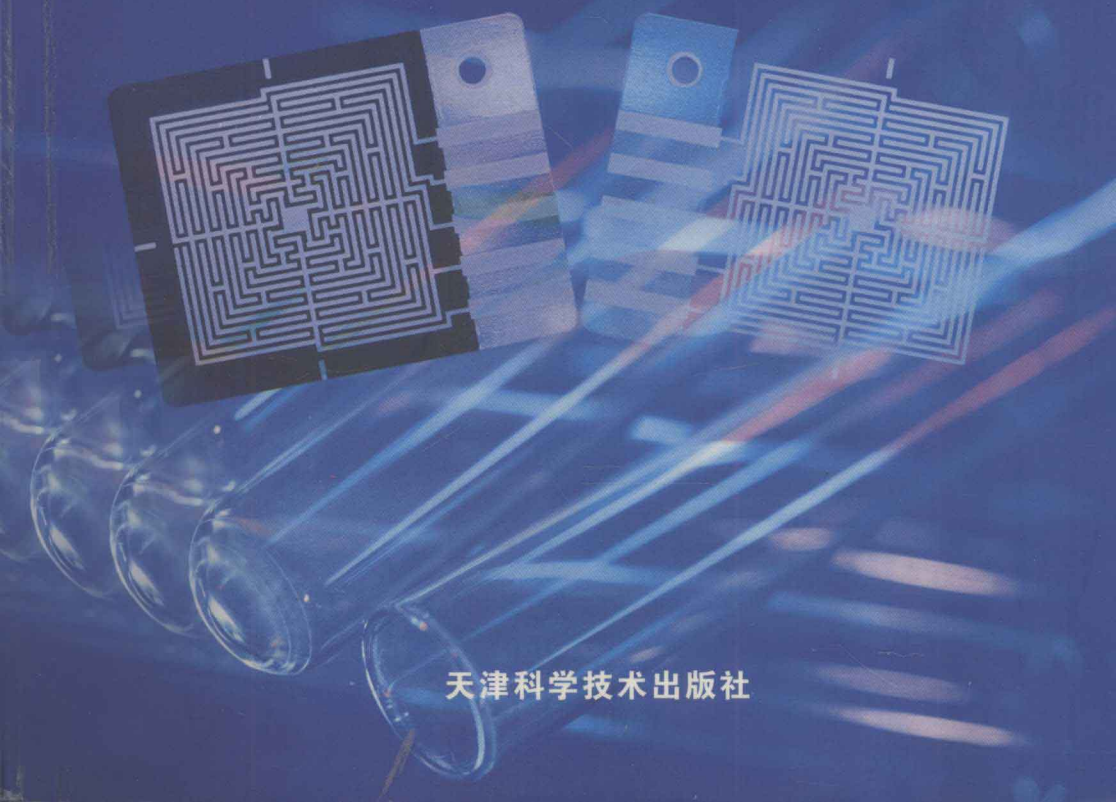


熵与信息光学

ENTROPY AND INFORMATION OPTICS

[美]杨振寰 著
陈历学 宋瑛林等 译
庄松林 校



天津科学技术出版社

熵与信息光学

ENTROPY AND INFORMATION OPTICS

[美]杨振寰 著
陈历学 宋瑛林等 译
庄松林 校

江苏工业学院图书馆
藏书章

天津科学技术出版社

图书在版编目 (CIP) 数据

熵与信息光学 / [美] 杨振寰著; 陈历学, 宋瑛林等译. 天津: 天津科学技术出版社, 2004.11

书名原文: Entropy and Information Optics

ISBN 7-5308-3844-X

I. 熵… II. ①杨… ②陈… ③宋… III. 熵-关系-信息光学-研究 IV. 0438

中国版本图书馆 CIP 数据核字 (2004) 第 111218 号

责任编辑: 王定一

责任印制: 王 莹

天津科学技术出版社出版、发行

出版人: 胡振泰

天津市西康路 35 号 邮编 300051 电话 (022) 23332393

网址: www.tjkjcs.com.cn

南开大学印刷厂印刷

开本 787×1092 1/16 印张 20.25 字数 281 000

2004 年 11 月第 1 版第 1 次印刷

定价: 46.00 元

前 言

光是空间中最重要信息载体之一。对未调制的光，我们不能从直接的观察中获得有用的信息；要利用光，就必须为此付出相应的代价！

正如我们所知道的，光不仅是维持生命所需的能量的主要来源，它也是一个很重要的信息载体。激光在 20 世纪 60 年代的出现促进了新的光通讯和处理系统的设计。而光纤通讯和光学信号处理的迅猛发展则以独特的方式揭示了光，或更准确地说光子和信息理论的关系。这种意义深远的熵和信息光学之间的关系是本书要阐述和讨论的主要内容。

