arXiv:2303.14539v1 [math.CO] 25 Mar 2023

倍周期序列的无重叠性对应概念

James D. Currie*温尼伯大学数学与统计学系，温尼伯，马尼托巴省 R3B 2E9，加拿大

2023年3月28日

摘要

优词是避开因子 11 和 1001，以及模式 0000 和 00010100 的二进制词。我们证明，优词与倍增词序列的关系，和无重叠词与图厄-莫尔斯序列的关系相同。我们证明了 Fife 定理关于优词的一个对应版本，展示了字典序最小和最大的无限优词，并确定了倍增词所避开的模式。

数学主题分类：68R15

1 引言

著名的图厄-莫尔斯序列 t 是由 $\mu(0) = 01$、 $\mu(1) = 10$ 给出的二元态射μ 的不动点，即

$$\mathbf{t} = \lim_{n\to\infty} \mu^n(0)$$

Thue [13]证明了 t 是无重叠的。

*由加拿大自然科学与工程研究委员会（NSERC）资助， [资助编号为 2017-03901]。

定理 1.1. 设 w 为一个无重叠二进制词。那么 $\mu(w)$ 是无重叠的。

他还证明，对于双向无限词和环形词，每个无重叠二进制词都是某个无重叠词在 μ 下的像 [14]。带有“端点”的词分析起来更复杂，但人们早已知道，有限无重叠二进制词也可以通过迭代 μ得到。（例如参见 Restivo 和 Salemi [11]。）

定理 1.2. 设 $w \in \{0,1\}^*$ 是一个有限无重叠词。那么我们可以写成 $w = a\mu(u)b$，其中 a、$b \in \{\epsilon, 0, 00, 1, 11\}$，且 u 是无重叠的。如果 $|w| \geq 7$，则这种分解是唯一的。如果 $\boldsymbol{w}$ 是一个单向无限无重叠词，那么我们可以写成 $\boldsymbol{w} = a\mu(\boldsymbol{u})$，其中某个单向无限无重叠词 $\boldsymbol{u}$ 满足 $a \in \{\epsilon, 0, 00, 1, 11\}$。

Restivo 和 Salemi 运用他们版本的定理 1.2，对二进制无重叠词进行了粗略计数。Kobayashi [8], 通过统计能够延伸为无限无重叠词的有限词, 给出了更好的计数结果，并得到了一个优良的下界。为此，Kobayashi 运用了 Fife [7] 刻画无限无重叠词的深刻定理。Carpi[4] 和 Cassaigne [5]通过正则语言刻画有限无重叠二进制词的集合，从而推广了 Fife 定理。

由于定理 1.2，t 在二进制无重叠词的研究中频繁出现。Berstel [1] 的以下结果便是一个例子（后来 Allouche 等人 [2]对其进行了大幅推广）：

定理 1.3。以 0 开头的单侧无限无重叠二进制词中字典序最大的词是 t。

t 与 μ 的关系也是确立 Shur [12]以下定理的关键：

定理 1.4。假设 t 遇到了模式 $p \in \{0,1\}^*$。那么，要么 p 是 t 的一个因子，要么 p 是 00100 和 11011 之一。

另一个著名的二元序列是倍周期序列，它是.....的不动点

$$\boldsymbol{d} = 0100010101000100010001010100010 1\cdots$$

二元态射 δ，其中 $\delta(0) = 01$，$\delta(1) = 00$。该序列在准晶体谱理论的上下文中得到了广泛研究。（例如参见 Damanik[6],。）

若二进制词 w 不含因子 11 或 1001，且不出现模式 0000 或 00010100，则称其为优良词。我们将证明 $\boldsymbol{d}$ 是优良的。事实上，我们将证明倍周期态射 δ 与优良词的关系，和 μ 与无重叠词的关系相同，即：

定理 1.5. 设 w 是优良的。则 $\delta(w)$ 是优良的。

定理 1.6. 设 w 是有限良词。则可以写成 $w = a\delta(u)b$，其中 $a \in \{\epsilon, 0, 1\}$、$b \in \{\epsilon, 0\}$ 和 u 是优良的。若 $|w| \geq 4$，则此分解唯一。若 $\boldsymbol{w}$ 是单侧无限良词，则可以写成 $\boldsymbol{w} = a\delta(\boldsymbol{u})$，其中 $\boldsymbol{u}$ 是某个单侧无限良词，且 $a \in \{\epsilon, 0, 1\}$。

我们以这些定理为基础，进而：

- 给出适用于优词的 Fife 定理版本，并刻画无限优词；
- 展示字典序极值的单侧无限优词；
- 刻画被 $\boldsymbol{d}$ 避免的二进制模式。

2 优良词

除非另有说明，我们的字和态射均定义在二进制字母表 $\{0, 1\}$ 上。为方便起见，我们将态射内联记录，即 $g = [g(0)\ g(1)]$。

引理 2.1. 设 u 是一个有限二进制词。假设 $\delta(u)$ 是优良的。那么 u 是优良的。

证明. 若 u 遇到模式 0000 或 00010100，则 $\delta(u)$ 也会遇到。如果 u 有因子 11 或 1001，那么 $\delta(u)$ 有因子 $\delta(11) = 0000$ 或 $\delta(1001) = 00010100$，因此会遇到模式 0000 或 00010100。 □

注记 2.2. 设 $w = w_1w_2w_3\cdots w_{2n}$ 且 $w_i \in \{0, 1\}$。字 w 可以写成 $w = \delta(u)$，其中存在某个 u，当且仅当对每个奇数位置 i，$w_i = 0$。

引理 2.3。设 w 是一个不含因子 11、1001 或 0000 的有限二进制词。那么我们可以写成 $w = a\delta(u)b$，其中 $a \in \{\epsilon, 0, 1\}$ 且 $b \in \{\epsilon, 0\}$.。若 $|w| \geq 4$，则此分解是唯一的。若 $\boldsymbol{w}$ 是一个不含因子 11、1001 或 0000 的单侧无限词，那么我们可以写成 $\boldsymbol{w} = a\delta(\boldsymbol{u})$，其中 $\boldsymbol{u}$ 是某个单侧无限词，且 $a \in \{\epsilon, 0, 1\}$。

证明. 首先，我们证明有限词的这种分解存在：若 $|w|_1 \leq 1$，则 w 是 0001000 的一个因子，结论可通过有限次检查得到。假设 $|w|_1 \geq 2$。若 10^k1 是 w的一个因子，则关于 w的条件迫使 $k = 1$ 或 $k = 3$成立。根据归纳法，若 $1u1$ 是 w 的一个因子，则 $|1u|$ 为偶数；因此，w 中所有 1 的位置都属于 w，且奇偶性相同。若 1 的位置为偶数，则令 $|a| = 0$；若 $|w|$ 为偶数（分别为奇数），则令 $|b|$ 为 0（分别为 1）。若 1 的位置为奇数，则令 $|a| = 1$；若 $|w|$ 为偶数（分别为奇数），则令 $|b|$ 为 1（分别为 0）。根据注记 2.2，我们可以将 $a^{-1}wb^{-1} = \delta$写成 $u(u)$。

