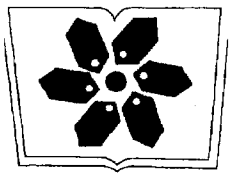
数学与现代科学技术丛书 1

量子纠错码

冯克勤 陈豪 著

科学出版社
www.sciencep.com



中国科学院科学出版基金资助出版

数学与现代科学技术丛书 1

量子纠错码

冯克勤 陈 豪 著

科 学 出 版 社
北 京

内 容 简 介

量子纠错是量子计算和量子通信得以实现的重要保证. 本书介绍量子纠错码的基本数学概念和理论、量子纠错码和经典纠错码之间的密切联系以及构作性能良好量子码的主要数学方法.

本书可作为数学、通信、计算和量子物理等专业的大学生、研究生和教师的教材或教学参考书, 也可供相关领域的科研人员阅读参考.

图书在版编目(CIP)数据

量子纠错码/冯克勤, 陈豪著. —北京: 科学出版社, 2010
(数学与现代科学技术丛书; 1)

ISBN 978-7-03-026383-4

Ⅰ.量… Ⅱ.①冯… ②陈… Ⅲ.①量子-纠错码 Ⅳ.①0157.4

中国版本图书馆 CIP 数据核字 (2010) 第 007030 号

责任编辑: 赵彦超 / 责任校对: 陈玉凤
责任印制: 钱玉芬 / 封面设计: 王 浩

科学出版社出版
北京东黄城根北街 16 号
邮政编码: 100717
http://www.sciencep.com

双青印刷厂印刷

科学出版社发行 各地新华书店经销
*
2010 年 3 月第 一 版 开本: B5(720×1000)
2010 年 3 月第一次印刷 印张: 11 3/4
印数: 1—3 000 字数: 221 000

定价: 38.00 元

(如有印装质量问题, 我社负责调换)

《数学与现代科学技术丛书》序

当代数学在向纵深发展的同时，被空前广泛地应用于几乎一切领域. 一方面，它与其他学科交汇，形成了许许多多交叉学科(例如，信息科学、计算机科学、系统科学、数学物理、数学化学、生物数学、数学语言学、数量经济学、金融数学、复杂性科学、科学计算等)；另一方面，它又被应用于高新技术的开发(例如，信息安全、信息传输、图像处理、语音识别、网络、海量数据处理、网页搜索、遥测遥感、交通管理、医疗诊断、手术方案、药物检验、商业广告等方面)，成为一些高新技术的核心. 应用数学的这种发展趋势急剧地扩展了数学的疆界，也深刻地改变了数学的面貌.

中国的经济正在迅猛发展，其中的科技含量也与日俱增. 为了提高自主创新能力，我国已经有不少数学工作者投身于这类应用数学的研究中，还有更多的数学工作者则正在密切关注这方面的进展，看好它的前景. 愈来愈多的人希望了解这类应用数学的现状，寻找入门之径.

《数学与现代科学技术丛书》是力图反映这个发展趋势的一套应用数学丛书，它将较全面地向我国读者介绍当今数学在现代科学技术各个领域中应用的状况，通过必要的准备知识，逐步把读者引向相关的研究前沿.

从事交叉学科研究和高新技术开发的应用数学家，除了要精通所需的数学知识外，还必须深入了解其所研究问题的来龙去脉. “建模”是应用数学研究实际问题的关键. 这也是一门数学艺术：从复杂的实际问题中抽象出关键的“量的关系”，使得既能反映出问题的基本特征，又能用现阶段的数学工具加以处理. 有鉴于此，这套丛书的一个特点就是：不但要介绍有关的数学理论和方法，还必须介绍问题的来源与背景、数学建模以及如何运用数学工具来解决实际问题.

本丛书适用于数学及相关专业的大学生和研究生，以及与数学有关的各专业科技工作者.

张恭庆

2009 年 11 月

前　言

随着数字计算机和数字通信的发展, 50 多年来, 离散数学 (组合学、图论、离散优化等)、数论、代数学以至代数几何在通信和计算机科学中得到重要的应用, 成为这些领域不可缺少的基本数学工具. 由于通信可靠性的实际需要, 从 20 世纪 50 年代末至今, 发展了深入系统的经典纠错码理论. 利用各种数学手段构作的性能良好的纠错码和实用的纠错译码算法, 在通信中得到实际运用.