在本书的写作过程中，我始终把注意力集中在熵和信息光学的几个有趣的交叉领域里而不是试图去覆盖庞大且复杂的内容。因为在实际工作中，这几个领域对高速大容量光通讯和光信息处理系统的研究和发展产生了冲击和影响。这种趋势在前些年就开始了，可以预见，它不仅会在将来继续存在而且会快速发展。因此对这些领域的知识结构和研究进展进行总结在我看来是很有意义的。这是我写作这本书的主要原因。

在宾夕法尼亚州州立大学 (The Pennsylvania State

University, USA), 本书的部分内容是作为光学信息处理课程的教学材料。这些课程主要是为高年级本科生和研究生开设的。我相信这本书对从事实际工作的工程师和科学家也有参考价值。

本书中文版的问世使我感到十分喜悦。我希望它对中文读者学习与研究光通讯和光学信息处理有所帮助。借此机会我对促进此书出版的母国光教授、周立伟教授, 译校此书的陈历学、宋瑛林、庄松林教授等表示衷心的感谢。对出版编辑策划的崔玉平先生、赵雪燕女士及《红外与激光工程》编辑部的同志表示诚挚的感谢, 没有他们的努力和所付出的代价, 这本书是不可能问世的!

楊振寰

2004 年 11 月

目 录

第 1 章 信息传输导论	(1)
1.1 信息测度	(4)
1.2 熵信息	(7)
1.3 信道	(14)
1.4 无记忆离散信道	(15)
1.5 可加噪声连续信道	(21)
1.6 总结	(30)
第 2 章 衍射与信号分析	(33)
2.1 衍射现象导论	(33)
2.2 菲涅耳-基尔霍夫原理	(36)
2.3 线性系统与傅里叶分析	(38)
2.4 有限带宽分析	(41)
2.5 信号的自由度	(46)
2.6 Gabor 信息元	(49)
2.7 信号探测	(52)
2.8 统计信号探测	(54)
2.9 信号恢复	(58)
2.10 信号模糊度	(61)
2.11 Wigner 信号表象	(64)
2.12 透镜的傅里叶变换性质	(66)
第 3 章 光学空间信道与编码原理	(71)

3.1 光学空间信道	(71)
3.2 光学空间编码与信息	(76)
3.3 不等尺寸分辨基元的空间信道	(80)
3.4 空间信道与编码的匹配	(84)
第4章 熵与信息	(90)
4.1 热力学基本定律	(90)
4.2 物理熵与信息	(94)
4.3 熵与信息的交换	(95)
4.4 典型实例	(97)
4.5 总结	(102)
第5章 伏魔师与熵耗费	(104)
5.1 永动机	(104)
5.2 Maxwell 妖魔	(106)
5.3 信息与伏魔师	(108)
5.4 再访伏魔师	(113)
5.5 Szilard 妖魔	(116)
5.6 衍射受限妖魔	(120)
5.7 最小的熵耗费	(122)
5.8 Gabor 第二类永动机	(124)
第6章 观测和信息	(129)
6.1 辐射法观测	(131)
6.2 同时观测	(135)
6.3 观测和信息	(138)
6.4 观测的精确性和可靠性	(141)
6.5 利用干涉和显微镜技术进行的观测	(149)
6.6 不确定度和观测	(156)

6.7 总结	(158)
第7章 像恢复与信息	(161)
7.1 像恢复	(161)
7.2 测不准原理与像恢复	(167)
7.3 分辨本领和信息	(171)
7.4 相干和数字像增强技术	(174)
7.5 被动信道传输中的信息丢失	(177)
7.6 模糊图像的恢复	(182)
第8章 信息传播的量子效应	(191)
8.1 问题的表述和熵的考虑	(192)
8.2 光子信道容量	(193)
8.3 信息理论方法	(198)
8.4 窄带光子信道	(202)
8.5 最佳信号功率分布, 一个特殊的实例	(208)
第9章 光学相干理论	(212)
9.1 关于相干性	(212)
9.2 空间相干性与时间相干性	(215)
9.3 相干与非相干处理	(218)
9.4 相干性的利用	(221)
9.5 总结	(227)
第10章 光学小波变换	(229)
10.1 小波变换技术	(229)
10.2 傅里叶频域中的处理	(230)
10.3 小波变换	(236)
10.4 光学实现	(241)
10.5 计算机模拟	(243)

