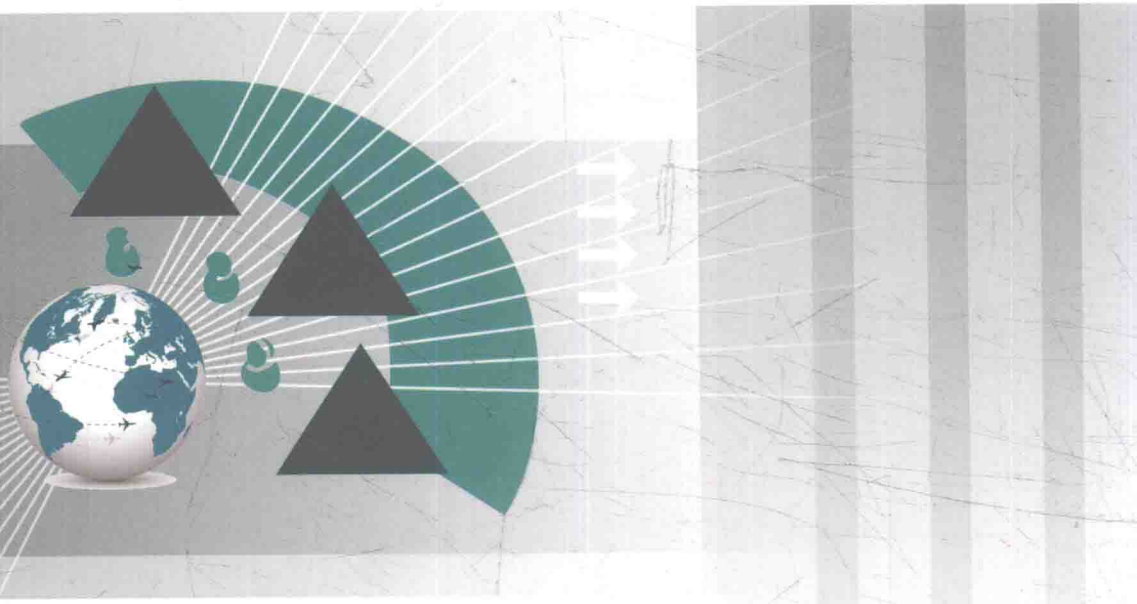
信息化网络平台研究丛书

格上基于身份的密码体制研究

陈莉 光焱 段然◎著



RESEARCH ON IDENTITY-BASED CRYPTOSYSTEM OVER LATTICE

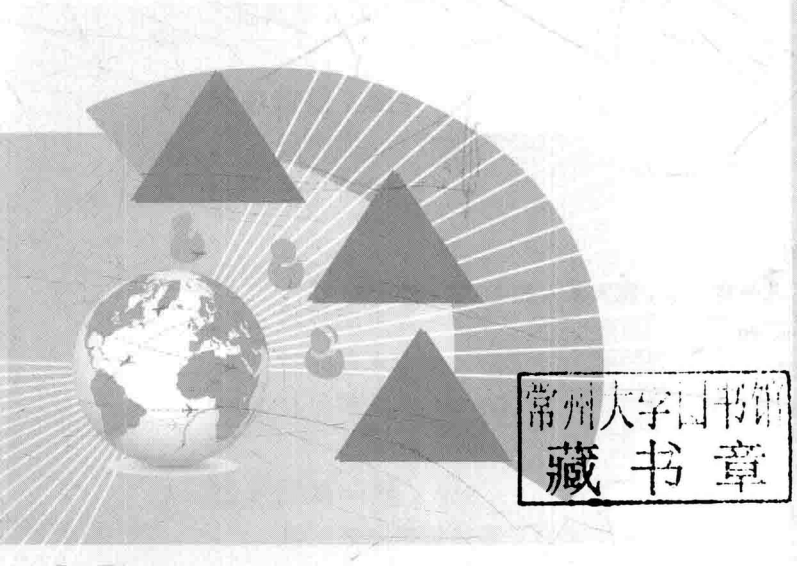
经济管理出版社
ECONOMY & MANAGEMENT PUBLISHING HOUSE