20 世纪后期, 量子计算和量子通信成为通信界、物理界和数学界的热门话题. 在理论上, 利用量子物理的并行计算机制, 可以极大地加快计算和通信速度. 实际上, 尽管目前无法预料何时能建成量子计算机, 量子计算和量子通信的真正应用似乎也很遥远, 但是实验的进展很快, 而数学理论则呈现出超前的趋势. 发达国家均以政府行为支持这方面的研究工作.

与经典的数字通信情形一样, 纠错问题是量子通信和量子计算得以实现的必要保障之一. 1995 年以前, 人们普遍认为解决量子纠错问题比经典通信情况要困难得多. 这是由于量子物理的特殊通信机制造成的 (不可复制性、测量与环境的相互影响、纠缠态、量子态的连续性等). 1995—1996 年, 量子纠错在物理方面取得重要突破, Peter Shor 和 Steane 在物理上把复杂的纠缠态错误归结和简化为只考虑在每个量子位上出现的几种错误类型 (即 Pauli 算子 σ_x, σ_z 和 σ_y). 基于此, Shor 构作出世界上第一个量子纠错码 $[[9,1,3]]$, 用 9 个量子位的码可纠 1 位量子错误, 码空间是 2 维复空间. 随后人们把所需的 9 位码长减至 7 位和最佳的 5 位. 1998 年, Calderbank, Rain, Shor 和 Sloane 等给出量子纠错码理论的数学形式和构作量子码的第一种系统而有效的数学方法, 并由此建立了经典纠错码和量子码之间的联系. 此后, 量子码的数学研究进展很快. 在短短的五年间, 利用经典纠错码不仅构作了许多好的量子码系列 (量子 RM 码、量子 RS 码、量子代数几何码等), 而且发现了构作量子码的其他方法, 也开展了多态量子码、界估计等方面的研究. 在实验方面, 2001 年已成功地运用量子码 $[[5,1,3]]$ 实现了量子纠错功能.

综上所述, 虽然量子纠错码理论只有十余年的历史, 但已成为计算机科学、通信、物理和数学的一个交叉和前沿领域, 成为发展迅速而又富有挑战性的一个研究方向. 本书的目的是介绍量子纠错码理论的基本数学概念和主要结果. 主要涉及量子码理论的数学侧面, 不讨论其物理机制和量子通信的具体机制. 对于物理文献中的某些结果, 我们试图给出纯数学的证明. 我们希望有更多数学界人士了解这个领域并从事这方面的工作.

我们首先介绍经典纠错码的相关内容, 这是由于经典纠错码是构作量子码的重要工具, 同时, 将这两种纠错码的相似和不同之处加以对比, 也有助于对量子码的理解. 然后讲述量子码, 介绍基本概念和主要数学结果. 由于这个领域仍处于发展阶段, 对于这个新的交叉领域, 特别是与量子物理和量子通信相关的部分, 作者的知识有限, 书中有不少不完备的叙述和欠缺, 欢迎大家指正. 代数几何码在构作量子码和改进量子码渐近界的工作中起到重要作用. 但是为了介绍这一工具, 需要较多的代数几何知识. 本书略去了这方面的内容.

作者的研究工作得到国家 “973” 计划项目 “数学机械化及其在信息技术中的应用”(2004CB3180000)、国家自然科学基金数学重大项目 “信息领域中的关键数学问题” 和国家自然科学基金 (10871068) 的资助.

冯克勤
清华大学

陈　豪
华东师范大学

目　录

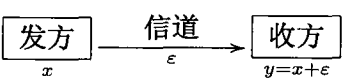
第 1 章　经典纠错码

本书的重点是量子纠错码, 但是它与经典纠错码有密切的联系. 首先, 量子纠错码尽管在物理实现机制上与经典纠错码不同, 但是许多基本的概念在数学形式上有相似之处, 如码长、信息位数、错误模式、纠错能力、最小距离、码的界估计和关于重量多项式的 MacWilliams 恒等式等. 对于两种纠错码的不同之处, 可相互参照以增强对量子纠错码的进一步理解. 其次, 许多好的量子纠错码是基于经典纠错码而构作出来的, 所以首先介绍经典纠错码.

经典纠错码已有半个世纪的历史, 内容丰富, 这里只介绍与本书所述量子纠错码有关的部分内容, 关于经典纠错码更全面和更详细的内容, 可参见有关书籍和文献, 如 Van Lint[L00] 或 MacWilliams, Sloane[MS77].