10.6 总结	(245)
第 11 章 光学模式识别	(247)
11.1 光学相关器	(247)
11.2 光盘相关器	(251)
11.3 基于光折变的相关器	(253)
11.4 光学神经网络	(254)
11.5 复合滤波器	(260)
11.6 总结	(263)
第 12 章 光学计算	(267)
12.1 基于逻辑的计算	(267)
12.2 光学互连与混洗	(271)
12.3 矢量矩阵乘法	(273)
12.4 收缩处理器	(275)
12.5 矩阵—矩阵处理	(278)
12.6 专家系统与人工智能	(280)
12.7 总结	(288)
第 13 章 纤维光学通信	(290)
13.1 光纤通信	(290)
13.2 光纤结构	(294)
13.3 光纤传输	(297)
13.4 光纤的种类	(302)
13.5 光纤光学通讯	(305)
13.6 总结	(308)
附录 A	(310)
附录 B	(312)
附录 C	(313)

第 1 章 信息传输导论

在物理世界中，光不仅是生命赖以生存的主要能源之一，同时也为我们提供了重要的信息资源。不难想像，如果没有光，今天的现代文明将永远不会出现。令人高兴的是，我们每个人都毫无例外地拥有一双虽不算完美但却十分敏锐的眼睛。在这个世界上，正是由于同时拥有睿智的头脑与非凡的双眼，人类才可以使自己优越于世界上所有的动物。毫无疑问，如果没有这双眼睛，人类不可能进化到今天的现代文明。正因为有了光，人类才能够去搜寻他们所需要的食物，才可以去追求他们喜爱的艺术，才能够去探索他们未知的世界。因此，从抽象艺术到复杂科学应用的广阔领域，光或者说是光学，为我们提供了非常有价值的信息资源。

本书的目的是讨论光学与信息传输之间的关系。然而，必须强调指出的是，我们的目的不是要覆盖整个光学与信息领域，只是重点介绍一些比较重要且读者感兴趣的内容。

在详细讨论光学与信息之间的关系之前，首先介绍一些信息传输的基本原理。值得提及的是，熵信息并不是由光学物理学者首先提出的，而是由一群立足于数学的电气工程师提出的，他们最初的兴趣是电子通讯。不过，从熵信息发现以来，人们对它在光学应用上的兴趣从来就没有减弱过。随着现代信息光学和光通讯的最新进展，光学和熵信息之间的联系正日益密切。

尽管人们似乎明白信息一词，但对其基本的理论概念可能并不是那么清楚。现在让我们先定义信息的含义。信息可以根据几种不同的标准来定义。事实上，信息可以根据所具有的相同的数学公式的应用来定义，在以下的各节将给出这些数学公式。从纯粹的数学角度讲，信息理论基本上就是概率论。从第 1.1 节将可以看到，没有概率，就没有信息理论。

但是，从物理学的角度讲，信息理论本质上就是熵理论。在第 4 章中将能看到，如果物理熵与信息熵之间没有本质的联系，则信息理论在物理学中将无用武之地。从通讯工程师的立场看，信息理论可被看作不确定论。例如，我们收到的消息不确定性越大，其所包含的信息量越多。

我们并不是对所有感兴趣领域的信息都定义，所以这里概括为：信息熵的美与伟大之处在于它在所有科学领域中的应用，从非常抽象的领域（如经济、音乐、生物、心理学）到极其复杂深奥的科学研究。不过在引论中，我们仅从实际通讯的观点来看信息的概念。例如，从信息论的观点出发，作为传信人，一个完美的说谎者与一个绝对的诚实者是一样的，当然，必须假设我们事先知道这个人是完美的说谎者或是绝对的诚实者。必须注意的是，不要认为一个人如不诚实就是说谎。众所周知，最成功的骗子是不会让人感觉到他像骗子的。因此，我们可以把信息论看成猜测游戏，实际上它就是一个游戏理论。

通常，一个信息传输系统可以用图 1.1 所示的方块图来表示。例如，在简单的光通信中，我们用一个以某种语言的书写字符的方式来表示一个消息(信息源)，例如中文、英文、法文、德文等， 然后选择一个适于我们通信的书写字符（代码）。选定书写字符并写到纸上后，信息还没有被传送，直到用可见光（发射体）照射时信息才传送出去，显然光是该信息的载体。当来自于书写字符的反射光到达你的眼睛（接收器）时，一个适当的解码过程（翻译）发生了，也就是说，由用户（你的大脑）