本专著由“河南省高校科技创新人才支持计划”（No.13HASTIT043）和河南省教育统计与数据分析研究中心资助出版

信息化网络平台研究丛书

格上基于身份的密码体制研究

陈莉 光焱 段然◎著



RESEARCH ON IDENTITY-BASED CRYPTOSYSTEM OVER LATTICE

经济管理出版社
ECONOMY & MANAGEMENT PUBLISHING HOUSE

图书在版编目（CIP）数据

格上基于身份的密码体制研究 / 陈莉，光焱，段然著. —北京：经济管理出版社，2017.9
ISBN 978-7-5096-5390-6

Ⅰ.①格… Ⅱ.①陈… ②光… ③段… Ⅲ.①密码学—高等学校—教材
Ⅳ.①TN918.1

中国版本图书馆 CIP 数据核字（2017）第 235738 号

组稿编辑：杨 雪
责任编辑：赵喜勤
责任印制：司东翔
责任校对：王淑卿

出版发行：经济管理出版社
（北京市海淀区北蜂窝 8 号中雅大厦 A 座 11 层 100038）
网 址：www.E-mp.com.cn
电 话：（010）51915602
印 刷：北京玺诚印务有限公司
经 销：新华书店
开 本：720mm×1000mm/16
印 张：16
字 数：238 千字
版 次：2018 年 3 月第 1 版 2018 年 3 月第 1 次印刷
书 号：ISBN 978-7-5096-5390-6
定 价：55.00 元

目录 Contents

上篇　基础篇

中篇 公钥密码体制

上篇　基础篇

本篇摘要

随着互联网技术的蓬勃发展，网络通信的安全性受到了各界的高度重视。作为网络安全的基石，公钥密码在保障网络通信安全的过程中发挥了重要作用。但随着量子计算机研究的推进，传统公钥密码算法的安全性受到了越来越大的挑战。基于格上困难问题的公钥密码体制对量子攻击算法免疫，是后量子时代的重要密码技术之一，具有重要的研究意义和应用价值。

云计算代表了IT领域迅速向集约化、规模化与专业化发展的趋势，被形象地比喻为当前信息技术领域正在发生的工业革命。但同时，云计算的发展也面临许多关键性问题，其中一些已成为云计算发展与应用中的重要制约因素，云中的数据安全和隐私保护问题正是这些问题中比较突出的一个。在云计算背景下，借助加密技术保护数据安全和隐私的传统方法面临着严重的局限性：由于服务器无法在不解密的情况下对用户的加密数据进行任何有效运算，因此在普通加密技术条件下，云计算的便利性与用户对于数据安全和隐私的需求成为一对难以调和的矛盾。全同态加密是近年来发展起来的一项新型加密技术，它允许任何人在无须解密的情况下，在加密数据上进行各种有意义的运算，因而为上述问题和矛盾提供了一条切实有效的解决途径，拥有广阔的应用前景。

本篇根据Peikert C. 撰写的*A Decade of Lattice Cryptography*一文，围绕公钥加密和格的历史背景、基础知识、格上基础困难问题、格上新型困难问题、格在体制设计上的应用、全同态加密体制、针对格上困难问题和体制的攻击以及针对公钥体制的泄露攻击等问题进行描述和讨论。

1 历史背景

1.1 公钥加密

公钥加密（Public Key Encryption，PKE）又称非对称密钥加密。与对称加密不同，公钥加密中用户的密钥由公钥、私钥组成，可以用来加密，同时也可以用来进行密钥管理、数字签名等。

1976 年，W. Diffie 等为解决传统密码体制（主要针对对称密码体制）中密钥分发困难、密文安全性高度依赖密钥等缺陷，第一次提出了公钥密码体制的概念，这是非对称密码学领域的开山之作。

1978 年，R. L. Rivest 等提出了基于大数分解问题的 RSA 体制；同年，R. C. Merkle 等提出了基于背包问题的背包体制。这两个体制是最早的两种公钥密码体制。目前常用的还有 T. El Gamal 提出的基于离散对数问题的 ElGamal 体制，以及 N. Koblitz 和 V. Miller 提出的基于椭圆曲线离散对数问题的椭圆曲线公钥密码体制（ECC 体制）。

