权衡关系示意图

图 4.19 给出了带宽效率平面的示意图。水平坐标轴上方为带宽效率大于 1 的区域,其中各点所代表信号均为带限信号,故该区域称为带宽受限区域;水平坐标轴下方为带宽效率小于 1 的区域,其中各点所代表信号均为功限信号,故称为功率受限区域。图左侧垂直的虚线位于 $E_b/N_0 = -1.6$ dB 处,代表香农极限。

对于带宽受限、功率受限的信道,其可传输信息速率存在上界,称为信道容量(或香农容量),其表达式为



$$C = 2W \times \frac{1}{2} \log_2 \left(1 + \frac{P}{N_0 W} \right) = W \log_2 \left(1 + \frac{P}{N_0 W} \right) \text{ bits/s} \quad (4.5-6)$$

其中, W 为信道带宽, P 为信号功率。上式说明, 信道容量和带宽、信号功率、噪声功率有关, 提高带信道宽或信号功率均可以提高信道容量。

若想实现可靠通信, 传输速率小于香农容量, 即 $R < C$, 对应图 4.19 中的白色区域, 该区域称为可达区域或者可实现区域。若 $R > C$, 则可靠通信不可实现, 对应图 4.19 中的灰色区域, 该区域称为不可达区域。两区域的分界曲线称为信道容量限, 即 $R = C$ 时的曲线, 其表达式为

$$R = W \log_2 \left(1 + \frac{P}{N_0 W} \right) \quad (4.5-7)$$

由带宽效率的定义 $r = R/W$ 则有

$$r = \log_2 \left(1 + \frac{P}{N_0 W} \right) \quad (4.5-8)$$

由

$$\mathcal{E}_b = \frac{\mathcal{E}}{\log_2 M} = \frac{P T_s}{\log_2 M} = \frac{P}{R} \quad (4.5-9)$$

有

$$r = \log_2 \left(1 + \frac{\mathcal{E}_b R}{N_0 W} \right) = \log_2 \left(1 + r \frac{\mathcal{E}_b}{N_0} \right) \quad (4.5-10)$$

进一步解得

$$\frac{\mathcal{E}_b}{N_0} = \frac{2^r - 1}{r} \quad (4.5-11)$$

上式即为以带宽效率和比特信噪比表征的信道容量限。

在图 4.19 所示的可达区域内部, 给出了差错概率 $P_b = 10^{-5}$ 、 M 不同时, QAM、PSK 和正交信号的带宽效率 r 和比特信噪比的对应关系。其中, QAM、PSK 位于带宽受限区域, 正交信号位于功率受限区域。曲线 P_{b1} 至 P_{b3} 的移动方向为差错概率提升的方向。

图 4.19 也可以解释不同调制格式差错性能、功率效率、带宽效率之间的权衡问题。沿着图中线条①所示方向移动, 从左到右, 在带宽效率不变的条件下, 比特信噪比需求变大, 但差错性能提升, 即在差错性能和比特信噪比(即功率效率)之间进行权衡。沿着图中线条②所示方向移动, 从下到上, 在比特信噪比需求不变的条件下, 差错性能下降, 但带宽效率提高, 即在差错性能和带宽效率之间进行权衡。沿着图中线条③所示方向移动, 从下到上, 在差错性能不变的条件下, 比特信噪比需求变大, 但带宽效率提高, 即在带宽效率和功率效率之间进行权衡。

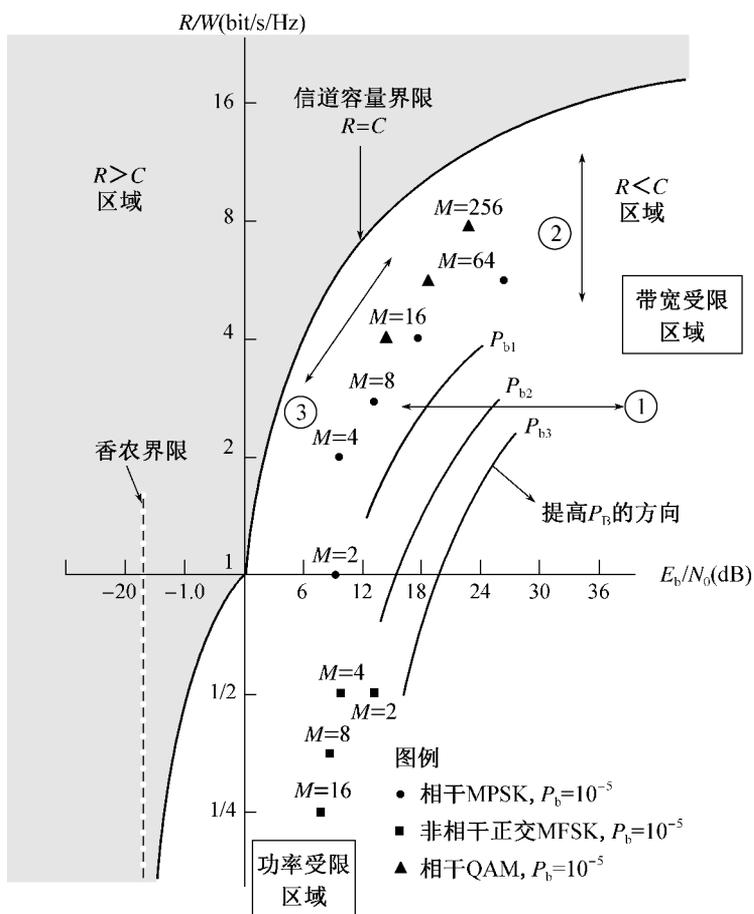


图 4.19 带宽效率平面

4.6 小结

由于信号在通过信道传输过程中普遍容易受到加性噪声的影响,本章重点考虑加性高斯白噪声,分析了 AWGN 信道下的最佳接收机设计理论与系统性能。

首先,构建了 AWGN 矢量信道模型,并在该模型下推导了 MAP 检测准则与 ML 检测准则的判决表达式。随后,由上述准则,得到最佳接收机的两种实现结构——相关接收机与匹配滤波器接收机。

