

# 边等周问题

L. H. Harper

2016 年 3 月 11 日

## 1 基本定义

一个图 (Graph)  $G = (V, E, \partial)$  由顶点集  $V$ 、边集  $E$  和标识每一条边对应一对顶点 (可以相同) 的边界函数  $\partial: E \rightarrow \binom{V}{1} \cup \binom{V}{2}$  构成。图经常用图表 (Diagram) 来表示, 其中顶点对应图表中的点, 边对应图表中连接一对点的曲线。对任意的图  $G$  和  $S \subseteq V$ , 我们定义:

$$\Theta(S) = \{e \in E \mid \partial(e) = \{v, w\}, v \in S, w \notin S\}$$

并称之为  $S$  的边界 (Edge-boundary)。那么对于一个给定的图  $G$  和  $k \in \mathbb{Z}^+$ , 边等周问题 (Edge-Isoperimetric Problem, EIP) 是对于所有的  $S \subseteq V$  并且  $|S| = k$ , 找出最小的  $|\Theta(S)|$ 。注意  $|\Theta(S)|$  是不变量 (Invariant), 即如果  $\phi: G \rightarrow H$  表示一个图同构, 那么对于  $\forall S \subseteq V_G$ , 都有  $|\Theta(\phi(S))| = |\Theta(S)|$ 。因此, 在一个自同构下等价的顶点子集有相同的边界。

自环, 即只对应一个顶点的边, 与 EIP 无关, 因此我们将忽略它们。我们大多数的图 (但并非全部) 是简单图 (Ordinary Graph), 即不含自环和多重边。一个简单图的表示可以缩短至  $(V, E)$ , 其中  $E \subseteq \binom{V}{2}$ ,  $\partial$  则是隐含的。

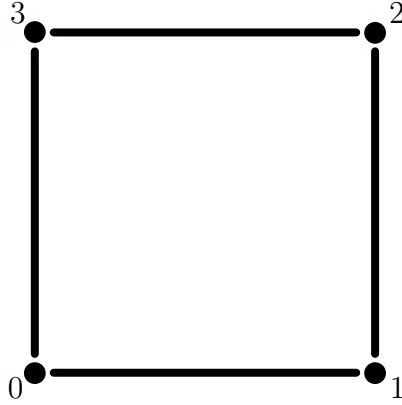


图 1:  $\mathbb{Z}_4$  的图表

## 2 例子

### 2.1 $K_n$ , $n$ -完全图

$K_n$  有  $n$  个顶点且  $E = \binom{V}{2}$ , 即每一对不同顶点之间都有一条边。对于每一个  $S \in V$  且  $|S| = k$ ,  $|\Theta(S)| = |S \times (V-S)| = k(n-k)$ 。因此  $K_n$  的 EIP 很简单: 任意一个  $k$ -集都是解。

### 2.2 $\mathbb{Z}_n$ , $n$ -环

对于  $\mathbb{Z}_n$ , 有  $V = \{0, 1, \dots, n-1\}$  和  $E = \{\{i, j\} \mid i-j \equiv \pm 1 \pmod{n}\}$ 。因此,  $\mathbb{Z}_3 = K_3$ ,  $\mathbb{Z}_4$  的图表见图 1。

现在我们从下面的一般性说明 (这些说明在后面也会有用), 来推导  $\mathbb{Z}_4$  以及  $\mathbb{Z}_n$  上 EIP 的解:

- (1) 对于  $|S| = k = 0$ , 在任意的图里都只有一个子集, 即空集  $\emptyset$ 。因此  $\min_{|S|=0} |\Theta(S)| = |\Theta(\emptyset)| = 0$ 。
- (2) 对于  $k = |V| = n$ , 也只有一个子集, 即  $V$ 。因此  $\min_{|S|=n} |\Theta(S)| = |\Theta(V)| = 0$ 。
- (3) 一个图被称为  $\delta$  度正则图 (Regular of Degree  $\delta$ ), 当它的每个顶点都正好对应  $\delta$  条边。在一个正则图中, 如果  $|S| = k = 1$  那么  $|\Theta(S)| = \delta$ , 因此

任意单点集都是一个解集。 $\mathbb{Z}_n$  是 2 度正则图；然而对于  $n = 4$  且  $k = 2$  的情况，有两个集合并不在  $\mathbb{Z}_n$  的对称性下等价： $\{0, 1\}$  和  $\{0, 2\}$ 。所有其他的 2-集都与这两者中的一个等价。 $|\Theta(\{0, 1\})| = 2$ ， $|\Theta(\{0, 2\})| = 4$ ，因此  $\min_{|S|=2} |\Theta(S)| = 2$ 。

(4) 对于  $\forall G$  和  $\forall S \subseteq V$ ，有

$$\Theta(V-S) = \Theta(S)$$

因此对于  $k > \frac{1}{2}|V|$ ， $\min_{|S|=k} |\Theta(S)| = \min_{|S|=n-k} |\Theta(S)|$ ，其中  $n = |V|$ 。这样我们就完成了  $\mathbb{Z}_4$  上 EIP 的解，见下表总结。

$k$	0	1	2	3	4
$\min_{ S =k}  \Theta(S) $	0	2	2	2	0

(5) 令

$$E(S) = \{e \in E \mid \partial(e) = \{v, w\}, v \in S, w \in S\}$$

$E(S)$  被称为  $S$  的导出边 (Induced Edges)。对于一个图，导出边问题 (Induced Edge Problem) 是对于所有的  $S \subseteq V$  并且  $|S| = k$ ，找出最大的  $|E(S)|$ 。

**引理 1.** 如果图  $G = (V, E, \partial)$  是一个  $\delta$  度正则图，那么对于  $\forall S \subseteq V$ ，都有

$$|\Theta(S)| + 2|E(S)| = \delta|S|$$

证明.  $\delta|S|$  表示  $S$  对应的边数，然而出现在  $E(S)$  中的边被计算了两次。□

**推论 1.** 如果  $G$  是一个正则图，那么  $S \subseteq V$  是导出边问题的一个解当且仅当它同时也是 EIP 的解。而且，对于  $\forall k$ ， $\min_{|S|=k} |\Theta(S)| = \delta k - 2 \max_{|S|=k} |E(S)|$ 。