1.2 格

格是多维欧式空间中离散点的规则排列，这些点的坐标可由空间中的一组相互独立的向量以整系数形式唯一表出。格在密码领域，最初用来分

析、攻击已有的密码体系。

1982 年，A. K. Lenstra 等提出了一种可以在多项式时间内近似求解格上最短向量问题的算法，称作 Lenstra-Lenstra-Lovász 算法（LLL 算法）。

1983 年，A. Shamir 利用该算法对 Merkle-Hellman 背包体制进行攻击，可以在多项式时间内破坏这一基于 NP 完全问题的公钥加密体制的安全性。1997 年，D. Coppersmith 利用该算法对使用小指数的 RSA 体制进行攻击。

1996 年，M. Ajtai 开创性地给出了第一个格上难题最坏情况到一般情况的归约，以及第一个基于被普遍接受的格上难题的可证安全的密码学模型。这也是第一个基于标准的最坏情况复杂度假设的加密函数。M. Ajtai 引入了最小整数解问题（Smallest Integer Solution，SIS）和基于该问题的单向函数，证明解决该问题的难度不低于近似求解一些格上难题的最坏情况。SIS 问题和该单向函数至今仍有着广泛应用。此后，在 1997 年，M. Ajtai 等给出了第一个基于格的公钥加密体制（AD 体制），同时也是第一个基于最坏情况复杂度假设的公钥加密体制。但该体制也存在只能进行单比特加密、公私钥尺寸大、效率低（公钥尺寸为 $\tilde{O}(n^4)$，私钥尺寸和运算时间为 $\tilde{O}(n^2)$）等缺点。

1997 年，基于 Ajt96 的工作以及 R. J. McEliece 的编码技术，O. Goldreich 等提出了一种基于格的公钥加密和签名体制（GGH 体制）。与 AD 体制不同的是，该体制没有最坏情况安全性保证，而只是进行了启发式的评估。在 1999 年，P. Q. Nguyen 就成功对具有可实际应用的参数的 GGH 体制进行了攻击（但对参数更高的 GGH 体制则无法进行攻击）；在 2006 年，GGH 签名体制更被证明是彻底不安全的。虽然该体制被证明是不安全的，但该体制的核心思想在格密码的后续研究中发挥了重大的作用，并由此诞生了许多具有特殊功能的密码学应用。

GGH 加密和签名体制的主要思想是将公钥定为某个格的一组“坏”基，而私钥则取该格的一组“好”基。直观地说，构成“坏”基的向量长度较长，正交性较差，而构成“好”基的向量长度相对较短。在密钥生成时，通常先生成一组“好”基，再通过乘以一个随机的幺模矩阵得到同一个格的一组“坏”基。此外，每个格都具有一个特殊的基，即埃尔米特正规形（Hermite Normal Form，HNF）。由于它可以通过任意一组格基较方便

地得到，因此可以看作性质最“坏”的一组格基，故也可使用其作为公钥。

1998 年，J. Hoffstein 等设计出了 NTRU 公钥加密体制（HPS 体制），这是第一个基于多项式环的加密体制。基于 NTRU 的密码体制的效率和密钥的空间利用率都较高，在参数选择恰当的情况下具有较强的抗攻击能力。HPS 体制与该体制所基于的问题同样没有明确的难度归约，既不存在从格上难题最坏情况到 NTRU 问题一般情况的归约，也不存在从求解 NTRU 问题到破坏 HPS 体制语义安全性的归约。

2002 年，受 NTRU 算法高效性的启发，D. Micciancio 将 Ajt96 的单向抗碰撞函数移植到多项式环$\mathcal{R}=\mathbb{Z}[X]/(X^n-1)$上，将密钥尺寸和运行时间从$\tilde{O}(n^2)$降低到$\tilde{O}(n)$。D. Micciancio 证明了该函数在假定近似求解 n 维循环格上某些问题最坏情况困难性的条件下，该函数是单向函数。与 Ajt96 的单向抗碰撞函数不同的是，该函数只能满足单向性，而无法满足抗碰撞性，不过可以通过轻微的改动而具有抗碰撞性。这些工作止步于只限于构造单向抗碰撞哈希函数，而并没有给出基于这些函数的公钥加密体制。但这些工作为后来的环上带错学习问题做了铺垫，而这一问题的用途则包括但不限于公钥加密体制。