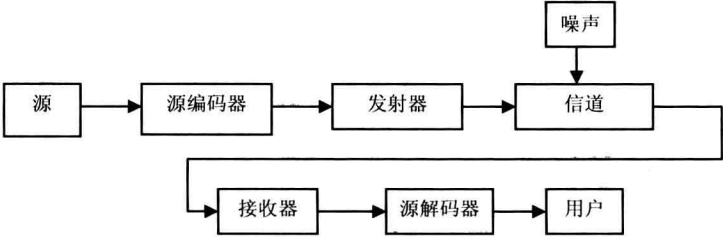


图 1.1 一个通信系统的方块图

进行了一个字符识别过程（解码）。从上面这个简单的例子可以看出，只有合适的编码过程还不够，还需要有一个合适的解码过程。如果让你看

一张英文报纸，即使这个光学信道非常理想（即没有噪声），你或许仍不能识别这种语言。这是因为一个正确的解码过程是需要提前知道编码表的（即恰当的信息存储），如事前熟悉英语。因此解码过程也称作识别过程。

信息论是一个很广泛的课题，不可能用几个小节的阐述而面面俱到。尽管如此，在论述中，我们将给出熵信息在光学上极其有益的应用。对信息论感兴趣的读者可以参考一下 Shannon 的经典论文^[1~3]和 Fano 的著作^[4]。

信息论有两种常用方式：一种是由 Wiener^[5,6]提出的，另一种则是由 Shannon^[1~3]提出的。尽管 Wiener 和 Shannon 提出的方式都是基于概率，但它们之间存在着本质的区别。

Wiener 工作的主要意义在于，一个信号（信息）被某些物理原因（如噪声，非线性畸变）破坏后可能被修复。基于此目的，Wiener 发展了相关探测、优化预测、匹配滤波等理论。然而 Shannon 的工作更进一步。Shannon 的工作证明，如果信号适当编码，则可以获得最优传输。这就是说，被传输的信号可以在传输前和传输后进行相应的处理。他指出，编码过程在一定程度上可以抵制通讯信道的干扰。因此，通过恰当的解码过程，信号可以最优还原。为此，Shannon 发展了信息测度、信道容量、编码过程等一系列理论。Shannon 理论的主要优点在于信道得到了充分的利用。

Shannon 提出的基本定理被视为在这一领域的工作中最令人惊叹的成果。这个定理可概述如下：给定一个有限存储的稳定信道，信道容量为 C ，如果这条消息的二进制信息传输速率 R 小于 C ，则存在这样的信道编码与解码过程，使得每个数据信息传输中的误差概率可以任意小。相反，如果信息传输速率 R 大于信道容量 C ，则不存在上述编码与解码过程，即这时数据信息传输中的误差概率不可以任意小。换句话说，信道中存在的随机干扰本身并不会限制信号传输的准确性。但是为获得任意高的传输精度，将限制传输速率。

最后，要再次指出 Wiener 与 Shannon 观点的不同之处。Wiener 认为，

被噪声干扰的可疑信号可以处理。Shannon 则认为，信号可以在信道传输前和传输后进行处理。当然这两个信息传输理论分支的目标基本是相同的，那就是忠实于原始信号的复原。

1.1 信息测度

前面已经讨论了信息传输的一般概念。在本节中将更详细地讨论这一主题。首要目的是定义信息测度，它在现代信息理论发展中是非常重要的。如图 1.2 所示，首先我们认为加到信道的信号为离散输入和离散输出消息集合。假设输入集合和输出集合分别表示为 $A=\{a_i\}$ 和 $B=\{b_j\}$, $i=1,2,\dots,M$, $j=1,2,\dots,N$ 。应当注意到， AB 形成了一个离散内积空间。

假设 a_i 是信道的一个输入事件， b_j 是相应的输出事件。现在来定义信息测度，即用接收到的事件 b_j 来表征 a_i 。也就是说，要定义一种信息

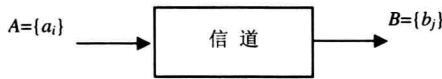


图 1.2 一个输入—输出信道

量的度量，即利用输出事件 b_j 来度量与之对应的输入事件 a_i 。可以看到，事件 a_i 通过信道传输会引起 a_i 在概率意义上的变化，即从先验概率 $P(a_i)$ 变到后效联合概率 $P(a_i/b_j)$ 。为测量这个变化，可以对这两个概率之比取对数。可以证明，这种定义方式对信息测度是很适合的。因此，输出事件 b_j 关于输入事件 a_i 所提供的信息量可以定义为：

$$I(a_i;b_j) \triangleq \log_2 \frac{P(a_i/b_j)}{P(a_i)} \quad \text{bit} \quad (1.1)$$