那么对于正则图，EIP 和导出边问题是等价的，我们将视它们为可等价互换的问题。通常 EIP 出现在实际应用中，而导出边问题则更易证明。EIP 还存在第三种自然变体：对于  $S \subseteq V$ ，令

$$\partial^*(S) = \{e \in E \mid \partial(e) \cap S \neq \emptyset\}$$

即  $S$  中的点对应的边的集合。

练习 1. 对于正则图，证明计算

$$\min_{\substack{S \subseteq V \\ |S|=k}} |\partial^*(S)|$$

与  $EIP$  等价。

回想一下，一棵树 (Tree) 是连通无环图。一个无环图也被称为森林 (Forest)，因为它是树——其连通分支的并集。

引理 2. 有  $n$  个顶点的树的边数为  $n-1$ 。有  $n$  个顶点的森林的导出边数则是  $n-t$ ，其中  $t$  为其连通分支数。

$\mathbb{Z}_n$  的任意适当的子集  $S$  都可以导出一个无环图，因此  $\max_{|S|=k} |\Theta(S)|$  会出现于当  $S$  是一个连通的集合，即区间时。因而，如果  $0 < k < n$ ， $\min_{|S|=k} |\Theta(S)| = 2k - 2(k-1) = 2$ 。

## 2.3 $Q_d$ , $d$ 阶立方

$d$  阶立方  $Q_d$  有顶点集  $\{0, 1\}^d$ ，即  $\{0, 1\}$  的  $d$  倍笛卡尔积。因此  $n = |V_{Q_d}| = 2^d$ 。 $Q_d$  中的两个顶点 (0 和 1 构成的  $d$  元组) 对应一条边，当且仅当它们正好有一位不相同。

练习 2. 找出  $m = |E_{Q_d}|$  对应的公式。

$Q_1$  和  $K_2$  同构， $Q_2$  和  $\mathbb{Z}_4$  同构，这些都已经可以通过  $EIP$  得到解决。一个 3 阶立方有 8 个顶点，12 条边，以及 6 个平面。 $Q_3$  的图表事实上是一个 3 阶立方的投影，见图 2。

我们可以用在前两个例子中发展出来的简便工具来解决  $Q_3$  上的  $EIP$ 。首先观察到  $Q_3$  有围长 (最短回路的长度) 为 4：因为 3 阶立方的对称群是传递的，所以任意顶点都和其他顶点一样。从一个顶点出发勾勒出路径，我们可以看到不存在长度为 3 的闭合回路。因此对于  $1 \leq k \leq 3$ ，根据引理 1 和引理 2，我们有

$$\begin{aligned} \min_{|S|=k} |\Theta(S)| &= 3k - 2 \max_{|S|=k} |E(S)| \\ &= 3k - 2(k-1) = k + 2 \end{aligned}$$

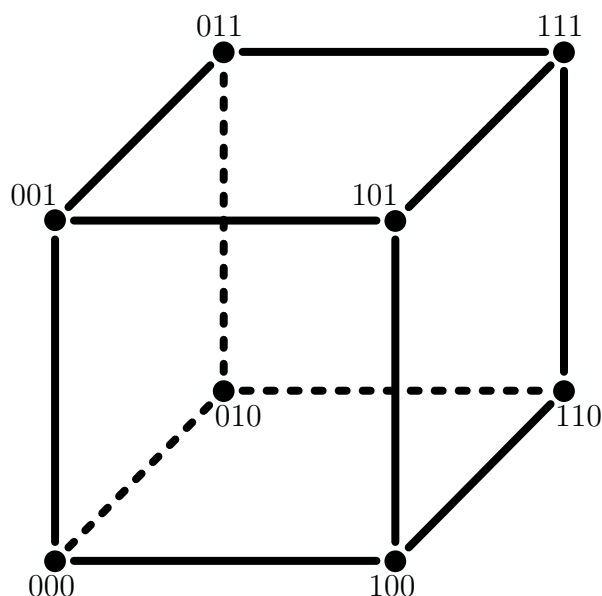


图 2:  $Q_3$  的图表

对于  $k = 4$ , 要么  $S$  导出一个回路, 在这种情况下是一个 4-环, 并且  $|\Theta(S)| = 4$ ; 要么  $S$  导出一个无环图, 并且根据上述可以得到  $|\Theta(S)| \geq 6$ 。对于  $k > 4 = \frac{8}{2}$ , 我们根据下面这个式子得到解

$$\min_{|S|=k} |\Theta(S)| = \min_{|S|=8-k} |\Theta(S)|$$

最后的解见下表总结。

$k$	0	1	2	3	4	5	6	7	8
$\min_{ S =k}  \Theta(S) $	0	3	4	5	4	5	4	3	0

为了将这个 EIP 的解推广到  $Q_d$  ( $d > 3$ ), 我们需要利用一些关于立方的简单结论, 这些结论的证明将被留作练习。一个  $d$  阶立方的  $c$  阶子立方 ( $c$ -subcube of the  $d$ -cube) 是  $Q_d$  的子图, 导出自一个包含所有在  $d-c$  个坐标下有相同 (固定) 值的顶点的集合。

**练习 3.** 证明  $d$  阶立方的任意  $c$  阶子立方与  $c$  阶立方同构。

**练习 4.** 计算  $d$  阶立方有多少个  $c$  阶子立方?

一个  $d$  阶立方的  $c$  阶子立方的邻居 (Neighbor) 是在  $d-c$  个固定坐标下有正好一个坐标不同的任意  $c$  阶子立方。

练习 5. 证明一个  $c$  阶子立方的所有邻居都互不相交。

练习 6. 证明一个  $c$  阶子立方的两个不同邻居（的顶点）之间没有边相连。

练习 7. 计算一个  $d$  阶立方的  $c$  阶子立方有多少个邻居？

$Q_d$  上的 EIP 最初由数据传输上的问题推动。W. H. Kautz [1]、E. C. Posner 和本书作者的研究引发了一种猜测：字典序编号方式（Lexicographic Numbering）的起始部分，即对于  $x \in V_{Q_d}$ ,

$$\text{lex}(x) = 1 + \sum_{i=1}^d x_i 2^{i-1}$$

是解集，然而这要如何证明？一个显而易见的尝试方法是对维度  $d$  使用归纳法。数学归纳法有一个看似自相矛盾的属性，它往往更容易证明一个更强的定理，因为一旦最初的情况得到验证，人们就可以假设该定理对归纳参数的较低值为真，以便建立下一次归纳。因此一个较强的假设可以产生一个更简单的证明。在本例中，该策略引发了这种推测——以下归纳步骤会生成所有解集：