我们已经证明，可以写成 $w = a\delta(u)b$，且 a、$b \in \{\epsilon, 0, 1\}$。假设 $b = 1$。回想一下，$|w|_1 \geq 2$。a 和 b 的位置奇偶性不同，因此不可能有 $a = 1$。由此可得 $|\delta(u)|_1 \geq 1$。那么 $\delta(u)1$ 的后缀必为 011、01001 或 00001，而这些都不是优良的。这就产生了矛盾，因此事实上 $b \in \{\epsilon, 1\}$。

现在我们证明有限词分解的唯一性：设$|w| \geq 4$ 和 w 有两种分解 $w = a_1\delta(u_1)b_1 = a_2\delta(u_2)b_2$。如果 $|a_1| = |a_2|$，我们就必须选择 $a_1 = a_2$、$u_1 = u_2$和 $b_1 = b_2$，因此这两种分解相同。不失一般性，设 $|a_1| = 0$、$|a_2| = 1$。那么根据引理 2.2，$a_1\delta(u_1)b_1$ 中的任何 1 的位置都是偶数，但$a_2\delta(u_2)b_2$ 中的任何 1 的位置也是偶数。由于 $a_1\delta(u_1)b_1 = w = a_2\delta(u_2)b_2$，我们得出 $|w|_1 = 0$，因此 w 以 0000 为前缀，这不可能。

现在假设 $\boldsymbol{w}$ 是一个单侧无限良词。对于每个非负的 n，令 $\boldsymbol{w}$ 的长度为 n 的前缀为 w_n。我们已证明有限词可以分解，因此将每个 $w_n = a_n\delta(u_n)b_n$ 写成.....，其中 a_n、$b_n \in \{\epsilon, 0, 1\}$. 这个字不能是 0000，因此 $|w_4|_1 > 0$。根据注记 2.2，w_4 中第一个 1 的位置决定了 a_4 以及所有后续的 a_n，因此对于 $n \geq 4$，有 $a_n = a_4$。这意味着，当 $n \geq 4$ 时，$\delta(u_n)$ 是 $\delta(u_{n+1})$ 的前缀，因此 u_{n+4} 是一个

对于所有 n，.....是 u_{n+5} 的前缀。令 $\boldsymbol{u} = = \lim_{n\to\infty} u_{n+4}$。那么

$$\begin{aligned}
\boldsymbol{w} &= \lim_{n\to\infty} w_n \\
&= \lim_{n\to\infty} w_{n+4} \\
&= \lim_{n\to\infty} w_{n+4} b_{n+4}^{-1} \\
&= \lim_{n\to\infty} a_4 \delta(u_{n+4}) \\
&= a_4 \lim_{n\to\infty} \delta(u_{n+4}) \\
&= a_4 \delta(\lim_{n\to\infty} u_{n+4}) \\
&= a_4 \delta(\boldsymbol{u}).
\end{aligned}$$

□

定理 1.6 的证明：这可直接由引理 2.1 和引理 2.3 得出。 □

称非擦除态射 g 为偶态射，若对每个字母 u，$|g(u)|$都是偶数。

引理 2.4. 设 p 是一个模式，w 是一个二进制词。假设$g(p)$是 $\delta(w)$的一个因子，其中 g 是一个偶态射。那么 w 包含模式 p。

证明。写作 $p = u_1 u_2 \cdots u_n$，其中 u_i是字母。写作 $\delta(w) = aU_1U_2\cdots U_n b$，其中对每个 i，$U_i = g(u_i)$。若 $|a|$ 是偶数，则 $w = \delta^{-1}(a)\delta^{-1}(U_1U_2\cdots U_n)\delta^{-1}(b)$，且 w 包含 p的实例 $\delta^{-1}(g(p))$。

若 $|a|$ 是奇数，则 0 是 a 以及每个 U_i的最后一个字母。因此 $\delta(w) = a0^{-1}0U_10^{-1}0U_20^{-1}0\cdots U_n0^{-1}0b$，且 w 包含 p的实例 $\delta^{-1}(h(p))$，其中 h 是定义在 p 的字母上的态射，满足 $h(u_i) = 0g(u_i)0^{-1}$。 □

引理 2.5. 设 w、u、v 是二进制词，并假设 uvu 是 $\delta(w)$ 的一个因子。若 $|u|_1 > 0$，则 $|uv|$ 是偶数。

证明. 由于 uvu 的周期为 $|uv|$，$\delta(w)$ 有一个因子 $1z1$，其中 $|1z| = |uv|$。由注记 2.2 可得此引理。 □

定理 1.5 的证明：首先，我们证明 $\delta(w)$ 不包含 11、0000、1001 或 00010100 中的任何一个作为因子。由注记 2.2，$\delta(w)$ 不含因子 11 或 1001。若 $\delta(w)$ 有因子 0000，写作 $\delta(w) = a0000b$，其中 a 和 b 是字。若 $|a|$ 是偶数，则 w 以前缀 $\delta^{-1}(a0000)$ 开头，而该前缀以

11 结尾。这是不可能的；若 $|a|$ 是奇数，则 a 的最后一个字母是 0，因此 w 以前缀 $\delta^{-1}((a0^{-1})0000)$ 开头，这同样以 11 结尾。最后，若 00010100 是 $\delta(w)$ 的一个因子，则 w 含有因子 $\delta^{-1}(00010100) = 1001$，这是不可能的。（根据注记 2.2，00010100 在 $\delta(w)$ 中的位置必为奇数）

现在假设 $\delta(w)$ 遇到模式 $p = 0000$，也就是说，对于某个非空的 X，$XXXX$ 是 $\delta(w)$ 的一个因子。由于 $\delta(w)$ 不以 0000 为因子，因此必有 $|X|_1 > 0$。利用引理 2.5 中的 $u = X$ 和 $v = \epsilon$，我们得出 $|X|$ 是偶数。于是，根据引理 2.4，w 会遇到 0000，这就产生了矛盾。

假设 $\delta(w)$ 遇到模式 $p = 00010100$，也就是说，对于某些非空的 X 和 Y，$XXXYXYXX$ 是 $\delta(w)$ 的一个因子。假设 $|X|_1 > 0$。由于 XX 是 $\delta(w)$ 的一个因子，在引理 2.5 中取 $u = X$ 和 $v = \epsilon$，可得 $|X|$ 是偶数。同样，由于 XYX 是 $\delta(w)$ 的一个因子，在引理 2.5 中取 $u = X$ 和$v = Y$，可得 $|XY|$ 是偶数。因此 $|Y|$ 是偶数。于是，根据引理 2.4，w 会遇到 00010100，这就产生了矛盾。因此我们得出 $|X|_1 = 0$，也就是说，对于某个 $n \geq 1$，有 $X = 0^n$。由于 $XX = 0^{2n}$ 是 $\delta(w)$ 的一个因子，但 0000 不是，因此我们得出 $n = 1$ 和 $X = 0$。于是 $\delta(w)$ 含有因子 $XXXYXYXX = 000Y0Y00$。

由于 0000 不是 $\delta(w)$ 的因子，Y的首字母是 1。若 $Y = 1$，则$\delta(w)$ 含有因子 00010100，这是不可能的。因此 $|Y| > 1$.