注意，对数的底数可以不是 2，但 2 是在信息论中最常用的底数，在本书中采用 2 作为底数。其他值的底数也经常用到，如 $\log_{10}$ 和 $\ln = \log_e$ 。这些不同底数的信息测度的单位分别是 hartleys 和 nats。hartleys 是以 R. V.

Hartley 命名的, 他首先建议用对数来度量信息^[7], 而 nats 是 natural unit (自然单位) 的缩写。在方程 (1.1) 中用的 bit 是二进制单位的缩写。

可以看到, 方程 (1.1) 对输入事件 a_i 和输出事件 b_j 具有对称性:

$$I(a_i; b_j) = I(b_j; a_i) \quad (1.2)$$

信息测度的这种对称性表示为:

$$\log_2 \frac{P(a_i, b_j)}{P(b_j)P(a_i)} = \log_2 \frac{P(b_j / a_i)}{P(b_j)}$$

根据方程 (1.2), 事件 b_j 相关于事件 a_i 提供的信息量和事件 a_i 关于事件 b_j 提供的信息量相同。方程 (1.1) 是由 Shannon 定义的信息测度, 或者称为互信息, 或者称为事件 a_i 和事件 b_j 之间信息转换的度量。

十分清楚, 如果输入与输出事件都是统计独立的, 也就是说, 如果 $P(a_i, b_j) = P(a_i)P(b_j)$, 则 $I(a_i; b_j) = 0$ 。

此外, 如果 $I(a_i; b_j) > 0$, 则 $P(a_i; b_j) > P(a_i)P(b_j)$, 这就是说, 存在一个更大的关于 a_i 和 b_j 的联合概率; 如果 $I(a_i; b_j) < 0$, 则 $P(a_i; b_j) < P(a_i)P(b_j)$, 这就是说, 存在一个更小的关于 a_i 和 b_j 的联合概率。

作为条件概率 $P(a_i / b_j) \leq 1$ 和 $P(b_j / a_i) \leq 1$ 的结果, 可以看出:

$$I(a_i; b_j) \leq I(a_i) \quad (1.3)$$

和

$$I(a_i; b_j) \leq I(b_j) \quad (1.4)$$

式中

$$I(a_i) \triangleq -\log_2 P(a_i) \quad (1.5)$$

$$I(b_j) \triangleq -\log_2 P(b_j) \quad (1.6)$$

$I(a_i)$ 和 $I(b_j)$ 分别定义为事件 a_i 和事件 b_j 各自的输入自信息与输出自信息。换句话说, $I(a_i)$ 和 $I(b_j)$ 分别表示由事件 a_i 和事件 b_j 的信道在输入与输出中提供的信息数量。由此得出结论, 当且仅当 $P(a_i, b_j) = 1$ 时, 事件

a_i 和事件 b_j 的互信息与事件 a_i 的自信息相等，即：

$$I(a_i; b_j) = I(a_i) \quad (1.7)$$

可以看出，如果方程 (1.7) 对于任意的 i 都成立，也就是对输入集合成立，则信道是无噪声的。但是，如果 $P(b_j / a_i) = 1$ ，则：

$$I(a_i; b_j) = I(b_j) \quad (1.8)$$

如果方程 (1.8) 对所有输出集合都是成立的，则信道是确定的。

需要强调的是，信息测度的定义可以扩展到更高阶内积空间。例如，我们可以对内积集合 ABC 定义一个互信息：

$$I(a_i; b_j / c_k) \triangleq \log_2 \frac{P(a_i / b_j c_k)}{P(a_i / c_k)} \quad (1.9)$$