- (1) 首先从空集  $\emptyset$  开始。
- (2) 对于已经构建好的集合  $S \subset V_{Q_d}$ ，选择任意一个  $x \in V_{Q_d} - S$  来扩充  $S$ ，使得导出边数的增量最大，即

$$|E(S \cup \{x\})| - |E(S)|$$

添加任意  $x \in V_{Q_d}$  都可使  $\emptyset$  的增量最大，因为  $|E(\{x\})| - |E(\emptyset)| = 0$ 。 $\{x\}$  的增量则必须是  $x$  的相邻顶点。那么对于  $k > 2$  的  $k$ -集，情况又该如何呢？答案是如果  $k = 2^c$ ，那么该集合必然是  $c$  阶子立方。我们刚刚已经验证了  $c = 0$  和  $1$  的情况。假设这对  $0, 1, \dots, c-1$  都成立，为了扩充一个  $2^{c-1}$ -集，即一个  $(c-1)$  阶子立方，我们只能选择一个使  $|E(S)|$  增量为  $1$  的顶点，即一个相邻的  $(c-1)$  阶子立方中的任意顶点。从一个相邻的子立方中选出一个顶点后，我们必须继续从同一个子立方中选点，直到用尽它所有的顶点，这是因为所选择的子立方中总会存在一个顶点使得  $|E(S \cup x)| - |E(S)| \geq 2$ ，

而其他子立方中的任意顶点都有  $|E(S \cup x)| - |E(S)| \leq 2$ 。当我们用尽所有相邻的  $(c-1)$  阶子立方时，我们就得到了一个  $c$  阶子立方。

总体而言，令

$$k = \sum_{i=1}^K 2^{c_i}, 0 \leq c_1 < c_2 < \cdots < c_K \leq \lfloor \log_2 k \rfloor$$

为  $k$  的二进制表示形式（注意  $K \leq \lfloor \log_2 k \rfloor$ ）。如果  $S \subseteq V_{Q_d}$  是一个互不相交的子立方的并集，其中每个  $c_i$  阶子立方（ $1 \leq i \leq K$ ）都满足：对于其他所有  $c_i$  阶子立方（ $j > i$ ），该  $c_i$  阶子立方处在该  $c_j$  阶子立方的一个邻居中，那么  $S$  就被称为立方集（Cubal Set）。立方集正好是依次使内部边数增量最大化所构造出来的集合。如果  $S$  是立方集且  $|S| = k$ ，那么（见练习 1）

$$|E(S)| = \sum_{i=1}^K (K - i) 2^{c_i} + c_i 2^{c_i - 1}$$

注意对于一个  $k$ -立方集  $S \subseteq V_{Q_d}$ ， $|E(S)|$  并非由  $d$  决定，而仅仅由  $k = |S|$  决定。这个函数很重要，我们把它记为  $E(k)$ 。 $E(k)$  有一个分形性质，见下面这个递推式：如果  $2^{d-1} \leq k < 2^d$ ，那么

$$E(k+1) - E(k) = E(k - 2^{d-1} + 1) - E(k - 2^{d-1}) + 1$$

这遵循  $k$ -立方集的递归结构。从  $k$  中减去最大的 2 的幂—— $2^{d-1}$ ，对应从  $S$  中去掉最大的子立方。这个子立方为集合中每一个余下的顶点提供了一个邻居。

**练习 8.** 证明如果  $S \subseteq V_{Q_d}$  是立方集，那么它的补集  $V_{Q_d} - S$  也是立方集。

**练习 9.** 证明任意两个  $k$ -立方集都同构，即存在能把一个  $d$  阶立方映射到另一个的自同构。

**定理 1.**  $S \subseteq V_{Q_d}$  在基数为  $k$  时有最大的  $|E(S)|$ ，当且仅当  $S$  是立方集。

**引理 3.** (Bernstein [2]) 对于  $\forall d$  和  $\forall k, t > 0$  使得  $k + t < 2^d$ ，都有

$$E(t) < E(k+t) - E(k) < E(2^d) - E(2^d - t)$$

引理 3 的证明. 对  $d$  进行归纳: 当  $d = 2$  时成立; 假设  $d-1 \geq 2$  时也成立, 考虑下面三种情况:

(1) 如果  $k \geq 2^{d-1}$ , 那么有

$$\begin{aligned} E(k+t)-E(k) &= \sum_{i=1}^t E(k+i)-E(k+i-1) \\ &= \sum_{i=1}^t [E(k+i-2^{d-1})-E(k+i-2^{d-1}-1)+1] \\ &= E(k+t-2^{d-1})-E(k-2^{d-1})+t \end{aligned}$$

并且两个不等式都遵循归纳假设。

(2) 如果  $k+t \leq 2^{d-1}$ , 那么左边的不等式遵循归纳假设; 右边的不等式遵循上述恒等式, 因而也遵循归纳假设。

(3) 如果  $k < 2_{d-1} < k+t$ , 那么

$$\begin{aligned} E(k+t)-E(k) &= [E(k+t)-E(2^{d-1})] + [E(2^{d-1})-E(k)] \\ &> [E(k+t-2^{d-1})] + [E(t)-E(t-(2^{d-1}-k))] \end{aligned}$$

分别根据情况 1 和 2

$$= E(t)$$

练习 10. 完成情况 3 的证明 (右边的不等式)。

□

定理 1 的证明. 我们已经注意到所有的  $k$ -立方集都有相同的导出边数, 即  $E(k)$ , 因此我们只需证明所有的最优集都是立方集。先前我们对  $d$  进行归纳, 并已证明了当  $d = 1, 2$  时为真。现在假设当维度  $d > 2$  时也为真。对于给定的  $k$  ( $0 < k < n = 2^d$ ) 并用上述 2 的幂的和来表示 (所以  $K < d$ ), 令  $S \subseteq V_{Q_d}$  为  $|S| = k$  时的最优集。如果我们根据第  $d$  位坐标, 把  $Q_d$  划分为两个  $(d-1)$  阶子立方  $Q_{d,0} = \{x \in V_{Q_d} \mid x_d = 0\}$  和  $Q_{d,1} = \{x \in V_{Q_d} \mid x_d = 1\}$ , 那么我们也得到了  $S$  的划分, 即  $S_0 = S \cap Q_{d,0}$  和  $S_1 = S \cap Q_{d,1}$ 。令  $|S_0| = k_0$ ,