假设 Y 以 0 结尾。若 Y 以 10 结尾，则 $\delta(w)$ 含有因子 1001（在 $Y0Y$ 中），这是不可能的。若 $|Y|$ 以 00 结尾，则 $000Y0Y00$ 以 $Y00$ 结尾，因此会出现 0000，这同样不可能。因此 Y 以 1 结尾，进而以 01 结尾。令 $Y = 1Z1$ 为某个非空字 Z。字 $\delta(w)$ 含有因子$0001Z101Z100$。

若 $|1Z1| = 3$，则 $\delta(w)$ 含有 $01Z101Z1 = 01010101$，它是 0000 的一个实例，而 0000 已证明不可能出现。由此可得 $|1Z1| \geq 4$。若 Z 以 01 开头，则 $Z10$ 以 0101 或 0100 开头。然而，若 $Z10$ 以 0101 开头，那么由于 Z 以 0 结尾，$Z101Z10$ 再次含有 01010101；若 $Z10$ 以 0100 开头，则 $0001Z10$ 以 00010100 开头，而这不是 $\delta(w)$ 的因子。我们得出结论，Z 并非以 01 开头，因此以 00 开头。

如果 $Z10$以 001 开头，那么 1Z10 中就有不可能出现的因子 1001。因此，$Z10$以 0001 开头。现在我们来考察 Z.的后缀。如果 Z 以 10 结尾，那么 $01Z$ 的结尾只能是 1010 或 0010。然而，如果 $01Z$ 以 1010 结尾，那么由于 Z 以 0 开头，$01Z101Z$ 就包含 10101010，这是 0000 的一个实例，不可能出现；如果$01Z$ 以 0010 结尾，那么 $101Z100$就以 00010100 结尾，而这不是 δ（w）的因子。由此可知，Z 不以 10 结尾，因此以 00 结尾。如果 $01Z$ 以 0100 结尾，那么 δ（w）包含 $01Z1$，于是就会出现不可能的因子 1001。因此，$01Z$ 以

1000 结尾。由于 Z 以 00 开头，这就意味着 $01Z101Z$ 中会出现 00010100，这是不可能的。 □

推论 2.6。周期倍增词 $\boldsymbol{d}$ 是优良的。

定理 2.7。设 w 是一个单侧无限二进制词。那么， $\boldsymbol{w}$ 是优良的，当且仅当 $\delta(\boldsymbol{w})$ 是优良的。

证明。若 w_n 是 w 的长度为 n 的前缀，则 $\mathbf{w} = \lim_{n\to\infty} w_n$，且 $\delta(\mathbf{w}) = \lim_{n\to\infty} \delta(w_n)$。由定理 1.5 和引理 2.1， w_n 是优良的，当且仅当$\delta(w_n)$ 是优良的。

$$\begin{aligned} \mathbf{w} \text{ is good} &\iff \text{each } w_n \text{ is good} \\ &\iff \text{each } \delta(w_n) \text{ is good} \\ &\iff \lim_{n\to\infty} \delta(w_n) = \delta(\lim_{n\to\infty} w_n) = \delta(\mathbf{w}) \text{ is good.} \end{aligned}$$

□

注 2.8。虽然二进制无重叠词的集合在取补和反转下封闭，但优词并非如此。例如，00101000 是优良的，但它的补和反转都不是优良的。

3 Fife 定理的类比

我们刻画单侧无限良词，并发展一套与 Fife 针对无重叠二进制词提出的理论相类比的理论；不过，我们采用 Ramersad 对 Fife 理论的阐述，并作适当修改，而非遵循 Fife 的原始论文。设 G 为单侧无限良词的集合。

引理 3.1. 设 $\boldsymbol{u}$ 是一个单侧无限良词。当且仅当 $01\boldsymbol{u}$ 优良时， $1\boldsymbol{u}$ 也优良。

证明。显然，如果 $01\boldsymbol{u}$ 是优词，那么 $1\boldsymbol{u}$ 是优词。假设 $1\boldsymbol{u}$ 是优词，但 $01\boldsymbol{u}$ 不是。根据定理 1.6，写作 $1\boldsymbol{u} = 1\delta(\boldsymbol{v})$，其中 $\boldsymbol{v}$ 满足某些条件。由于$1\boldsymbol{u}$ 是优词， $01\boldsymbol{u}$ 必须有一个前缀，它要么是因子 11 或 1001，要么是模式实例 g(0000) 或 g (00010100)，其中 $g = [X, Y]$ 是某个非擦除态射。显然，11 和 1001 都不可能是 $01\boldsymbol{u}$ 的前缀，因此 $01\boldsymbol{u}$ 有一个前缀 $XXXX$ 或 $XXXYXYXX$，其中 X 和 Y 均非空。

假设 $01\boldsymbol{u}$ 有前缀 $XXXX$。记作 $X = 0X'$。由于 $|XXXX|$是偶数，字 $01\boldsymbol{u} = \delta(0\boldsymbol{v})$ 以 $XXXX0 = 0X'0X'0X'0X'0$ 为前缀。但此时优词 $1\boldsymbol{u}$ 包含四次幂 $X'0X'0X'0X'0$，这就产生了矛盾。

现在假设 $01\boldsymbol{u}$ 以 $XXXYXYXX$ 为前缀。记 $X = 0X'$。若 $X' = \epsilon$，则 $X = 0$。然而，此时 $01\boldsymbol{u}$ 的长度为 2 的前缀是 $XX = 00$，这是不可能的。因此 $X' \neq \epsilon$，从而 01 是 X 的前缀，于是$|X|_1 > 0$。由于 $01\boldsymbol{u} = \delta(0\boldsymbol{v})$，考察 $XXXYXYXX$ 中的第二个 XX，由引理 2.5，取 $u = X$、$v = \epsilon$，可知 $|X|$ 是偶数。同样，$\delta(0\boldsymbol{v})$ 的一个因子是 XYX，因此应用引理 2.5，取 $u = X$、$v = Y$，可知 $|XY|$ 也是偶数。由于 $|X|$ 和 $|XY|$ 都是偶数，$|Y|$ 是偶数。由于 $|XXX|$ 是偶数，$XXX0$ 是 $01\boldsymbol{u} = \delta(0\boldsymbol{v})$ 的前缀，因此可记为 $Y = 0Y'$。由于 $|XXXYXYXX|$ 是偶数，字 $01\boldsymbol{u} = \delta(0\boldsymbol{v})$ 以 $XXXYXYXX0 = 0X'0X'0X'0Y'0X'0Y'0X'0X'0$ 为前缀。但此时，优词 $1\boldsymbol{u}$ 包含 $X'0X'0X'0Y'0X'0Y'0X'0X'0 = g'$（00010100），其中 $g' = [X'0, Y'0]$。这就产生了矛盾。 □