然后,根据判决表达式计算差错概率,进一步研究了各类数字调制系统的接收性能。对于二进制信号,选取双极性信号与正交信号两种特例进行分析,并推导了一般二进制信号的差错概率表达式,证明差错概率取决于两种码元信号的相关系数 ρ_r 和比特信噪比 E_b/N_0 。相关系数越小,差错概率越低。对于多进制信号,推导了 PAM、PSK、QAM 与正交信号的差错概率表达式或差错概率边界。在比特信噪比 E_b/N_0 一定的条件下,随着进制数 M 的增大,PAM、PSK、QAM 的差错概率提升,而正交信号的差错概率下降。



最后,为了更全面地对各调制方式进行比较,以带宽效率、功率效率、差错性能为指标,阐述了它们之间的权衡关系。

4.7 思考题

4-1 简述相关接收和匹配滤波的异同点,在什么条件下两者能够给出相同的输出信噪比?

4-2 在 AWGN 信道中,等概率发送 N 维信号 $S_m(t), m=1, 2, \dots, M$, 接收信号为 $r(t)$ 。写出接收信号矢量 \mathbf{r} 的似然函数 $p(\mathbf{r}|\mathbf{S}_m)$ 及其 ML 意义下的最佳接收机框图,并说明它的输出是 ML 检测结果。

4-3 OOK(On-Off Keying)信号是一种特殊的二进制振幅键控(2ASK)信号,它通过如下信号来发送信息:

$$\begin{cases} s_0(t)=0 & 0 \leq t \leq T \\ s_1(t)=A \sin 2\pi f_1 t & 0 \leq t \leq T \end{cases}$$

假设发射信号是等概的,试求高斯白噪声信道下,该系统的差错概率表达式,并与双极性信号进行比较。

4-4 设 2FSK 信号为

$$\begin{cases} s_0(t)=A \sin 2\pi f_0 t & 0 \leq t \leq T \\ s_1(t)=A \sin 2\pi f_1 t & 0 \leq t \leq T \end{cases}$$

且 $f_0=2/T, f_1=2f_0, s_0(t)$ 和 $s_1(t)$ 等概率出现。

(a) 画出相关接收机结构框图。

(b) 设发送码元 010,画出接收机各点的时间波形。

(c) 设高斯白噪声信道的单边功率谱密度为 $n_0/2(\text{W/Hz})$,求该系统的差错概率。

4-5 假设有如图 4.20 所示两个低通等效信号,它们在加性高斯白噪声信道上传输二进制序列。接收信号可以表示为:

$$r_i(t)=s_i(t)+z(t), \quad 0 \leq t \leq T; i=1, 2$$

其中, $z(t)$ 为零均值高斯白噪声,其双边功率谱密度为 $n_0/2(\text{W/Hz})$ 。

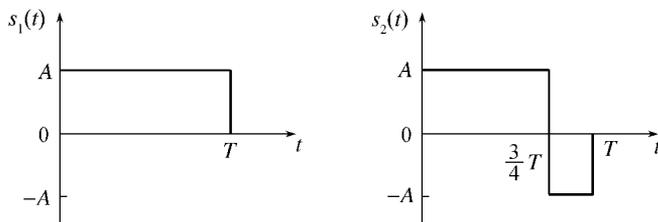


图 4.20 两个低通等效信号

(a) 求 $s_1(t), s_2(t)$ 的能量和互相关系数 ρ_{12} 。



(b) 假设接收机采用两个匹配滤波器的相干检测方法实现,画出对应的最佳接收机结构。

(c) 画出两个匹配滤波器的等效低通冲激响应和无噪声情况下的输出波形。

(d) 求该二进制系统的差错概率。

4-6 有一数字通信系统采用 QAM 在音带电话信道上以 1 600 符号/秒的速率传输信息,假设信道为加性高斯白噪声信道。试求:当差错概率要求低于 10^{-6} 、速率要求分别为 3 200 b/s、6 400 b/s、12 800 b/s 时,所需要的 \mathcal{E}_b/N_0 。

4-7 假设某带限为 W 的 AWGN 信道,其双边功率谱为 $N_0/2$,且传输信号的平均传输功率为 P_{av} ,求:

(a) 推导该信道的信道容量公式。

(b) 说明其物理含义。

(c) 推导带宽效率平面图中信道容量限曲线的表达形式。

4-8 若要为某一数字通信系统选择调制方式,应重点考虑哪些指标?它们之间存在怎样的关系?

参考文献

- [1] Proakis, J. G., & Salehi, M. Digital Communications (5th ed.) [M]. McGraw-Hill Science/Engineering/Math, 2007.
- [2] 樊昌信,曹丽娜. 通信原理(第7版)[M]. 北京:国防工业出版社,2012.
- [3] 郭勇,周冬梅. 数字通信原理[M]. 成都:电子科技大学出版社,2010.
- [4] 冯穗力等. 数字通信原理(第2版)[M]. 北京:电子工业出版社,2016.
- [5] Claude E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, vol. 27, no. 3, pp. 379 - 423, July 1948; and vol. 27, no. 4, pp. 623 - 656, October 1948.