$|S_1| = k_1$ ，我们不妨假设  $k_0 \geq k_1$ 。如果  $k_1 = 0$ ，定理遵循归纳假设，因此就假设  $k_1 > 0$ 。因为  $E(S)$  中的边的端点要么都在  $S_0$  中，要么都在  $S_1$  中，要么一个在  $S_0$  中，一个在  $S_1$  中，所以

$$|E(S)| \leq \max_{|S|=k_0} |E(S)| + \max_{|S|=k_1} |E(S)| + k_1$$

如果  $S_1$  是基数为  $k_1$  的立方集，那么根据归纳，我们有  $|E(S_1)| = E(k_1) = \max_{|S|=k_1} |E(S)|$ 。  $S_1$  在  $Q_{d,0}$  中的邻居与  $S_1$  同构，因此也是立方集。根据立方集的递归构造， $\exists S_0 \in Q_{d,0}$  且  $|S_0| = k_0$ ，它是立方集，并且包含  $S_1$  的邻居。因此我们可以取到  $|E(S)|$  的上界，而且每个取到上界的集合都必须是两个立方集的并集。令  $k_0 = \sum_{i=1}^{K_0} 2^{c_{i,0}}, 0 \leq c_{1,0} \leq c_{2,0} \leq \dots \leq c_{K_0,0}$ ， $k_1$  与其类似。因为  $k_0 + k_1 = k$ ，所以只有三种可能的情况：

- (1)  $c_{K_0,0} = c_K$ ：那么我们可以假设  $S_0 - Q_{c_{K_0,0}}$  和  $S_1$  处在  $Q_{c_{K_0,0}}$  的两个不同的邻居中，因此  $S$  不是立方集。根据练习 5 可知  $S_0 - Q_{c_{K_0,0}}$  中的顶点和  $S_1$  中的顶点没有边相连。令  $k'_0 = k_0 - 2^{c_{K_0,0}} > 0$ ，我们有  $k'_0 + k_1 \leq 2^{c_{K_0,0}}$ 。如果我们去掉  $S_1$  并添加相同数量的顶点至  $S_0$ ，以此来修改  $S$ ，那么引理 3 表明  $|E(S)|$  会增大。这便与  $S$  是最优的相矛盾。
- (2)  $c_{K_0,0} = c_K - 1$  且  $c_{K_1,1} = c_K - 1$ ：  $Q_{c_{K_0,0}}$  和  $Q_{c_{K_1,1}}$ ，这两个  $(c_K - 1)$  阶子立方是邻居，因此构成了一个  $c_K$  阶子立方。  $S_0 - Q_{c_{K_0,0}}$  和  $S_1 - Q_{c_{K_1,1}}$  每一个都处在相邻的  $(c_K - 1)$  阶子立方中，这两个子立方构成了一个与第一个相邻的  $c_K$  阶子立方。根据归纳假设， $S$  肯定是立方集。
- (3)  $c_{K_0,0} = c_K - 1$  且  $c_{K_1,1} < c_K - 1$ ：与情况 1 同理，我们假设  $S_0 - Q_{c_{K_0,0}}$  和  $S_1$  处在  $Q_{c_{K_0,0}}$  的两个不同的邻居中，因此它们之间没有边相连。令  $k'_0 = k_0 - 2^{c_{K_0,0}} > 0$ ，我们有

$$\begin{aligned} k'_0 + k_1 &= k_0 - 2^{c_{K_0,0}} + k_1 \\ &= k - 2^{c_K - 1} \\ &\geq 2^{c_K} - 2^{c_K - 1} \\ &= 2^{c_K - 1} \end{aligned}$$

如果我们从  $S_1$  中去掉  $2^{c_{K_0}} - k'_0$  个顶点并把它们补充到包含  $S_0 - Q_{c_{K_0},0}$  的那个集合中, 使其成为  $Q_{c_{K_0},0}$  的邻居, 以此来修改  $S$ , 那么引理 3 表明这会使  $|E(S)|$  增大, 再一次与  $S$  是最优的相矛盾。

□

### 3 布局问题上的应用

组合等周问题频繁地出现在通信工程、计算机科学、物理科学和数学本身这些领域。我们不希望在这里涵盖所有的应用, 但会给出一个具有代表性的样例。我们选择专注于布局问题, 它们出现在电气工程中, 当一个人把一些电气回路的布线图在机箱上“勾画出来”, 即把每一个元件和线放置在机箱上。布线图本质上就是一个图, 其中电气元件成了顶点, 连接它们的线成了边。对于机箱上的顶点和边的任意给定布局, 都存在一定的成本或者说性能度量, 我们则希望对其进行优化。

#### 3.1 线长问题

假设我们想把元件 (图  $G = (V, E, \partial)$  中的顶点) 放置在一个线形的机箱上, 每一个都与先前的那个元件相距一个单位, 以这样的方式来最小化连接它们的线的总长度。精确地说, 我们将  $G$  的一种顶点编号方式 (Vertex-numbering) 定义为一个一对一的函数

$$\eta: E \rightarrow \{1, 2, \dots, n\}, \text{ 其中 } n = |V|$$

$\eta$  的值域中的整数可根据线形机箱上的位置来区分。那么  $\eta$  的总线长 (Wire-length) 则是

$$wl(\eta) = \sum_{\substack{e \in E \\ \partial(e) = \{v, w\}}} |\eta(v) - \eta(w)|$$