2003 年，O. Regev 对 AD 体制进行了较大的改进，最显著的莫过于通过引入高斯测量和谐波分析来设计和分析基于格的公钥加密体制。这些技术使算法和分析更加简单，大幅降低了所基于的格上难题最坏情况的近似因子的大小。在 AD 体制中，如果体制的安全性被破坏，那么就可以求解近似因子为$\gamma\approx n^7$的近似唯一最短向量问题（γ-approximate unique-Shortest Vector Problem，$uSVP_\gamma$）；而在 Regev 体制中，近似因子只有约$\gamma=\tilde{O}(n^{1.5})$。虽然体制的公钥尺寸同样为$\tilde{O}(n^4)$，私钥尺寸和运算时间也均为$\tilde{O}(n^2)$，故在保证相近安全性的条件下，该体制所需要的格维数更低，因此具有更小的密钥尺寸和更低的运算时间。

2 基础知识

2.1 符号和定义

$\mathbb{Z}^+$、$\mathbb{Z}$、$\mathbb{Q}$、$\mathbb{R}$ 分别表示正整数集、整数集、有理数集和实数集。向量以粗斜体小写字母形式表示，如 $\boldsymbol{a}$。矩阵以粗斜体大写字母形式表示，如 $\boldsymbol{A}$。常量以正体表示，如 1。变量以斜体表示，如 a。多项式以斜体表示，如 a。矩阵 $\boldsymbol{A}$ 及向量 $\boldsymbol{a}$ 的转置用 $\boldsymbol{A}^{\mathrm{T}}$ 和 $\boldsymbol{a}^{\mathrm{T}}$ 表示。对于 2 的方幂 n，用 $\mathcal{R}$ 表示多项式环 $\mathbb{Z}[x]/(x^n+1)$，用 $\mathcal{R}_{\mathbb{Q}}$ 表示 $\mathbb{Q}[x]/(x^n+1)$，环上元素 f，g 的多项式乘法记作 fg。对于正整数 k，$[k]$ 表示集合 $\{0, \cdots, k-1\}$。$x \leftarrow \mathcal{D}$ 表示从概率分布 $\mathcal{D}$ 中随机选取变量 x，$x \xleftarrow{R} A$ 表示从集合 A 中随机均匀选取变量 x。向量 $\boldsymbol{a} \in \mathbb{Z}_q^n$ 可表示为 $a=(a_0, \cdots, a_{n-1})$；多项式 $a \in \mathcal{R}_q$ 表示为 $a=(a_0, \cdots, a_{n-1})$。$a_{i,j}$ 表示 a_i 的第 j 比特（最低位为第 0 比特，最高位为第 $l-1$ 比特）。$\boldsymbol{c}_i$ 表示矩阵 $\boldsymbol{C}$ 的第 i 行。

定义 2.1 对于函数 $f(x)$，$g(x)$，如果存在常数 c，$d>0$，使得对于任意的 $x>d$，有 $|f(x)| \leqslant c \cdot |g(x)|$，则 $f(x)=O(g(x))$。

定义 2.2 对于函数 $f(x)$，$g(x)$，如果对于任意的 $\varepsilon>0$，存在常数 $x>0$，使得对于任意的 $x>c$，有 $f(x)/g(x) \leqslant \varepsilon$，则 $f(x)=o(g(x))$。

定义 2.3 对于函数 $f(x)$，$g(x)$，如果 $f(x)=O(g(x))$ 且 $g(x)=$

$O(f(x))$，则$f(x)=\Theta(g(x))$。

定义 2.4 对于函数$f(x)$，$g(x)$，如果存在常数c，使得$f(x)=O(g(x)\cdot\log^{c}(g(x)))$，则$f(x)=\tilde{O}(g(x))$。

定义 2.5 对于函数$f(x)$，如果存在常数$c>0$，使得$f(x)=O(x^{c})$，则$f(x)=\mathrm{poly}(x)$。

定义 2.6 对于函数$f(x)$，如果对任意常数$c>0$，均有$f(x)=o(x^{-c})$，则称$f(x)$为x的可忽略函数，记作$f(x)=\mathrm{negl}(x)$。

定义 2.7 如果一个概率可以用可忽略函数来表示，那么就称这个概率为可忽略概率。如果一个概率可以用$1-\mathrm{negl}(x)$来表示，那么就称这个概率为足够大的概率。