注 3.2. 这个结果利用了 $\boldsymbol{u}$ 是单侧无限的这一事实。如果 u =010101，那么 $1u$ 是优良的，但 $01u = (01)^4$ 不是。

令 G 为单侧无限良词的集合。对于 $w \in \{0,1\}^*$，令$G_w = G \cap w\{0,1\}^\omega$。

引理 3.3. 设 $\boldsymbol{w}$ 是一个单边无限二进制词。

(a) $\delta(\boldsymbol{w}) \in G \iff \boldsymbol{w} \in G$(b) $1\delta(\boldsymbol{w}) \in G \iff 0\boldsymbol{w} \in G$(c) $0\delta(\boldsymbol{w}) \in G \iff (1\boldsymbol{w} \in G)$ or $(\boldsymbol{w} \in G_{001})$

注 3.4. (c) 中的情形彼此不相交，因为若 001 是 w 的前缀，则 1w 的前缀为 1001，因此不是优词。

(a) 的证明：这正是定理 2.7。

(b) 的证明：如果 $0\boldsymbol{w} \in G$，那么根据定理 2.7，$\delta(0\boldsymbol{w}) = 01\delta(\boldsymbol{w}) \in G$，因此特别有 $1\delta(\boldsymbol{w}) \in G$。

反过来，假设 $1\delta(\boldsymbol{w}) \in G$。将引理 3.1 应用于前缀，得到 $01\delta(\boldsymbol{w}) = \delta(0\boldsymbol{w}) \in G$。根据定理 2.7，$0\boldsymbol{w} \in G$。

（c）的证明：若 $1\boldsymbol{w} \in G$，则由定理 2.7 $\delta(1\boldsymbol{w})$，有 $= 00\delta(\boldsymbol{w}) \in G$，因此特别有 $0\delta(\boldsymbol{w}) \in G$。

假设 $\boldsymbol{w} \in G_{001}$。由定理 2.7, $\delta(\boldsymbol{w}) \in G$。假设 $0\delta(\boldsymbol{w}) \notin G$。因此它有一个前缀 $XXXX$ 或 $XXXYXYXX$，其中 X 非空。由于 001 是 w 的前缀，$p = 0010100$是 0δ(w) 的前缀。0δ(w) 的前缀XXX 的周期为 $|X|$。若 $|X| < |p|$，则 $|X|$ 是 p的周期，而其最小周期为 5。由此得出 $|X| \geq 5$。这意味着 $|X|_1 > 0$。XXX 中的第二个 XX 是 δ(w) 的一个因子。由引理 2.5，取 $u = X$,$v = \epsilon$，可得 $|X|$ 为偶数。因此 $|X| \geq 6$。由于 $0^{-1}X$ 是 $\delta(\boldsymbol{w})$ 的奇数长度前缀，0 是 X的后缀。

若 $|X| = 6$，则 δ(w) 的前缀 $0^{-1}XXX$ 以 $0XX = 0\ 001010$001010 结尾，其前缀为 00010100。这不可能，因为 δ(w) $\in G$。若$|X| > 6$，则 $p = 0010100$ 是 X 的前缀，而 0 是 X 的后缀，因此 δ(w) 的因子 XX 包含 0 0010100，这同样不可能。

反过来，假设 0δ(w) $\in G$。那么根据定理 2.7, $\boldsymbol{w} \in G$，我们将证明要么 $1\boldsymbol{w} \in G$, 要么 $\boldsymbol{w} \in 001(0, 1)^*$。假设$1\boldsymbol{w} \notin G$，且 001 不是 w 的前缀。由此可知，$1\boldsymbol{w}$ 的一个前缀形式为 11、1001、$XXXX$ 或 $XXXYXYXX$，其中 X 和/或 Y 非空。

根据定理 1.6，写作 $\boldsymbol{w} = a\delta(\boldsymbol{u})$，其中 $a \in \{\epsilon, 0, 1\}$。

如果 11 是 $1\boldsymbol{w}$ 的前缀，那么 $a = 1$，且 10 是 $\boldsymbol{w}$ 的前缀。然而，此时 $0\delta(\boldsymbol{w})$ 的前缀为 $0\delta(10) = 00001$，因此并非优良。

假设 1001 是 $1\boldsymbol{w}$ 的前缀，那么 001 是 w 的前缀，这就导致矛盾。

假设 $XXXX$ 是 $1\boldsymbol{w}$ 的非空前缀。那么 $XXXX$ 的最后 XX 是 $\delta(\boldsymbol{u})$ 和 $|X|_1 > 0$ 的因子，因此根据引理 2.5, $|X|$ 是偶数。由于 X 是 $1\boldsymbol{w}$ 的前缀，X 的第一个字母是 1。由于 XX 是 $\boldsymbol{w}$ 的因子，且 $\boldsymbol{w}$ 不含因子 11，X 的最后一个字母是 0。写作 $X = 1X'0$。

字 $\boldsymbol{w}$ 有前缀 $X'01X'01X'01X'0c$，其中某个 $c \in \{0, 1\}$。这意味着 $0\delta(\boldsymbol{w})$ 有前缀

$$\begin{aligned} 0\delta(X'01X'01X'01X'0c) &= 0\delta(X'0)\delta(1)\delta(X'0)\delta(1)\delta(X'0)\delta(1)\delta(X'0)\delta(c) \\ &= 0\delta(X'0)00\delta(X'0)00\delta(X'0)00\delta(X'0)0\overline{c} \\ &= (0\delta(X'0)0)^4\overline{c} \end{aligned}$$

而 $0\delta(\boldsymbol{w})$ 含有一个四次幂。这与矛盾相悖，因为 $0\delta(\boldsymbol{w}) \in G$。

假设 $XXXYXYXX$ 是 $1\boldsymbol{w}$的前缀。这意味着 XX 和XYX 是 $\delta(\boldsymbol{u})$的因子，且 $|X|_1 > 0$。由引理 2.5 可知，$|X|$ 和 $|Y|$ 都是偶数。X 的首字母是 1。由于 XX 和 YX 是 $\boldsymbol{w}$的因子，且 $\boldsymbol{w}$ 没有因子 11，因此 X 和 Y 各自的末字母是

令 $X = 1X'0$ 和 $Y = cY'0$，其中 $c \in \{0,1\}$。由于 $X'01$ 是 $a\delta(\boldsymbol{u})$的偶数长度前缀，注记 2.2 可推出 $a = \epsilon$。