对于一个图  $G = (V, E, \partial)$  来说,

$$wl(G) = \min\{wl(\eta) \mid \eta \text{ 是 } G \text{ 的一种顶点编号方式}\}$$

记住, 一个有  $n$  个顶点的图有  $n!$  种顶点编号方式。

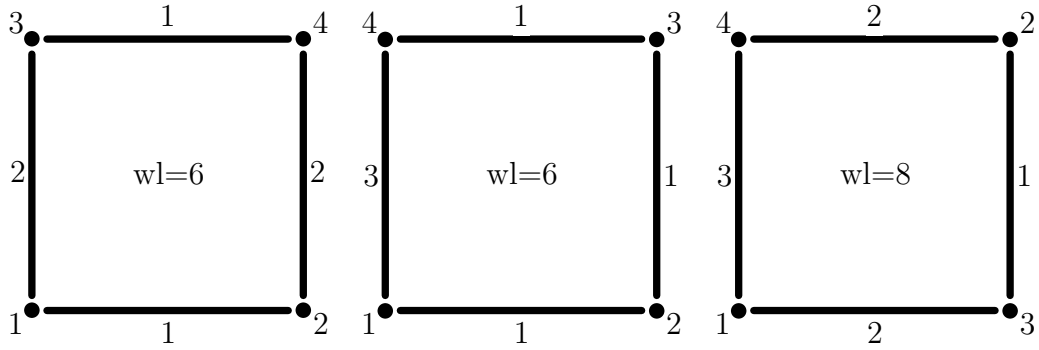


图 3:  $Q_2$  的编号方式

### 3.1.1 例子

平方的图有  $4! = 24$  种顶点编号方式，但它同时还有 8 种对称情况。任意两种对称的编号方式都有相同的线长。图 3 的三种编号方式代表了  $24/8 = 3$  种编号方式的等价类。前两种编号方式有最小的线长  $wl$ ，第三种则有最大的。因此  $wl(Q_2) = 6$ 。然而  $wl(Q_3)$  就没有那么容易得出了，因为  $Q_3$  有  $8! = 40320$  种编号方式以及 48 种对称情况。虽然  $40320/48 = 840$  也并不是那么大，但是要如何才能系统地生成这 840 种编号方式等价类的代表呢？我们现在来看看如何解决这些明显的困难。

为了最小化一个和，譬如  $wl$ ，一个显而易见的策略是分别最小化每一个加数。这些最小值的和就成了和的最小值的下界，并且我们可以期望这会是一个不错的下界，甚至是精确的。然而这并不适用于  $wl(\eta)$  的定义，因为对于每一条边  $e \in E$ ， $\partial(e) = \{v, w\}$ ，都有

$$\min_{\eta} |\eta(v) - \eta(w)| = 1$$

### 3.1.2 $wl$ 的另一种表示形式

对于给定的一个编号方式  $\eta$  和一个整数  $k$ ， $0 \leq k \leq n$ ，令

$$S_k(\eta) = \eta^{-1}(\{1, 2, \dots, k\}) = \{v \in V \mid \eta(v) \leq k\}$$

即编号方式  $\eta$  下的前  $k$  个顶点的集合。那么下面我们就得到了线长的另一种表示形式。

引理 4.

$$wl(\eta) = \sum_{k=0}^n |\Theta(S_k(\eta))|$$

证明. 注意  $S_0(\eta) = \eta^{-1}(\emptyset) = \emptyset$ . 令

$$\chi(e, k) = \begin{cases} 1 & \text{如果 } \partial(e) = \{v, w\}, \eta(v) \leq k < \eta(w) \\ 0 & \text{其他情况} \end{cases}$$

那么

$$\begin{aligned} wl(\eta) &= \sum_{\substack{e \in E \\ \partial(e) = \{v, w\}}} |\eta(v) - \eta(w)| = \sum_{e \in E} \sum_{k=0}^n \chi(e, k) \\ &= \sum_{k=0}^n \sum_{e \in E} \chi(e, k) = \sum_{k=0}^n |\Theta(S_k(\eta))| \end{aligned}$$

□

练习 11. (*Steiglitz-Bernstein [3]*) 假设线形机箱实际上是一条实线, 上面有  $n$  个位置  $s_1 < s_2 < \dots < s_n$  用来放置元件. 布局  $\eta$  将顶点  $v \in V$  放置在位置  $s_{\eta(v)}$  上, 证明  $\eta$  有总线长为

$$wl(\eta) = \sum_{k=0}^n (s_{k+1} - s_k) |\Theta(S_k(\eta))|$$

线长的这种作为一个和的新表示形式给了我们它的另一个下界。

推论 2. 对于任意的图  $G$ ,

$$\begin{aligned} wl(G) &= \min_{\eta} wl(\eta) = \min_{\eta} \sum_{k=0}^n |\Theta(S_k(\eta))| \\ &\geq \sum_{k=0}^n \min_{\eta} |\Theta(S_k(\eta))| = \sum_{k=0}^n \min_{|S|=k} |\Theta(S)| \end{aligned}$$

定理 2. 对于  $G = (V, E, \partial)$  的任意顶点编号方式  $\eta_0$ , 如果它的所有初始段  $S_k(\eta_0)$  ( $0 \leq k \leq n$ ) 都是  $G$  上 EIP 的解, 那么它本身就是  $G$  上线长问题的一个解。

推论 3.  $\mathbb{Z}_n$  的编号方式  $\eta_0(i) = 1 + i$  是  $\mathbb{Z}_n$  上线长问题的一个解, 并且

$$\begin{aligned} wl(\mathbb{Z}_n) &= \sum_{k=0}^n \min_{|S|=k} |\Theta(S)| \\ &= \sum_{k=1}^{n-1} 2 = 2(n-1) \end{aligned}$$

定义 1. 如果  $\{T_i\}_{i=1}^n$  是一个完全有序集合的序列, 那么它们乘积的字典序  $T_1 \times T_2 \times \cdots \times T_n$  是一个由  $x < y$  定义的全序,  $x < y$  成立当  $\exists m$  使得  $x_1 = y_1, x_2 = y_2, \dots, x_{m-1} = y_{m-1}$  且  $x_m < y_m$ 。一个集合的任意全序都对应一个编号方式, 其中集合的第一个 (最小的) 元素编号为 1, 第二个编号为 2, 以此类推。0 和 1 构成的  $d$  元组 ( $Q_d$  的顶点) 上的字典序对应的编号方式是  $lex(x) = 1 + \sum_{i=1}^d x_i 2^{i-1}$

推论 4.  $Q_d$  的字典序编号方式是  $Q_d$  上线长问题的一个解。

练习 12. 证明  $wl(Q_d) = 2^{d-1}(2^d - 1)$ 。

练习 13. 对于练习 11 中的更一般的线形机箱, 证明  $lex$  同时也是它的  $Q_d$  上线长问题的一个解。

### 3.2 4 阶德布鲁因图

德布鲁因图是一个有向图 (Directed Graph, Digraph), 即存在两个边界函数  $\partial_{\pm}: E \rightarrow V$  分别标识每一条边的头和尾。有向图的图表中的每条边一般都有一个箭头指向它的头。 $d$  阶德布鲁因图 (the deBruijn Graph of Order  $d$ )  $DB_d$  有与  $Q_d$  相同的顶点集 (0 和 1 构成的所有  $d$  元组的集合), 而它的边集则完全不同。 $E_{DB_d}$  是 0 和 1 构成的所有  $(d+1)$  元组的集合, 并且

$$\begin{aligned} \partial_{-}(x_1, x_2, \dots, x_d, x_{d+1}) &= (x_1, x_2, \dots, x_d) \\ \partial_{+}(x_1, x_2, \dots, x_d, x_{d+1}) &= (x_2, \dots, x_d, x_{d+1}) \end{aligned}$$