定义 2.8 在有限的值域空间Ω中，两个随机变量X和Y的概率分布的统计距离定义为：

$$SD(X,\ Y)=\frac{1}{2}\sum_{\omega\in\Omega}|\ \Pr[X=\omega]-\Pr[Y=\omega]\ |$$

如果两个概率分布的统计距离不超过ε，则称它们为ε-统计接近的。

定义 2.9 对于两组随机变量$\{X_n\}$，$\{Y_n\}$，如果它们的统计距离是n的可忽略函数，那么就称这两组概率分布是统计不可区分的。

定义 2.10 对于两组随机变量$\{X_n\}$，$\{Y_n\}$，如果对于所有的多项式时间算法$\mathcal{A}$，均有

$$|\ \Pr[\mathcal{A}(1^{n},\ X_n)]-\Pr[\mathcal{A}(1^{n},\ Y_n)]\ |=\mathrm{negl}(n)$$

那么就称这两组概率分布是计算不可区分的。

定义 2.11 向量$\boldsymbol{a}$的n-范表示为：

$$\|\boldsymbol{a}\|_{n}=\sqrt[n]{\sum_{i}|\ a_i\ |^{n}}$$

通常用向量的2-范（又称欧几里得范，简写为$\|\boldsymbol{a}\|$）来表示向量的长度，向量的无穷范（即$\|\boldsymbol{a}\|_{\infty}$）与该向量绝对值最大分量的绝对值相等。

定义 2.12 矩阵的长度定义为其长度最长的列向量的长度：

$$\|\boldsymbol{X}\|=\max_{i}\|\boldsymbol{x}_i\|$$

可以近似地将向量集合S的长度定义为其中长度最长的向量的长度。

2.2 格的基本定义

定义 2.13 令 $\mathbb{R}^m$ 为一 m 维欧氏空间。给定 n 个线性无关向量 $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n \in \mathbb{R}^m$，由这些向量生成的格 Λ，记作 $\mathcal{L}(\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n)$，定义如下：

$$\mathcal{L}(\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n) = \{\sum_{i=1}^{n} x_i \boldsymbol{b}_i : \boldsymbol{x}_i \in \mathbb{Z}\}$$

称 m 为格 Λ 的维数，n 为格 Λ 的秩。

定义 2.14 定义 2.13 中生成格 Λ 的线性无关向量组 $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n\}$ 称为格 Λ 的一组基向量（格基），可用如下的矩阵表示：

$$\boldsymbol{B} = [\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n] \in \mathbb{Z}_q^{m \times n}$$

其中每列为基的一个向量。因此，格也可以通过如下的形式表示：

$$\mathcal{L}(\boldsymbol{B}) = \{\boldsymbol{B} \cdot \boldsymbol{x} : \boldsymbol{x} \in \mathbb{Z}^n\}$$

$\tilde{\boldsymbol{B}}$ 表示由 $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n$ 经过格拉姆-施密特正交化得到的向量集。

定义 2.15 对于格 $\mathcal{L}(\boldsymbol{B})$ 和其一组基 $\boldsymbol{b} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n\}$，格 $\mathcal{L}(\boldsymbol{B})$ 关于这组基的基本平行六面体为如下的半开平行六面体：

$$\mathcal{P}(\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n) = \{\sum_{i=1}^{n} x_i \boldsymbol{b}_i : x_i \in [0, 1)\}$$

定义 2.16 格 Λ 的基本平行六面体的容积称为 Λ 的行列式 $\det(\Lambda)$。

命题 2.17 格 Λ 的行列式 $\det(\Lambda)$ 等于它的一组基 $\boldsymbol{B}$ 的行列式 $\det(\boldsymbol{B})$。

此结论可由基本平行六面体和格的行列式的定义直接得到。

定义 2.18 对于矩阵 $\boldsymbol{A} \in \mathbb{Z}_q^{n \times m}$，其中 $m = O(n\log n)$，$q = q(n)$ 是一个诸如 $O(n^2)$ 的次数较小的多项式，$\mathbb{Z}_q^m$ 上满足如下关系的向量 $\boldsymbol{e}$ 的集合称为模格：

$$\Lambda_q^{\perp}(A) = \{e \in \mathbb{Z}_q^m \mid \boldsymbol{A} \cdot e = \boldsymbol{0} \bmod q\}$$