1.1　经典纠错码及其基本数学问题

数字通信的最简单模型可以表示成如下形式:



发方希望把信息 x 传输给收方, 但是在传输过程中出现错误 ε, 所以收方收到的是 $y = x + \varepsilon$. 我们希望收方在收到 y 之后有能力发现 y 有错 (检错), 进而希望能决定出错误 ε, 从而可以正确地恢复所传送的信息 $x = y - \varepsilon$(纠错).

设想需要传送 16 个信息 $\{0, 1, 2, \cdots, 15\}$, 它们可以是任何具体信息 (如电传打字的符号或者是汉字). 可以用这些数的二进制展开

$$n = a_3 \cdot 2^3 + a_2 \cdot 2^2 + a_1 \cdot 2 + a_0 \quad (a_i \in \{0, 1\}) \quad (0 \leqslant n \leqslant 15)$$

然后把 n 表示成二元域 $\mathbb{F}_2 = \{0, 1\}$ 上长为 4 的向量 (a_3, a_2, a_1, a_0). 从而这 16 个信息可表示成 $\mathbb{F}_2$ 上 4 维向量空间 $V = \mathbb{F}_2^4$ 的全部向量:

$$0 = (0000), 1 = (0001), 2 = (0010), 3 = (0011), \cdots,$$
$$13 = (1101), 14 = (1110), 15 = (1111)$$

如果发方想把数字 2 传给对方, 即传出向量 $x=(0010)$. 假如信道发生错误, 比如最左一位出错, 也就是错误向量为 $\varepsilon=(1000)$, 那么收方得到向量 $y=x+\varepsilon=(0010)+(1000)=(1010)$. 收方得到 y 之后无法判别是否有错, 因为 $y=(1010)$ 也是有意义的, 它可能是发方传来的信息 $10=(1010)$(即信道无错), 也可能是发出信息 $5=(0101)$ 而信道有错误 (1111) 等. 所以, 这个通信系统没有任何检错和纠错能力.

为了使通信系统有纠错能力, 需要把表示信息的向量长度加大. 一个最简单的例子是**重复码**, 把每个信息 (a_3,a_2,a_1,a_0) 重复 3 次而传送 $(a_3,a_2,a_1,a_0,a_3,a_2,a_1,a_0,a_3,a_2,a_1,a_0)$, 成为长 12 的二元向量. 于是, 16 个信息编成向量空间 $\mathbb{F}_2^{12}$ 中的一个子集合:

$$C=\{(a_3,a_2,a_1,a_0,a_3,a_2,a_1,a_0,a_3,a_2,a_1,a_0)\mid a_3,a_2,a_1,a_0\in\mathbb{F}_2\}$$

C 中的向量叫做**码字**, 它们是有意义的, 而 $\mathbb{F}_2^{12}$ 中其他 $2^{12}-2^4$ 个向量均不是码字, 是没有意义的.

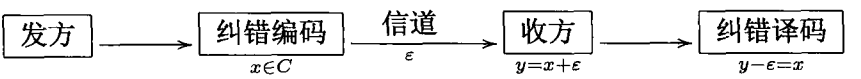
现在把代表数字 2 的码字 $x=(0010\,0010\,0010)$ 传给对方. 如果信道只产生 1 位错误, 例如仍是最左一位出错, 即 $\varepsilon=(1000\,0000\,0000)$, 则收方得到 $y=x+\varepsilon=(1010\,0010\,0010)$ 不是码字, 于是收方发现有错. 进一步, 假如信道在向量的 12 位中只发生 1 位错误, 那么容易看出错误在最左边, 因为把 y 的 12 位分成三节 (每节 4 位), 后两节完全一样, 可知错在第 1 节, 从而找到第 1 节的错位, 于是正确的信息为 $x=(0010\,0010\,0010)$. 如果信道发生 2 位的错误, 比如 $\varepsilon=(1000\,0100\,0000)$, 则收到的 $y=(1010\,0110\,0010)$ 仍不是码字, 即仍可检查 2 位错误, 但不能有效地纠错. 这表明上述重复码可检查 2 位错误, 也可以纠正 1 位错误, 但是没有更好的检查和纠错能力. 例如, 当信道有 3 位错 $\varepsilon=(1000\,1000\,1000)$ 时, 收到 $y=x+\varepsilon=(1010\,1010\,1010)$ 是码字, 所以不能发现错误. 这个重复码只能检查 2 位错和纠正 1 位错, 这是因为不同码字之间至少有 3 位不同, 并且存在两个码字恰好有 3 位不同 (如 $x=(0010\,0010\,0010)$ 和 $x'=(1010\,1010\,1010)$).