关于德布鲁因图的更多信息, 请参考 [4]。 $DB_4$  的图表见图 4。

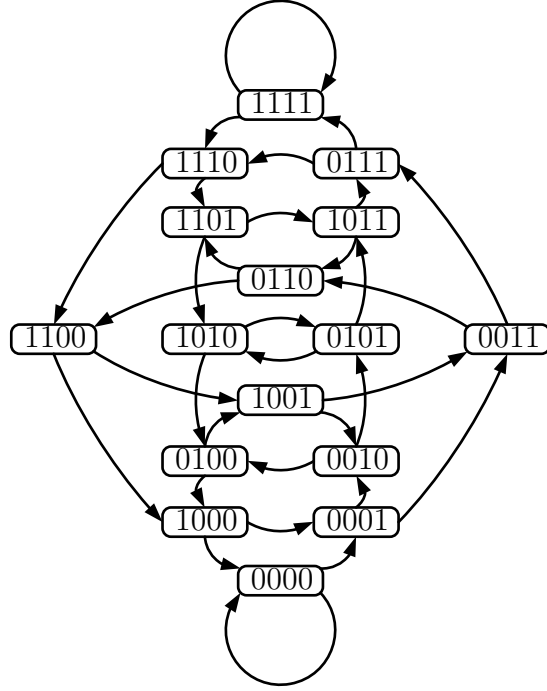


图 4: 4 阶德布鲁因图

对于一个编号方式  $\eta$ ，它的作为边界的和的线长表示形式，及其在求解  $\mathbb{Z}_n$  和  $Q_d$  上的线长问题上的应用，表明了一种可最小化任意图的线长的启发式方法：给顶点编号为  $1, 2, \dots, k-1, k, \dots, n$ ，从而最小化

$$|\Theta(S_k(\eta))| - |\Theta(S_{k-1}(\eta))|$$

即每增加一个顶点所带来的边际增量。把这个启发式方法应用到  $\underline{DB}_4$ （去掉自环和边的方向的  $DB_4$ ）上，我们就得到了编号方式  $\eta_0$ ，见图 5。

下表列出了  $k$  的取值范围在 0 到 16 时  $|\Theta(S_k(\eta_0))|$  的值：

$k$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$ \Theta(S_k(\eta_0)) $	0	2	4	4	6	6	6	6	6	6	8	6	6	4	4	2	0

从这可以看出  $S_{10}(\eta_0)$  不是 EIP 的解。事实上并不存在一个所有的初始段都是 EIP 的解的编号方式，因为如果存在的话，它就会产生于我们用来构造  $\eta_0$  的这个过程中。然而，我们可以断言，对于  $k \neq 10$ ， $S_k(\eta_0)$  是 EIP 的

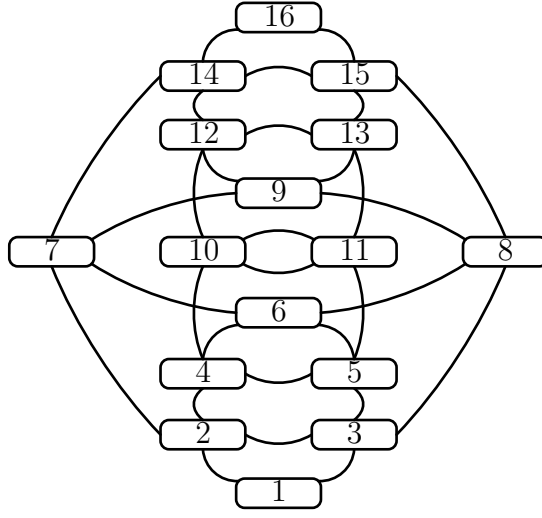


图 5: 编号方式  $\eta_0$  下  $\underline{DB}_4$  的图表

解。因此有不等式可导致定理 2，即

$$wl(\underline{DB}_4) > \sum_{k=0}^n \min_{|S|=k} |\Theta(S)| = 2 \times 34 + 6 = 74$$

而

$$wl(\eta_0) = 76$$

因为对于每一个顶点  $v \in V_{\underline{DB}_4}$ ，度数  $\delta(v)$  都是偶数，并且对于所有的  $S \subseteq V_{\underline{DB}_4}$ ，

$$|\Theta(S)| = \sum_{v \in S} \delta(v) - 2|E(S)|$$

也是偶数（引理 1 中的恒等式的泛化），所以对于任意  $\underline{DB}_4$  的编号方式  $\eta$ ，我们都有  $wl(\eta) \geq 76$ ，因此  $\eta_0$  是它的线长问题的一个解。

**练习 14.** 找出  $\underline{DB}_4$  上所有的  $EIP$  的解集，其中  $k \leq 8$ （无需证明）。

### 3.3 划分问题

$G = (V, E, \partial)$  的一个划分（Partition）是一个集合  $\pi \subseteq 2^V$  满足

- (1)  $\forall B \in \pi, B \neq \emptyset$ ;

(2)  $\forall A, B \in \pi$ , 要么  $A = B$ , 要么  $A \cap B = \emptyset$ ;

(3)  $\bigcup_{B \in \pi} B = V$ 。

$B \in \pi$  被称为划分的一个块 (Block)。一个划分  $\pi$  被认为是均匀的 (Uniform), 当  $\forall A, B \in \pi$ ,  $||A| - |B|| \leq 1$ 。如果  $|\pi| = p$ , 那么这就等价于

$$\forall B \in \pi, \left\lfloor \frac{n}{p} \right\rfloor \leq |B| \leq \left\lceil \frac{n}{p} \right\rceil, \text{ 其中 } n = |V|$$

事实上, 如果  $j$  是  $n$  除以  $p$  的余数 (即  $n \equiv j \pmod{p}$ ,  $0 \leq j < p$ ), 那么  $G$  的任意均匀  $p$ -划分都有  $j$  个  $\lceil \frac{n}{p} \rceil$ -块和  $(p - j)$  个  $\lfloor \frac{n}{p} \rfloor$ -块。

一个划分  $\pi$  的边界 (Edge-boundary) 是

$$\begin{aligned} \Theta(\pi) &= \{e \in E \mid \partial(e) = \{v, w\}, v \in A, w \in B, A \neq B\} \\ &= \bigcup_{B \in \pi} \Theta(B) \end{aligned}$$