与上述讨论相似，可以看出：

$$I(a_i; b_j / c_k) = I(b_j; a_i / c_k) \quad (1.10)$$

$$I(a_i; b_j / c_k) \leq I(a_i / c_k) \quad (1.11)$$

和

$$I(a_i; b_j / c_k) \leq I(b_j / c_k) \quad (1.12)$$

式中

$$I(a_i / c_k) \triangleq -\log_2 P(a_i / c_k) \quad (1.13)$$

和

$$I(b_j / c_k) \triangleq -\log_2 P(b_j / c_k) \quad (1.14)$$

表示条件自信息。

而且，由方程 (1.1) 可以看出：

$$I(a_i; b_j) = I(a_i) - I(a_i / b_j) \quad (1.15)$$

$$I(a_i; b_j) = I(b_j) - I(b_j / a_i) \quad (1.16)$$

从上述定义可以看出：

$$I(a_i b_j) \triangleq -\log_2 P(a_i, b_j) \quad (1.17)$$

内积集合 AB 在点 (a_i, b_j) 处的自信息定义为：

$$I(a_i; b_j) = I(a_i) + I(b_j) - I(a_i b_j) \quad (1.18)$$

与此相反，

$$I(a_i b_j) = I(a_i) + I(b_j) - I(a_i; b_j) \quad (1.19)$$

在本节结束之际，需要指出，对于互信息 $I(a_i; b_j)$ （即通过信道传输的信息量）存在一个上限或者是 $I(a_i)$ ，或者是 $I(b_j)$ ，取决于信道的性质。如果信道是无噪声的，则互信息 $I(a_i; b_j)$ 与事件 a_i 的输入自信息 $I(a_i)$ 相等。然而，如果信道是确定的，则交互信息等于事件 b_j 的输出自信息 $I(b_j)$ 。而且，如果信道的输入和输出是统计独立的，那么将不能传输任何信息。应当注意到，当联合概率 $P(a_i, b_j) < P(a_i)P(b_j)$ 时， $I(a_i; b_j)$ 是负的，也就是说，事件 b_j 关于事件 a_i 提供的信息，跟统计独立的情况相比较更加恶化。最后，应当清楚，信息测度的定义可以推广到更高阶的内积集合，即 $ABC\cdots$ 内积空间。

1.2 熵信息

在第 1.1 节，我们定义了信息测度。从中可知信息论确实是概率论的一个分支。

在这一节中，我们把信息测度看成一个随机变量，也就是说，信息测度是一个随机事件。这样，信息测度就可以用概率分布 $P(I)$ 来描述，这里 I 是自信息、条件信息或互信息。

由于信息测度通常由集合平均来表征，所提供信息的平均量可以由集合平均得到：

$$E[I] = \sum_I IP(I) \quad (1.20)$$

式中， E 表示集合平均；求和是对所有的 I 进行的。

把方程 (1.5) 中 a_i 的自信息代入方程 (1.20)，由输入集合 A 提供的自信息平均值为：

$$E[I(a_i)] = \sum_I IP(I) = \sum_{i=1}^M P(a_i) I(a_i) \quad (1.21)$$

式中, $I(a_i) = -\log_2 P(a_i)$ 。

为了使用符号方便, 略去下标 i , 这样, 方程 (1.21) 可以改写为:

$$I(A) \triangleq -\sum_A P(a) \log_2 P(a) \triangleq H(A) \quad (1.22)$$

上式中的求和是对输入集合 A 进行的。

相似地, 在信道的输出端给出的自信息的平均值可以写为:

$$I(B) \triangleq -\sum_B P(b) \log_2 P(b) \triangleq H(B) \quad (1.23)$$

事实上, 方程 (1.22) 和方程 (1.23) 是 Shannon^[1~3]信息理论的出发点。这两个方程本质上与统计热力学中的熵方程具有同样的形式。由于和熵表述具有完全相同的形式, $H(A)$ 和 $H(B)$ 经常用来表述信息熵。在以下几章中可以看到, 方程 (1.22) 和方程 (1.23) 不仅在数学形式上与熵方程相似, 而且揭示了科学与信息论之间的深刻关系^[8~10], 也揭示了光学与信息论之间的深刻关系^[11,12]。

值得注意的是, 从通讯理论的观点看, 熵 H 主要是对不确定性的度量, 而从统计热力学的角度上说, 熵 H 是无序的度量。

另外, 根据方程 (1.22) 和方程 (1.23), 有:

$$H(A) \geq 0 \quad (1.24)$$

式中, $P(a)$ 是一个恒为正的量。如果 $P(a)=1$ 或 $P(a)=0$, 方程 (1.24) 取等号。可以得到:

$$H(A) \leq \log_2 M \quad (1.25)$$

式中, M 为输入事件集合 A 中不同事件的个数, 即 $A=\{a_i\}, i=1, 2, \dots, M$ 。可以看出, 当且仅当 $P(a)=1/M$ 时, 即所有的输入事件都是等概率事件时, 方程 (1.25) 取等号。