那么 $\boldsymbol{w}$ 以 $X'01X'01X'01Y'01X'01y'01X'01X'0d$ 为前缀，其中 $d \in\{0,1\}$。因此 $0\delta(\boldsymbol{w})$ 以

$$0\delta(X'01X'01X'0cY'01X'0cY'01X'01X'0d)= \ 0\delta(X')0100\delta(X')0100\delta(X')010\overline{c}\delta(Y')0100\delta(X')010\overline{c}\delta(Y')0100\delta(X')0100\delta(X')010\overline{d}= \ h(00010100)\overline{d}$$

其中 $h = [0\delta(X')010, \ \overline{c}\delta(Y')010]$。这是一个矛盾，因为 $0\delta(\boldsymbol{w}) \in G$。

□

假设 $w \in \{0,1\}^*$ 有后缀 $\delta^n(01)$，$n \geq 0$，并令 n 尽可能大。写作 $w = y\delta^n(01)$。在 w 上定义映射 α、β 和 γ，如下所示

$$\begin{aligned}\alpha(w) &= w\delta^n(00) = y\delta^{n+1}(01)\\ \beta(w) &= w\delta^n(0100) = y\delta^{n+1}(001)\\ \gamma(w) &= w\delta^n(010100) = y\delta^{n+1}(0001)\end{aligned}$$

例如，若 $w = 0001000101$，则 $y = 00$, $n = 2$, $\delta^n(0) = 0100$,$\delta^n(1) = 0101$，因此

$$\begin{aligned}\alpha(w) &= 00\ 0100\ 0101\ 0100\ 0100\\ \beta(w) &= 00\ 0100\ 0101\ 0100\ 0101\ 0100\ 0100\\ \gamma(w) &= 00\ 0100\ 0101\ 0100\ 0101\ 0100\ 0101\ 0100\ 0100\end{aligned}$$

设 $\boldsymbol{f} = f_1f_2f_3\cdots$，其中每个 $f_i \in B = \{\alpha,\beta,\gamma\}$。假设 w 有某个后缀 $\delta^n(01)$。那么 w 是 $f_1(w)$ 的真前缀，而后者是 $f_2(f_1(w))$ 的真前缀，后者又是 $f_3(f_2(f_1(w)))$ 的真前缀，依此类推。我们将无限复合 $\boldsymbol{x} = (\cdots \circ f_3 \circ f_2 \circ f_1)(01)$ 定义为单侧无限词

$$\boldsymbol{x} = \lim_{n\to\infty} f_n(f_{n-1}(\cdots f_2(f_1(w))\cdots)$$

它以每个 $f_n(f_{n-1}(\cdots f_2(f_1(w))\ \cdots)$ 为前缀。沿用 Rampersad 的记法，我们使用记号 $\boldsymbol{x} = w \bullet \boldsymbol{f}$。

令 $w \in B^k$。则

$$01 \bullet w\boldsymbol{f} = (01 \bullet w)\delta^k(01)^{-1}\delta^k(\boldsymbol{x}) \tag{1}$$

定义集合 I 和 F，使得

$$I = (\beta+\gamma)(\alpha\alpha)^*\alpha(\gamma+\beta\beta) \cup \gamma(\alpha\beta)^*\alpha\gamma$$

$$F = B^\omega - B^* I B^\omega$$

定理 3.5. 设 $\boldsymbol{x} \in \{0,1\}^\omega$。若 $\boldsymbol{x}$ 以 01 开头，则 $\boldsymbol{x}$ 是优良的，当且仅当对于某个 $\boldsymbol{f} \in F$，$\boldsymbol{x} = 01 \bullet \boldsymbol{f}$。

我们沿用 Berstel 的记号，Rampersad 也使用了这一记号。这里 I 表示“理想”，而 $B^* I B^\omega$ 是由 I生成的理想，由 F的禁因子组成。

令 $W = \{\boldsymbol{f} \in B^\omega : \ 01 \bullet \boldsymbol{f} \in G\}$. 要证明定理 3.5，只需证明 $W = F$。令 $L \subseteq \Sigma^\omega$，并令 $x \in \Sigma^*$。我们将（左）商 $x^{-1}L$ 定义为 $x^{-1}L = \{\boldsymbol{y} \in \Sigma^\omega : x\boldsymbol{y} \in L\}$。下面的引理给出了集合 W的商所满足的若干恒等式。利用（1）和引理 3.3 可以证明这些恒等式。它们表明，W 恰好是图 3 中给出的自动机 A_{01} 上的无限标记路径的集合。这些路径正是省略了 I中的因子的标记路径，因此 $W = F$。由此，证明引理 3.11 即可证明定理 3.5。

引理 3.6。以下恒等式成立：

(a) $W = \alpha^{-1}W$;

(b) $(\beta\alpha\alpha)^{-1}W = (\beta\beta)^{-1}W = \beta^{-1}W$

(c) $(\beta\alpha\beta)^{-1}W = \gamma^{-1}W = (\beta\gamma)^{-1}W = (\gamma\gamma)^{-1}W$

(d) $(\beta\alpha)^{-1}W = (\gamma\alpha)^{-1}W$

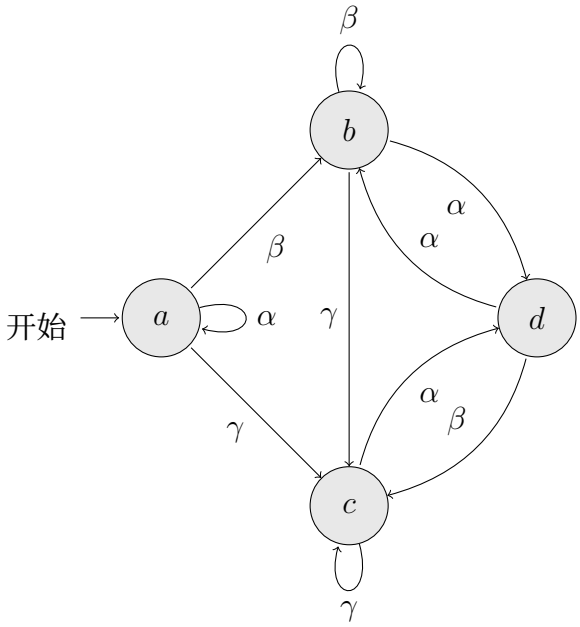
(e) $(\beta\alpha\gamma)^{-1}W = (\gamma\beta)^{-1}W = \emptyset$

每组恒等式对应于与这些恒等式标签相同的 A_{01} 状态。图中未显示非接受汇点 (e)。

证明。设 $01 \bullet \boldsymbol{f} = \boldsymbol{x}$ 是优良的。
(a)

$$\alpha\boldsymbol{f} \in W \iff 01 \bullet \alpha\boldsymbol{f} \in G \iff (01 \bullet \alpha)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \iff$$