从这个例子可总结出以下几点:

(1) 16 个信息在输出时, 本来每个信息用长为 4 的向量 (每位为 0 或 1) 就可以 ($k=4=\log_2 16$ 叫信息位数). 但是, 为使通信系统有纠错能力, 我们把每个信息用 12 位来传送 ($n=12$ 叫码长), 通信速度降为原来的 $\dfrac{k}{n}=\dfrac{4}{12}=\dfrac{1}{3}$(叫做效率或信息率). 所以, 这是牺牲了速度 (或效率) 才使得通信系统具有纠错能力的.

(2) 不同码字之间的“相异”位数愈多, 检错和纠错能力就愈好.

(3) 通信系统具有纠错能力, 就是增加了“纠错编码”和“纠错译码”两个环节. 纠错编码是将原始 $K(=16)$ 个信息编成 $V=\mathbb{F}_2^n(n=12)$ 中的 K 个码字, 使不同码字的“相异位”很多, 从而有很好的纠错能力. 而纠错译码是采用有效的方式把收到的 y 恢复成正确信息 x. 于是, 一个有纠错能力的通信系统表示成



有了以上直观描述, 现在可以抽象出经典纠错码的严格数学概念.

定义 1.1.1 一个 q 元纠错码 C 是 q 元有限域 $\mathbb{F}_q$ 上 n 维向量空间 $\mathbb{F}_q^n$ 中的一个非空子集合. C 中向量 $c=(c_1,\cdots,c_n)\in C$ 叫做码字, n 叫做码长, C 中码字个数 $|C|$ 记成 K, 而 $k=\log_q K$ 叫做码 C 的信息位数, $\dfrac{k}{n}$ 叫做码 C 的信息率. 由 $1\leqslant K\leqslant q^n$ 可知 $0\leqslant k\leqslant n$, 从而 $0\leqslant\dfrac{k}{n}\leqslant 1$.

还需要刻画码 C 的纠错能力.

定义 1.1.2 对于 $\mathbb{F}_q^n$ 中每个向量 $v=(v_1,\cdots,v_n)$, 用 $w_H(v)$ 表示 v 的非零分量的个数, 叫做向量 v 的 Hamming 重量(weight), 即

$$w_H(v)=\#\{i\mid 1\leqslant i\leqslant n, 0\neq v_i\in\mathbb{F}_q\}$$

而两个向量 $v=(v_1,\cdots,v_n)$ 和 $u=(u_1,\cdots,u_n)$ 的 Hamming 距离定义为这两个向量相异位的个数, 表示成 $d_H(u,v)$, 即

$$d_H(u,v)=\#\{i\mid 1\leqslant i\leqslant n, v_i\neq u_i\}=w_H(u-v)$$

为简单起见, 在讲量子码之前, 把 $d_H(u,v)$ 和 $w_H(v)$ 简记为 $d(u,v)$ 和 $w(v)$. 不难验证, Hamming 距离有以下三个基本性质:

(H1) $d(u,v)\geqslant 0$, 并且 $d(u,v)=0$ 当且仅当 $u=v$;

(H2) $d(u,v)=d(v,u)$;

(H3) (三角不等式)$d(u,w)\leqslant d(u,v)+d(v,w)$.

定义 1.1.3 设 C 是码长为 n 的 q 元纠错码 (即 $C\subseteq\mathbb{F}_q^n$), $|C|\geqslant 2$. 定义码 C 的最小距离为 C 的所有不同码字之间的 Hamming 距离的最小值, 表示为 $d(C)$, 即

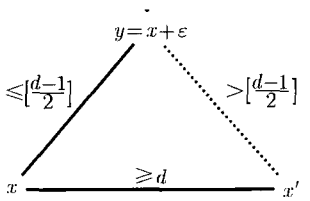
$$d(C)=\min\{d(c,c')\mid c,c'\in C, c\neq c'\}$$

下面结果表明, 最小距离这个概念确实刻画了码的纠错能力.