注意后一种表示形式中的并集不是互不相交的, 而是任意  $e \in E$  且  $\partial(e) = \{v, w\}$  都正好被包含了两次, 一次来自包含  $v$  的块, 另一次来自包含  $w$  的块。

$G$  的边界划分问题 (Edge-boundary Partition Problem) 是对于  $G$  的所有均匀  $p$ -划分  $\pi$ , 找出最小的  $|\Theta(\pi)|$ 。  $G$  可以被当做一个要布局在  $p$  块芯片上的布线图, 其中的元件要尽可能平均地分布在这些芯片上。问题则是给芯片分配元件, 使得连接不同芯片上元件的线数最小。边界划分问题的一个变体是边宽划分问题 (Edge-width Partition Problem), 即对于  $G$  的所有均匀  $p$ -划分  $\pi$ , 找出最小的  $\max_{B \in \pi} |\Theta(B)|$ 。

**引理 5.** 如果  $n \equiv j \pmod{p}$ ,  $0 \leq j < p$ , 那么

$$\min_{\substack{|\pi|=p \\ \pi \text{ 均匀}}} |\Theta(\pi)| \geq \frac{1}{2} \left( j \min_{|B|=\lceil \frac{n}{p} \rceil} |\Theta(B)| + (p - j) \min_{|B|=\lfloor \frac{n}{p} \rfloor} |\Theta(B)| \right)$$

同样还有

$$\min_{|\pi|=p} \max_{B \in \pi} |\Theta(B)| \geq \max \left\{ \min_{|B|=\lceil \frac{n}{p} \rceil} |\Theta(B)|, \min_{|B|=\lfloor \frac{n}{p} \rfloor} |\Theta(B)| \right\}$$



证明. 如果  $\pi_0$  是一个最优均匀  $p$ -划分, 那么

$$\begin{aligned}\min_{|\pi|=p} |\Theta(\pi)| &= |\Theta(\pi_0)| \\ &= \frac{1}{2} \sum_{B \in \pi_0} |\Theta(B)|\end{aligned}$$

因为每一条边  $e \in \Theta(\pi_0)$  都被计算了两次

$$\geq \frac{1}{2} \left( j \min_{|B|=\lceil \frac{n}{p} \rceil} |\Theta(B)| + (p-j) \min_{|B|=\lfloor \frac{n}{p} \rfloor} |\Theta(B)| \right)$$

□

### 3.3.1 例子

(1) 对于  $\mathbb{Z}_n$ , 我们得到

$$\begin{aligned}\min_{|\pi|=p} |\Theta(\pi)| &\geq \frac{1}{2} \left( j \min_{|B|=\lceil \frac{n}{p} \rceil} |\Theta(B)| + (p-j) \min_{|B|=\lfloor \frac{n}{p} \rfloor} |\Theta(B)| \right) \\ &= \frac{1}{2} (j \cdot 2 + (p-j) \cdot 2) = p\end{aligned}$$

同样还有

$$\begin{aligned}\min_{|\pi|=p} \max_{B \in \pi} |\Theta(B)| &\geq \max \left\{ \min_{|B|=\lceil \frac{n}{p} \rceil} |\Theta(B)|, \min_{|B|=\lfloor \frac{n}{p} \rfloor} |\Theta(B)| \right\} \\ &= \max\{2, 2\} = 2\end{aligned}$$

这些下界可以在把  $\mathbb{Z}_n$  均匀地划分为  $p$  个区间时取到, 因此  $\mathbb{Z}_n$  上的边界和边宽划分问题都得到了解决。

(2) 对于  $Q_d$ ,  $p = 2^a$  是 2 的幂,  $\lceil \frac{n}{p} \rceil = \frac{2^d}{2^a} = 2^{d-a} = \lfloor \frac{n}{p} \rfloor$ , 因此我们得到

$$\begin{aligned}\min_{|\pi|=p} |\Theta(\pi)| &\geq \frac{1}{2} \left( j \min_{|B|=\lceil \frac{n}{p} \rceil} |\Theta(B)| + (p-j) \min_{|B|=\lfloor \frac{n}{p} \rfloor} |\Theta(B)| \right) \\ &= \frac{1}{2} \cdot 2^a \min_{|B|=2^{d-a}} |\Theta(B)| = \frac{1}{2} \cdot 2^a \cdot (d-a) 2^{d-a} \\ &= (d-a) 2^{d-1}\end{aligned}$$

同样不出意料地, 这个下界, 以及与边宽对应的那个下界, 可以在把  $Q_d$  划分为  $(d-a)$  阶子立方时取到。不过令人惊讶的是, 根据 Bezrukov [6] 的观察, 对于  $p$  的其他 (也许是全部) 的值, 这些下界依然是精确的。

**定理 3.**  $\forall d > a$ , 都存在一个均匀  $(2^a + 1)$ -划分, 能把  $Q_d$  划分为立方集。

**练习 15.** 在阅读下面的证明之前, 先证明  $a = 1$  时的特例。

证明. 我们从  $Q_d$  中的一个  $\lfloor \frac{2^d}{2^a+1} \rfloor$ -立方集  $B_1$  开始。因为所有的  $k$ -立方集都在图同构下等价 (见练习 9), 我们不妨令  $B_1 = S_k(\text{lex})$ , 其中  $k = \lfloor \frac{2^d}{2^a+1} \rfloor$ 。因为  $\lfloor \frac{2^d}{2^a+1} \rfloor < 2^{d-a}$ , 所以  $B_1$  是  $Q_d$  的  $(d-a)$  阶子立方的子集, 它的最后  $a$  位坐标均为固定值 0。这些固定坐标的  $2^a$  个值每一个都给出  $B_1$  的一份拷贝, 因此都是立方集。以这些固定坐标的字典序把它们标号为  $B_1, B_2, \dots, B_{2^a}$ 。再把元素  $S_{k+1}(\text{lex}) - S_k(\text{lex})$  加入到  $B_1$ , 把对应的元素加入到  $B_2, \dots, B_j$ , 其中  $j \equiv 2^d \pmod{2^a+1}$ 。那么这些前  $j$  个块就成了  $S_{k+1}(\text{lex})$  的拷贝。我们断言