图 1：“Fife”自动机 A_{01} 用于 G_{01}



~~0100(0100)~~$^{-1}\delta(\boldsymbol{x})$

$\in G \iff \boldsymbol{x} \in G \iff 01 \bullet \boldsymbol{x} \in G \iff \boldsymbol{f} \in W$使得 $\alpha^{-1}W = W$。

(b)

$$\beta \boldsymbol{f} \in W \iff 01 \bullet \beta \boldsymbol{f} \in G \iff (01 \bullet \beta)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \iff$$

$$01\ \text{~~0100(0100)~~}^{-1}\delta(\boldsymbol{x}) \in G \iff \delta(0\boldsymbol{x}) \in G \iff 0\boldsymbol{x} \in G$$

$$\beta\beta \boldsymbol{f} \in W \iff 01 \bullet \beta\beta \boldsymbol{f} \in G \iff (01 \bullet \beta\beta)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \iff$$

$$010100\ \text{~~01000101(01000101)~~}^{-1}\delta^2(\boldsymbol{x}) \in G \iff \delta(0\delta(0\boldsymbol{x})) \in G \iff$$

$$\delta(0\delta(0\boldsymbol{x})) \in G \iff 0\delta(0\boldsymbol{x}) \in G \iff$$

$$10\boldsymbol{x} \in G \text{ or } 0\boldsymbol{x} \in G_{001} \iff 0\boldsymbol{x} \in G$$

这里我们利用 01 是 $\boldsymbol{x}$ 的前缀这一事实，因此 $10\boldsymbol{x} \notin G$ 和 $0\boldsymbol{x} \in$001(0, 1)$^{\omega}$。

$\beta\alpha\alpha\boldsymbol{f} \in W \iff 01 \bullet \beta\alpha\alpha\boldsymbol{f} \in G \iff (01 \bullet \beta\alpha\alpha)\delta^3(01)^{-1}\delta^3(\boldsymbol{x}) \in G \iff 01$
$0100010101000100(0100010101000100)^{-1}\delta^3(\boldsymbol{x}) \in G \iff \delta(0\delta^2(\boldsymbol{x})) \in G$

$$\iff 10\delta^2(\boldsymbol{x}) \in G \text{ or } 0\delta^2(\boldsymbol{x}) \in G_{001} \iff 0\delta^2(\boldsymbol{x}) \in G$$
$$\iff 1\delta(\boldsymbol{x}) \in G \text{ or } \delta(\boldsymbol{x}) \in G_{001} \iff 1\delta(\boldsymbol{x}) \iff 0\boldsymbol{x} \in G$$

这里再次注意到，01 是 $\delta(\boldsymbol{x})$ 和 $\delta^2(\boldsymbol{x})$ 的前缀。我们已经证明了 $(\beta\alpha\alpha)^{-1}W = (\beta\beta)^{-1}W = \beta^{-1}W$，如所需。

(c)

$$\gamma\boldsymbol{f} \in W \iff 01 \bullet \gamma\boldsymbol{f} \in G \iff (01 \bullet \gamma)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \iff$$
$$0101 \ \cancel{0100(0100)^{-1}}\delta(\boldsymbol{x}) \in G \iff \delta(00\boldsymbol{x}) \in G \iff 00\boldsymbol{x} \in G$$

$\beta\alpha\beta\boldsymbol{f} \in W \iff 01 \bullet \beta\alpha\beta\boldsymbol{f} \in G \iff (01 \bullet \beta\alpha\beta)\delta^3(01)^{-1}\delta^3(\boldsymbol{x}) \in G \iff$
$0101000101\ 0100010101000100(0100010101000100)^{-1}\delta^3(\boldsymbol{x}) \in G \iff$
$\delta(0\delta^2(0\boldsymbol{x})) \in G \iff \iff 0\delta^2(0\boldsymbol{x}) \in G \iff 1\delta(0\boldsymbol{x}) \in G$ 或 $\delta(0\boldsymbol{x}) \in G_{001}$
$\iff 1\delta(0\boldsymbol{x}) \in G \iff \delta(00\boldsymbol{x}) \in G \iff 00\boldsymbol{x} \in G$

$$\beta\gamma\boldsymbol{f} \in W \iff 01 \bullet \beta\gamma\boldsymbol{f} \in G \iff (01 \bullet \beta\gamma)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \iff$$
$$0101000100 \ \cancel{01000101(01000101)^{-1}}\delta^2(\boldsymbol{x}) \in G \iff \delta(0\delta(00\boldsymbol{x})) \in G \iff$$
$$0\delta(00\boldsymbol{x}) \in G \iff 100\boldsymbol{x} \in G \text{ or } 00\boldsymbol{x} \in G_{001} \iff 100\boldsymbol{x} \in G$$

$$\gamma\gamma\boldsymbol{f} \in W \iff 01 \bullet \gamma\gamma\boldsymbol{f} \in G \iff (01 \bullet \gamma\gamma)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G$$
$$\iff 010101000100 \ \cancel{01000101(01000101)^{-1}}\delta^2(\boldsymbol{x}) \in G$$
$$\iff \delta^2(100\boldsymbol{x}) \in G \iff 100\boldsymbol{x} \in G$$

最后，假设 $00\boldsymbol{x} \in G$。那么 0001 是 $00\boldsymbol{x}$ 的前缀，根据定理 1.6，我们可以写成 $00\boldsymbol{x} = \delta(10\boldsymbol{y})$，其中 $10\boldsymbol{y} \in G$。于是由引理 3.1，

$010\boldsymbol{y} \in G$。因此 $\delta(010\boldsymbol{y}) = 0100\boldsymbol{x} \in G$，特别是 $100\boldsymbol{x} \in G$。由此我们得出 $100\boldsymbol{x} \in G \iff 00\boldsymbol{x} \in G$，这就得到了所需结果。

(d)

$$\beta\alpha\boldsymbol{f} \in W \iff 01 \bullet \beta\alpha\boldsymbol{f} \in G \iff (01 \bullet \beta\alpha)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \iff$$

$$01\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \iff \delta(0\delta(\boldsymbol{x})) \in G \iff$$

$$0\delta(\boldsymbol{x}) \in G \iff 1\boldsymbol{x} \in G \text{ or } \boldsymbol{x} \in G_{001} \iff 1\boldsymbol{x} \in G$$

$$\gamma\alpha\boldsymbol{f} \in W \iff 01 \bullet \gamma\alpha\boldsymbol{f} \in G \iff (01 \bullet \gamma\alpha)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \iff$$

$$0101\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \iff \delta(00\delta(\boldsymbol{x})) \in G$$

$$\iff \delta^2(1\boldsymbol{x}) \in G \iff 1\boldsymbol{x} \in G$$

因此 $(\beta\alpha)^{-1}W = (\gamma\alpha)^{-1}W$。

(e)

$$\beta\alpha\gamma\boldsymbol{f} \in W \iff 01 \bullet \beta\alpha\gamma\boldsymbol{f} \in G \iff (01 \bullet \beta\alpha\gamma)\ \delta^3(01)^{-1}\delta^3(\boldsymbol{x}) \in G \iff$$

$$01010001010100010\cancel{1\ 0100010101000100(0100010101000100)}^{-1}\delta^3(\boldsymbol{x}) \in G$$

$$\iff \delta(0\delta^2(00\boldsymbol{x})) \in G \iff 0\delta^2(00\boldsymbol{x}) \in G \iff$$