定理 1.1.4　*设纠错码 C 的最小距离 $d=d(C)$, 则此码可检查 $\leqslant d-1$ 位错, 也可纠正 $\leqslant \left[\dfrac{d-1}{2}\right]$ 位错 (这里, 对每个实数 α, $[\alpha]$ 表示满足 $[\alpha]\leqslant\alpha<[\alpha]+1$ 的整数 $[\alpha]$, 即 α 的整数部分).*

证明　设发出码字 $x\in C$, 信道中传输时有 $\leqslant d-1$ 位错误, 即 $w(\varepsilon)\leqslant d-1$. 则收方得到 $y=x+\varepsilon$. 如果有错, 即 $1\leqslant w(\varepsilon)\leqslant d-1$, $\varepsilon\neq 0$, 则 $y\neq x$. 另一方面, $d(y,x)=w(y-x)=w(\varepsilon)\leqslant d-1$, 而对 C 中任何码字 x' ($x'\neq x$), 均有 $d(x',x)\geqslant d$, 因此 $y\neq x'$. 这就表明 y 不是 C 中任何码字, 从而可检查 $\leqslant d-1$ 位错.

现在设 $w(\varepsilon)\leqslant \left[\dfrac{d-1}{2}\right]$, 收到 $y=x+\varepsilon$ 之后, $d(x,y)=w(\varepsilon)\leqslant \left[\dfrac{d-1}{2}\right]$. 而对于其他码字 $x'\in C$ ($x\neq x'$), 由三角形不等式知



$$\begin{aligned} d(y,x')&\geqslant d(x',x)-d(x,y)\\ &\geqslant d-\left[\frac{d-1}{2}\right]\\ &>\left[\frac{d-1}{2}\right] \end{aligned}$$

这就表明 x 是所有码字中与 y 的 Hamming 距离最小的唯一码字, 收方把 y 译成与之距离最小的码字, 就达到正确纠错的目的, 所以 C 可以纠正 $\leqslant \left[\dfrac{d-1}{2}\right]$ 位错. 证毕. □

这个定理虽然简单, 却是整个经典纠错码理论的基础.

以上给出 q 元纠错码 C 的三个基本参数: 码长 n、信息位数 k(或者用 $K=q^k=|C|$), 以及 C 的最小距离 $d=d(C)$. 把这个纠错码表示成 $(n,K,d)_q$ 或者 $[n,k,d]_q$. 若略去 d, 也可表成 $(n,K)_q$ 或 $[n,k]_q$.

经典纠错码理论最基本的研究课题有以下两个:

(1) 构作性能良好的纠错码;

(2) 对于好的纠错码, 研制出好的纠错编码和纠错译码算法, 使得这些好的纠错码在工程上得以有效的使用.

关于问题 (1). 由纠错码三个基本参数 n,k,d 的实际意义可知, 一个好的码 C 即指 $\frac{k}{n}$ 很大 (通信效率高) 并且 d 很大 (纠错能力强), 但是二者相互制约不可兼得. 比如在固定码长 n 时, 如果 k 很大 (即码字个数 $K=q^k$ 很大), 一般来说, 最小距离 d 不会很大. 确切地说, 参数 n,k,d 之间有一些约束条件, 这就是下节给出的纠错码的一些界的估计, 而达到这些界的码就是在某种意义下最好的纠错码.

关于问题 (2). 一般来说, 纠错编码比较容易实现, 即把 K 的信息编成 $\mathbb{F}_q^n$ 中 C 的 K 个码字. 但是译码算法比较困难, 是工程中非常关心的问题. 所谓译码算法, 就是在收到 $y=x+\varepsilon$ 之后是否有好的算法决定正确的码字 x. 定理 1.1.4 的证明给出了一个译码算法: 每次收到 y 之后, 与 C 中所有码字都加以比较, 其中, 与 y 的 Hamming 距离最小的那个码字 x 就是发送的正确信息. 显然, 这不是一个好的算法. 为了研制出好的译码算法, 需要设计具有各种附加性质的纠错码 (如线性、自同构、对偶性和其他几何性质或组合性质). 这不是本书的重点, 对此亦不作详细介绍.

1.2 纠错码的界

本节介绍经典纠错码三个参数 n,k(或者 K),d 以及 q 之间的一些关系, 并讲述三个界. 以下总假设 $K\geqslant 2$(即至少有 2 个码字).

定理 1.2.1(Hamming 界) *若存在纠错码* $C=(n,K,d)_q$, *则*

$$q^n\geqslant K\left(\sum_{i=0}^{[\frac{d-1}{2}]}(q-1)^i\binom{n}{i}\right)$$

其中 $\binom{n}{k}=n!/i!(n-i)!$.