$$\bigcup_{i=1}^{2^a} B_i$$

是立方集: 它包括一个  $(a_i + a)$  阶子立方的并集, 其中每一个的指数满足  $\sum 2^{a_i} = k$  并且每一个  $(a_i + a)$  阶子立方都处在所有更大的子立方的邻居中。它还包括一个  $b_l$  阶子立方的并集, 其中每一个的指数满足  $\sum 2^{b_l} = j$ ,  $b_1 < b_2 < \dots$ 。因为所有的  $b_l$  阶子立方都处在一个  $a$  阶子立方中, 它处在所有的  $(a_i + a)$  阶子立方的一个邻居中, 所以这个断言得到了证明。因此, 根据练习 8,  $B_{2^a+1} = V_{Q_d} - \bigcup_{i=1}^{2^a} B_i$  是一个大小为  $\lfloor \frac{2^d}{2^a+1} \rfloor$  的立方集, 从而

$$\pi = \{B_1, B_2, \dots, B_{2^a}, B_{2^a+1}\}$$

是所求的划分。 □

**引理 6.** 如果存在一个均匀  $p$ -划分能把  $Q_d$  划分为立方集, 那么也存在一个均匀  $2p$ -划分能把  $Q_{d+1}$  划分为立方集。

证明. 留作练习。 □

Bezrukov [6] 继续用类似的方法证明了对于固定的  $p$ , 当  $d \rightarrow \infty$  时, 存在一个均匀  $p$  划分能把  $Q_d$  划分为渐进最优集。

## 4 评论

一个人总是要不断超越自我，否则还要天堂做什么？(Robert Browning)

当我在写给本章奠定基础的论文 [5] 的时候，我还只是一个初期的研究生。当时的我毫无疑问超越了自我，甚至也许足以让 Browning 也赞同这个说法。很幸运，我并不需要等到上了天堂才有机会弥补我的不足。在我的论文发表两年后，Bernstein 发表了一篇后续论文 [2]，指出我忽视了一种情况（定理 1 论据中的情况 3）并填补了这个空缺。当时 Bernstein 的补丁（引理 3）看似复杂得令人失望，现在，有了将近四十年该教材的使用经验，我了解到 Bernstein 的引理正是当时所需的。它不仅涵盖了遗漏的情况 3，还将情况 1 中含糊不清的论据替换为一个清晰明确的说法，同时还包含着更深刻的见解。

寻找最优均匀 2-划分的问题是 EIP 的一个很吸引人的特例，它在文献中有许多不同的名称，比如“图二分”或者“最小平衡割”。

我们所谓的线长问题出现在许多不同的应用中，因此在文献中也有不同的名称。我们之所以称之为“线长”是因为它既简洁又具有描述性。 $\mathbb{Z}_n$  上线长问题的解由 Lehman [7] 发表于 1963 年。它的应用是显示如何用  $n$  个不同重量但由均匀的弹簧相连的珠子来构造一个手镯，以便最小化基本频率。 $Q_d$  上线长问题的解的原始应用 [5] 是将数据从集合  $\{1, 2, \dots, 2^d\}$  编码到 0 和 1 的  $d$ -元组，以便在一个有噪声（二元对称）但任意位发生错误概率都较低的信道上传输时最小化平均绝对误差。这类线长问题的原始实例是 4 阶德布鲁因图 [8]，它出现在解码电路的布线图中。最小化线长意味着最小化自感。

$Q_d$  上的线长问题，即在本章出现的一个应用，实际上曾是我的工作的出发点。我在喷气推进实验室（JPL）老板 Ed Posner 向我提出了这个问题。然而他所想的应用并不是最小化线长，而是最小化在二元对称的信道上传输线性数据时的平均绝对误差。当时 JPL 的摄像机正在传输第一份月球表面的近拍图像。图像的像素是灰度，有 64 级，从白到黑。它们被编码为 0 和 1 的 6 元组，用以传输到地球，然后再被解码为灰度并重组为图像。问题是由于发送器仅被最低限度供电，在传输过程中一个 0 会偶尔变成一

个 1，或者相反。这最终就会导致一个错误的灰度（取决于编码）并且恶化收到的图像。当时面临的挑战是要证明工程师使用的编码，即字典序编码（编号方式），能最小化平均绝对误差。

作为一名本科生，在解决问题上我深深折服于 G. Polya 的著作 [9]。Polya 的论题是解决问题是可以学习的，尤其是存在有效的策略可以有意识地施加于上。其中之一便是将一个猜想 A 归约为——假定是更简单的——猜想 B 和 C 的结合。我在致力于解决 Posner 的问题上遵循了 Polya 的建议，这就导致产生了第 3.1 小节里的思路。当我看到 Posner 的猜测可归约为字典序编号方式的初始段解决了 EIP 这一猜想，并且这个猜想似乎适用于对  $d$  进行归纳时，我很自信我有了重大发现。自从认识到 Polya 的方法对我自己如此有帮助，我就一直在向年轻的数学家们推荐它。

能够给出  $Q_d$  上 EIP 的所有解集是一个偶然的错误。在那时我以为这对于归纳证明的逻辑来说是必要的。现在回想起来，情况显然并非如此，但它确实导致了一个更强有力的结果，使得这个结果能被更加灵活地应用起来。注意 Bezrukov 在边界划分问题上的应用就利用了这种灵活性，然而这是发生在这篇文章 [5] 发表的 33 年后了。

## 参考文献

- [1] W. H. Kautz, *Optimized data encoding for digital computers*, Convention Record I. R. E., 1954, pp. 47-57
- [2] A. J. Bernstein, *Maximally connected arrays on the  $n$ -cube*, SIAM J. Appl. Math., vol. 15, pp. 1485–1489, 1967
- [3] K. Steiglitz, A. J. Bernstein, *Optimal binary coding of ordered numbers*, J. SIAM, vol. 13, pp. 441-443, 1965
- [4] S. W. Golomb, *Shift Register Sequences*, Aegean Park Press, 1982
- [5] L. H. Harper, *Optimal assignments of numbers to vertices*, J. SIAM, vol. 12, pp. 131-135, 1964

- [6] S. L. Bezrukov, *On  $k$ -partitioning the  $n$ -cube*, Proc. International Conference on Graph Theory Concepts in Computer Science, Como, Italy, 1997
- [7] A. Lehman, *A result on rearrangements*, Israel J. Math, vol. 1, pp. 22-28, 1963
- [8] L. H. Harper, *Chassis layout and isoperimetric problems*, Jet Propulsion Lab. SPS, vol. 11, pp. 37-66, 1970
- [9] G. Polya, *Mathematics and Plausible Reasoning* (2 vols.), Princeton University Press, 1954