$1\delta(00\boldsymbol{x}) \in G$ 或 $\delta(00\boldsymbol{x}) \in G_{001} \iff 1\delta(00\boldsymbol{x}) \in G \iff 000\boldsymbol{x} \in G$但是 $000\boldsymbol{x}$ 的前缀是0000。因此 $(\beta\alpha\gamma)^{-1}W = \emptyset$。

$$\gamma\beta\boldsymbol{f} \in W \iff 01 \bullet \gamma\beta\boldsymbol{f} \in G \iff (01 \bullet \gamma\beta)\ \delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \iff$$

$$\underline{01010100\ 01000101}(01000101)^{-1}\delta^2(\boldsymbol{x}) \in G \iff \delta^2(10\boldsymbol{x}) \in G \iff 10\boldsymbol{x} \in G$$

但是 $10\boldsymbol{x}$ 的前缀是1001，所以 $(\gamma\beta)^{-1}W = \emptyset$。

□

我们有

$$G = G_{01} \cup G_{001} \cup G_{0001} \cup G_1.$$

我们可以将定理 3.5 写作

$$G_{01} = 01 \bullet W.$$

根据引理 3.1， $\boldsymbol{x} \in G_1$ 当且仅当 $0\boldsymbol{x} \in G_{01}$，因此

$$G_1 = 0^{-1}G_{01}.$$

此外，由定理 1.6 和定理 1.5,

$$G_{0001} = \delta(G_1)$$

因为 $G_{10} = G_{1}$。若 $\boldsymbol{x} \in G_{001}$，则可将 $\boldsymbol{x} = 0\delta(0\boldsymbol{y})$ 写成某个 $\boldsymbol{y}$。然而，由引理 3.3,

$$\begin{aligned}
& 0\delta(0\boldsymbol{y}) \in G \\
\iff\ & 10\boldsymbol{y} \in G \text{ or } 0\boldsymbol{y} \in G_{001} \\
\iff\ & 010\boldsymbol{y} \in G \text{ or } 0\boldsymbol{y} \in G_{001} \text{ by Lemma 3.1} \\
\iff\ & 010\boldsymbol{y} \in G_{01} \text{ or } 0\boldsymbol{y} \in G_{001} \\
\iff\ & 0\boldsymbol{y} \in (01)^{-1}G_{01} \text{ or } 0\boldsymbol{y} \in G_{001} \\
\iff\ & 0\delta(0\boldsymbol{y}) \in 0\delta((01)^{-1}G_{01} \cup \in G_{001}), \text{ so that}
\end{aligned}$$

$$G_{001} = 0\delta((01)^{-1}G_{01} \cup \in G_{001}).$$

综上,

定理 3.7。

$$\begin{aligned}
G_{01} &= 01 \bullet W \\
G_{001} &= 0\delta((01)^{-1}G_{01} \cup \in G_{001}) \\
G_{0001} &= \delta(G_1) \\
G_1 &= 0^{-1}G_{01}
\end{aligned}$$

由于 I 是一个正则语言，可以按照 Kobayashi[?] 的方法，枚举 G_{01} 的有限前缀（从而也能枚举 G_{0001}、 G_1）。这将给出长度为 n 的优词数量的一个下界。若要枚举 G 的所有前缀，则需要处理 G_{001} 方程中的递归，并能为长度为 n 的优词数量给出更好的下界。

若能消去 G_{001} 方程中的递归就好了。对于重叠，Rampersad 对 Fife 定理的表述具有如下形式（无递归），其中 $\hat{\mathcal{I}}$, $\mathcal{F}$, $\hat{\alpha}$, β, $\hat{\gamma}$, $\hat{\bullet}$ 等分别对应于 I, F, α, β, γ, $\bullet$ 等:

定理 3.8。设 $\boldsymbol{x} \in \{0,1\}^\omega$。

1. 如果 $\boldsymbol{x}$ 以 01 开头，那么当且仅当对于某个 $\boldsymbol{f} \in \mathcal{F}$， $\boldsymbol{x} = 01\hat{}\bullet \boldsymbol{f}$， $\boldsymbol{x}$ 才是无重叠的。2. 如果 $\boldsymbol{x}$ 以 001 开头，那么当且仅当对于某个 $\beta\boldsymbol{f} \in \mathcal{F}$， $\boldsymbol{x} = 01\hat{}\bullet \boldsymbol{f}$， $\boldsymbol{x}$ 才是无重叠的。

注 3.9. 在此定理的第二部分中，该条件可以改写^为： $\boldsymbol{x} \in 001(0,1)^\omega$ 不是无重叠的，当且仅当 $\beta\boldsymbol{f}$ 在 $\mathcal{I}$ 中有一个因子，其中 $\boldsymbol{f}$ 描述了 $\boldsymbol{x}$ 的规范分解。特别地，如果 $\boldsymbol{f}$ 在 $\mathcal{I}$ 中有一个因子，那么 $\boldsymbol{x}$ 就不是无重叠的。对于优词，似乎无法得到类似的结果；我们将看到，存在一些字 $\boldsymbol{x} \in G_{001}$，其描述为 $\boldsymbol{x} = 001 \bullet \boldsymbol{f}$，使得 $\boldsymbol{f}$ 有一个因子 $\beta\alpha\beta\beta \in I$。另一方面， $\alpha\alpha\alpha\gamma \notin I$，但不能是任何满足 $001\bullet\boldsymbol{f} \in G.$的 $\boldsymbol{f}$ 的因子

设 $\boldsymbol{f} \in B^\omega$，使得 $001 \bullet \boldsymbol{f} = \boldsymbol{x}$。设 $w \in B^k$。于是

$$001 \bullet w\boldsymbol{f} = (001 \bullet w)\delta^k(01)^{-1}\delta^k(\boldsymbol{x}) \tag{2}$$

令 F_{001} 为可在图 3 的自动机 A_{001} 上行走的无限字集合。

定理 3.10。设 $\boldsymbol{x} \in \{0,1\}^\omega$。若 $\boldsymbol{x}$ 以 001 开头，则 $\boldsymbol{x}$ 优良，当且仅当对于某个 $\boldsymbol{f} \in F_{001}$，有 $\boldsymbol{x} = 001 \bullet f$。