证明 对于每个向量 $v\in\mathbb{F}_q^n$ 和整数 r $(0\leqslant r\leqslant n)$, 以 $B(v,r)$ 表示与 v 的 Hamming 距离 $\leqslant r$ 的所有向量构成的集合:

$$B(v,r)=\{u\in\mathbb{F}_q^n\mid d(u,v)\leqslant r\}$$

叫做以 v 为球心、r 为半径的球. 不难算出, 这个球的体积 (即球中向量个数)$|B(v,r)|$ 为

$$N_r=\sum_{i=0}^{r}(q-1)^i\binom{n}{i}$$

它与球心 v 的位置无关. 现在取 $r=\left[\dfrac{d-1}{2}\right]$, 并且考虑以 C 中所有码字为中心的 K 个球 $B(c,r)$ $(c\in C)$. 如果 c 和 c' 是 C 中不同的码字, 则 $d(c,c')\geqslant d$. 由于 $2r=2\left[\dfrac{d-1}{2}\right]\leqslant d-1<d$, 由三角不等式可知球 $B(c,r)$ 和 $B(c',r)$ 不相交. 这表明 K 个球 $B(c,r)$ $(c\in C)$ 两两不相交, 它们填到整个空间 $\mathbb{F}_q^n$ 之中, 因此

$$q^n=|\mathbb{F}_q^n|\geqslant\sum_{c\in C}|B(c,r)|=K\cdot N_r\qquad\left(r=\left[\frac{d-1}{2}\right]\right)$$

这就证明了定理 1.2.1. □

定义 1.2.2　达到 Hamming 界 (即定理 1.2.1 中的不等式变为等式) 的码叫做*完全码*.

注　完全码是一类好的纠错码. Hamming 界也叫做*球填充界*. 是否存在完全码 $(n,K,2r+1)_q$ 相当于如下的一个有限几何问题: 能否有一些 (K 个) 半径为 r 的球, 不重叠地填满整个空间 $\mathbb{F}_q^n$? 20 世纪 70 年代已决定了全部完全码的参数: 除了 2.2 节的 Hamming 码系列之外, 还有两个 Golay 码 $[23,12,7]_2$ 和 $[11,6,5]_3$ 是非平凡的完全码. 而平凡的完全码是指 $C=\mathbb{F}_q^n$(参数为 $(n,q^n,1)_q$) 和参数为 $(n,K,d)_q=(2r+1,2,2r+1)_2$ 的二元码 (对每个 $r\geqslant 1$), 如码 $C=\{(0,\cdots,0),(1,\cdots,1)\}\subseteq\mathbb{F}_2^{2r+1}$.

定理 1.2.3(Singleton 界)　*若存在纠错码* $[n,k,d]_q$, *则*

$$n\geqslant k+d-1$$

证明　设 C 是纠错码 $[n,k,d]_q$, 对每个 $a\in\mathbb{F}_q$, 考虑长为 $n-1$ 的码

$$C_a=\{(c_1,\cdots,c_{n-1})\in\mathbb{F}_q^{n-1}\mid(c_1,\cdots,c_{n-1},a)\in C\}$$

易知 $C_a(a\in\mathbb{F}_q)$ 这 q 个码彼此不相交, 合起来为 C, 于是有 $\sum\limits_{a\in\mathbb{F}_q}|C_a|=|C|=K$. 所以存在某个 $a\in\mathbb{F}_q$, 使得 $|C_a|\geqslant\dfrac{K}{q}$, 而 C_a 的参数为 $\left(n-1,\geqslant\dfrac{K}{q},\geqslant d\right)$. 以上证明了若存在码 $(n,K,d)_q$ 并且 $d\geqslant 2$, 则存在码 $\left(n-1,\geqslant\dfrac{K}{q},\geqslant d\right)_q$. 归纳下去, 便得存在 q 元码 $\left(d,\geqslant\dfrac{K}{q^{n-d}},d\right)_q$. 但是, 码长和最小距离为 d 的 q 元码最多有 q 个码字. 因此 $\dfrac{K}{q^{n-d}}\leqslant q$, 即 $K\leqslant q^{n-d+1}$. 由 $K=q^k$ 可知 $k\leqslant n-d+1$. 证毕. □

定义 1.2.4 满足 $n=k+d-1$ 的 q 元码叫做 MDS 码[①].