对于向量 $\boldsymbol{u} \in \mathbb{Z}_q^n$，定义：

$$\Lambda_q^{u}(A) = \{\boldsymbol{e} \in \mathbb{Z}_q^m \mid \boldsymbol{A} \cdot \boldsymbol{e} = \boldsymbol{u} \bmod q\}$$

$$f_A(e) = \boldsymbol{A} \cdot \boldsymbol{e} \bmod q (e \in \mathbb{Z}_q^m)$$

定义 2.19 格 Λ^* 称为格 Λ 的对偶格，如果 Λ^* 的基的组成向量 $\boldsymbol{B}^* = [\boldsymbol{b}_1^*, \boldsymbol{b}_2^*, \cdots, \boldsymbol{b}_n^*]$ 满足对所有的 $\boldsymbol{b} \in \boldsymbol{B}$ 和所有的 $\boldsymbol{b}^* \in \boldsymbol{B}^*$，有 $\langle \boldsymbol{b}, \boldsymbol{b}^* \rangle \in \mathbb{Z}$。

定义 2.20 给定格 Λ 的一组基 $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_n\}$，其对偶基 $\{\boldsymbol{b}_1^*, \boldsymbol{b}_2^*, \cdots, \boldsymbol{b}_n^*\}$ 满足

$$i = j \Leftrightarrow \langle \boldsymbol{b}_i, \boldsymbol{b}_j^* \rangle = 1,\ i \neq j \Leftrightarrow \langle \boldsymbol{b}_i, \boldsymbol{b}_j^* \rangle = 0 (i, j \in [1, n])$$

定义 2.21 对于 $n \in \mathbb{Z}^+$，令 $\Lambda \in \mathbb{R}^n$ 为 n 维格。对于 $\mathbb{R}^n$ 上的任意向量 $\boldsymbol{c}$ 及任意实数 $\sigma > 0$，定义

$$\rho_{\sigma, c}(x) = \exp\left(\frac{-\pi \| x - c \|^2}{\sigma^2}\right)$$

$$\rho_{\sigma, c}(\Lambda) = \sum\nolimits_{x \in \Lambda} \rho_{\sigma, c}(x)$$

则格 Λ 上以 c 为中心，以 $\sigma / \sqrt{2\pi}$ 为标准差的离散高斯分布 $\mathcal{D}_{\Lambda, \sigma, c}(\cdot)$ 表示为：

$$\forall y \in \Lambda,\ \mathcal{D}_{\Lambda, \sigma, c}(y) = \frac{\rho_{\sigma, c}(y)}{\rho_{\sigma, c}(\Lambda)}$$

当 $c = 0$ 时，该分布简写为 $\mathcal{D}_{\Lambda, \sigma}(y)$，称为 $\mathbb{R}$ 上的错误分布，记为 χ。

定义 2.22 如果整数上的概率分布 $\{\chi^m\}$ 满足 $\Pr_{e \leftarrow \chi^m}[|e| > \beta] = \mathrm{negl}(n)$，则称 $\{\chi^m\}$ 为 β-有界分布。

2.3 环的基本定义

2.3.1 空间 H

在标准嵌入（见下文）下考虑分圆数域和理想格时，使用子空间 $H \subseteq \mathbb{C}^{\mathbb{Z}_m^*}$ 会很方便，其中 H 定义如下：

$$H = \{x \in \mathbb{C}^{\mathbb{Z}_m^*} : x_i = \overline{x_{m-i}},\ \forall i \in \mathbb{Z}_m^*\}$$

令 $n = \varphi(m)$，可以验证作为内积空间，H 与 $\mathbb{R}^{[n]}$ 同构。

如果 $\boldsymbol{B} = \{\boldsymbol{b}_j\} \subset H$ 是线性无关的向量（即 H 的 $\mathbb{R}$-basis），它的对偶基 $\boldsymbol{B}^* = \{\boldsymbol{b}_j^*\}$ 满足 $\langle \boldsymbol{b}_j, \overline{\boldsymbol{b}_k^*} \rangle = \delta_{jk}$，其中 δ_{jk} 是克罗内克函数（即当 $j = k$ 时