下节将给出完全码和 MDS 码的例子.

Hamming 界和 Singleton 界都是纠错码的参数所满足的必要性条件. 下面的界则是纠错码存在性的一个充分性条件.

定理 1.2.5 设 $2 \leqslant d \leqslant n$, $K \geqslant 1$. 如果

$$K \cdot \sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i < q^n$$

则存在参数为 $(n, K+1, \geqslant d)_q$ 的纠错码.

证明 任取 $\mathbb{F}_q^n$ 中一个向量 c, 考虑以 c 为中心、$d-1$ 为半径的球 $B(c, d-1)$, 它的体积为 $V = \sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i$. 由定理假设知 $V < q^n$, 所以 $\mathbb{F}_q^n$ 中存在向量 c' 不属于球 $B(c, d-1)$, 即 $d(c, c') \geqslant d$. 所以码 $\{c, c'\}$ 的参数为 $(n, 2, \geqslant d)_q$. 一般地, 设 $1 \leqslant M \leqslant K$, 并且存在参数为 $(n, M, \geqslant d)_q$ 的码 $\{c_1, \cdots, c_M\} \subset \mathbb{F}_q^n$. 以 S 表示 M 个球 $B(c_i, d-1)(1 \leqslant i \leqslant M)$ 的并集. 则

$$|S| \leqslant \sum_{i=1}^{M} |B(c_i, d-1)| = MV \leqslant KV < q^n$$

这表明存在向量 $c_{M+1} \in \mathbb{F}_q^n \setminus S$. 由 $c_{M+1} \notin S$ 可知 $d(c_i, c_{M+1}) \geqslant d (1 \leqslant i \leqslant M)$. 这表明码 $\{c_1, \cdots, c_M, c_{M+1}\}$ 的参数为 $(n, M+1, \geqslant d)_q$. 现在取 $M=K$, 定理即证. □

注 定理 1.2.5 中的界是 Gilbert 和 Varshamov 于 1952 年独立给出的, 称作 GV 界. 定理 1.2.5 的证明是非构造性的, 即当定理 1.2.5 中的不等式成立时, 定理只是证明了纠错码的存在性, 并没有明显地给出具体构作方式. 一直到 30 年之后, 发明了代数几何码, 才于 1982 年找到构作纠错码的具体方法, 其参数达到甚至超过 GV 界.

1.3 线 性 码

定义 1.3.1 $\mathbb{F}_q^n$ 的每个 $\mathbb{F}_q$ 向量子空间 C 均叫做码长为 n 的 q 元线性码.

① 这个名称的来源是: 这种码等价于一种组合设计方案, 叫做极大距离可分 (maximal distance separable) 的设计.

对于线性码 C, 若维数为 $k=\dim_{\mathbb{F}_q} C$, 则 $K=|C|=q^k$, 从而 k 就是该线性码的信息位数. 另一方面, 由于 C 是线性的, 它包含零向量, 并且对 C 中任意两个码字 c 和 c', $c-c'$ 也是码字. 由此可知, 线性码 C 的最小距离为

$$\begin{aligned} d(C) &= \{d(c,c') \mid c,c' \in C, c \neq c'\} \\ &= \{w(c-c') \mid c,c' \in C, c \neq c'\} \\ &= \{w(c) \mid 0 \neq c \in C\} \end{aligned}$$

即 $d(C)$ 等于 C 中非零码字 Hamming 重量的最小值.

参数为 $[n,k]_q$ 的线性码 C 是 $\mathbb{F}_q^n$ 的一个 k 维 $\mathbb{F}_q$ 向量子空间, 从而可以选取 C 的一组 $\mathbb{F}_q$ 基 $v_1,\cdots,v_k$, 其中

$$v_i=(a_{i1},a_{i2},\cdots,a_{in}) \quad (a_{ij}\in\mathbb{F}_q, 1\leqslant i\leqslant k, 1\leqslant j\leqslant n)$$

于是每个码字唯一表示成

$$\begin{aligned} c &= b_1v_1+\cdots+b_kv_k \quad (b_i\in\mathbb{F}_q) \\ &= (b_1,\cdots,b_k)G \end{aligned}$$

其中 G 是 $\mathbb{F}_q$ 上秩为 k 的 $k\times n$(k 行 n 列) 矩阵

$$G=\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{pmatrix}$$

G 叫做线性码 C 的一个生成 (矩) 阵, 可以先把 $K=q^k$ 个原始信息分别表示成 $\mathbb{F}_q^k$ 中的全部向量 $(b_1,\cdots,b_k)(b_i\in\mathbb{F}_q)$, 然后作映射

$$\varphi:\mathbb{F}_q^k\to\mathbb{F}_q^n, \quad \varphi(b_1,\cdots,b_k)=(b_1,\cdots,b_k)G$$

这是 $\mathbb{F}_q$ 线性的单射, 象集合 $\varphi(\mathbb{F}_q^k)=\mathrm{Im}\varphi$ 就是线性码 C. 这就是线性码的纠错编码, φ 把信息由 k 位的 $b=(b_1,\cdots,b_k)$ 编成码长为 n 的码字 $\varphi(b)=bG=b_1v_1+\cdots+b_kv_k$.

线性码由它的一组基所决定, 从而可以由一个生成阵所决定. 另一方面, 一个线性码 $[n,k]_q$(即 $\mathbb{F}_q^n$ 的一个 k 维 $\mathbb{F}_q$ 向量子空间) 是 n 元 $(x_1,\cdots,x_n)$ 齐次线性方程组

$$\begin{cases} b_{11}x_1+b_{12}x_2+\cdots+b_{1n}x_n=0 \\ b_{21}x_1+b_{22}x_2+\cdots+b_{2n}x_n=0 \\ \cdots\cdots\cdots \\ b_{n-k,1}x_1+b_{n-k,2}x_2+\cdots+b_{n-k,n}x_n=0 \end{cases} \tag{1.3.1}$$

的解空间, 其中

$$H=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n-k,1} & b_{n-k,2} & \cdots & b_{n-k,n} \end{pmatrix}=\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-k} \end{pmatrix}$$

是 $\mathbb{F}_q$ 上秩为 $n-k$ 的 $(n-k)\times n$ 方阵 (H 的秩为 $n-k$, 相当于 (1.3.1) 中 $n-k$ 个线性方程是 $\mathbb{F}_q$ 线性无关的). H 叫做线性码 C 的一个校验 (矩) 阵. 线性码 C 也可由校验阵 H 所决定, 因为由方程组 (1.3.1) 可知, 对于 $c=(c_1,\cdots,c_n)\in\mathbb{F}_q^n$, 有 $c\in C$ 当且仅当 $Hc^{\mathrm{T}}=0$(长为 $n-k$ 的零向量). 所以用校验阵 H 可以很方便地检查向量 c 是否为码字.

事实上, 还可以用校验阵来纠错, 使线性码的译码变得容易 (从略). 下面的结果表明, 用校验阵还可以决定线性码的最小距离.

定理 1.3.2 设 C 是线性码 $[n,k]_q$, 将它的一个校验阵写成列向量的形式:

$$H=\left(w_1^{\mathrm{T}},\cdots,w_n^{\mathrm{T}}\right)$$

其中

$$w_j^{\mathrm{T}}=\begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{n-k,j} \end{pmatrix}\quad (1\leqslant j\leqslant n)$$

则 C 的最小距离为 d 当且仅当 H 的任意 $d-1$ 列均 $\mathbb{F}_q$ 线性无关, 并且 H 有 d 列是 $\mathbb{F}_q$ 线性相关的.

证明 设 $c=(c_1,c_2,\cdots,c_n)$ 是 $\mathbb{F}_q^n$ 中 Hamming 重量为 l 的向量, 即 c 中有 l 个分量 $c_{j_1},c_{j_2},\cdots,c_{j_l}$ 不为零, 而其余分量均为零. 由于 H 是码 C 的校验阵, 所以

$$c\in C\ \Leftrightarrow\ 0=Hc^{\mathrm{T}}=(w_1^{\mathrm{T}},\cdots,w_n^{\mathrm{T}})\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}=c_{j_1}w_{j_1}^{\mathrm{T}}+\cdots+c_{j_l}w_{j_l}^{\mathrm{T}}$$

这就表明 C 中每个重量为 l 的码字 c 对应着 H 中有 l 列向量是 $\mathbb{F}_q$ 线性相关的. 所以 C 的最小距离 (即非零码字的最小重量) 就等于 $w_1^{\mathrm{T}},\cdots,w_n^{\mathrm{T}}$ 中线性相关的最少个数. 证毕. □