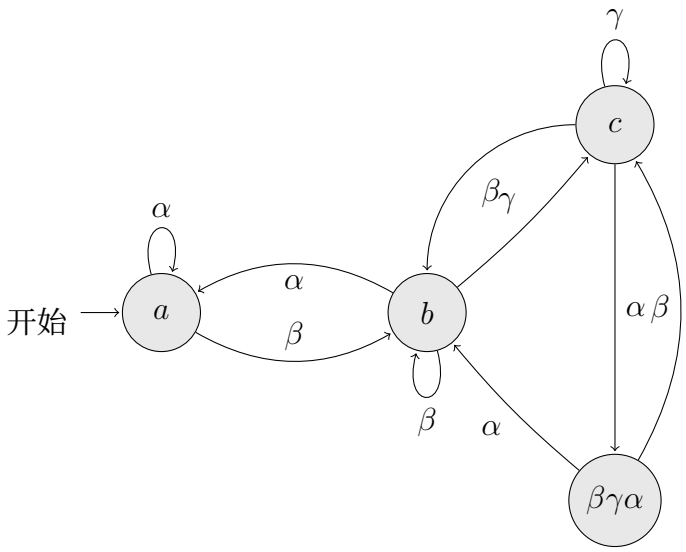
Let $W_{001} = \{\boldsymbol{f} \in B^\omega : 001 \bullet \boldsymbol{f} \in G\}$. We prove that $W_{001} = F_{001}$.

下面引理中的实体对应于 A_{001} 的状态，但状态 1 $\beta\gamma\alpha$除外；该状态以其从状态 a 出发的最短路径命名。图中省略了非接受汇点 (d)。因此，证明这些恒等式即可证明定理 3.10。注意到 $\beta\alpha\beta\beta \in I \cap F_{001}$。然而， $\alpha\alpha\alpha\gamma$ 是 F 中各字的前缀，但不能是任何 $\boldsymbol{f} \in F_{001}$ 的因子。

引理 3.11。以下恒等式成立：

(a) $(\beta\alpha)^{-1}W_{001} = W_{001} = \alpha^{-1}W_{001}$; (b) $(\beta\gamma\alpha\alpha)^{-1}W_{001} = \beta\gamma\beta)^{-1}W_{001} = (\beta\beta)^{-1}W_{001} = \beta^{-1}W_{001}$(c) $(\beta\gamma\alpha\beta)^{-1}W_{001} = \beta\gamma\gamma^{-1}W_{001} =(\beta\gamma)^{-1}W_{001}$(d) $(\beta\gamma\alpha\gamma)^{-1}W_{001} = \gamma^{-1} = \emptyset$

图 2：“Fife”自动机 A_{001} 用于 G_{001}



证明。令 $001 \bullet \boldsymbol{f} = \boldsymbol{x}$ 优良。
(a)

$$\alpha \boldsymbol{f} \in W_1 \iff 001 \bullet \alpha \boldsymbol{f} \in G \iff (001 \bullet \alpha)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \iff 0100(0100)^{-1}\delta(\boldsymbol{x}) \in G \iff 1\boldsymbol{x} \in G \text{ 或 } \boldsymbol{x} \in G_{001} \iff \boldsymbol{x} \in G \iff \boldsymbol{f} \in W_{001}$$

$$\beta\alpha \boldsymbol{f} \in W_{001} \iff 001 \bullet \beta\alpha \boldsymbol{f} \in G \iff (001 \bullet \beta\alpha)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \iff$$

$$001\,01000101(01000101)^{-1}\delta^2(\boldsymbol{x}) \in G \iff 0\delta(0\delta(\boldsymbol{x})) \in G \iff$$

$$10\delta(\boldsymbol{x}) \in G \text{ 或 } 0\delta(\boldsymbol{x}) \in G_{001} \iff 0\delta(\boldsymbol{x}) \in G \iff 1\boldsymbol{x} \in G \text{ 或 } \boldsymbol{x} \in G_{00\ \ 1}$$

$$\iff \boldsymbol{x} \in G \iff \boldsymbol{f} \in G$$

使得 $(\beta\alpha)^{-1}W_{001} = \alpha^{-1}W_{001} = W_{001}$。
(b)

$$\beta \boldsymbol{f} \in W_{001} \iff 001 \bullet \beta \boldsymbol{f} \in G \iff (001 \bullet \beta)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \iff$$

$$001\,0100(0100)^{-1}\delta(\boldsymbol{x}) \in G \iff 0\delta(0\boldsymbol{x}) \in G \iff 10\boldsymbol{x} \in G \text{ 或 } 0\boldsymbol{x} \in G_{001} \iff 0\boldsymbol{x} \in G \iff 10\boldsymbol{x}$$

$$\beta\beta\boldsymbol{f} \in W_{001} \iff 001 \bullet \beta\beta\boldsymbol{f} \in G \iff (01 \bullet \beta\beta)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \iff$$

$$010100\ \cancel{01000101(01000101)}^{-1}\delta^2(\boldsymbol{x}) \in G \iff \delta(0\delta(0\boldsymbol{x})) \in G$$

$$\iff 0\delta(0\boldsymbol{x}) \in G \iff 10\boldsymbol{x} \in G \text{ or } 0\boldsymbol{x} \in G_{001} \iff 10\boldsymbol{x} \in G$$

$$\beta\gamma\alpha\alpha\boldsymbol{f} \in W_{001} \iff 001\bullet\beta\gamma\alpha\alpha\boldsymbol{f} \in G \iff (01\bullet\beta\gamma\alpha\alpha)\delta^4(01)^{-1}\delta^4(\boldsymbol{x}) \in G \iff$$

$$001010001000\delta^4(\boldsymbol{x}) \in G$$

$$\iff 0\delta(0\delta^2(1\delta(\boldsymbol{x}))) \in G \iff 10\delta^2(1\delta(\boldsymbol{x})) \in G \text{ 或 } 0\delta^2(1\delta(\boldsymbol{x})) \in G_{001} \iff$$
$$0\delta^2(1\delta(\boldsymbol{x})) \in G \iff 1\delta(1\delta(\boldsymbol{x})) \in G \text{ 或 } \delta(1\delta(\boldsymbol{x}) \in G_{001} \iff 1\delta(1\delta(\boldsymbol{x})) \in G$$

$$\iff 01\delta(\boldsymbol{x}) \in G \iff \delta(0\boldsymbol{x}) \in G \iff 0\boldsymbol{x} \in G \iff 10\boldsymbol{x} \in G$$

因此 $((\beta\gamma\alpha\alpha)^{-1}W_{001} = \beta\gamma\beta)^{-1}W_{001} = (\beta\beta)^{-1}W_{001} = (\beta\alpha)^{-1}W_{001} = \beta^{-1}W_{001}$。

(c)

$$\beta\gamma\boldsymbol{f} \in W_{001} \iff 001 \bullet \beta\gamma\boldsymbol{f} \in G \iff (001 \bullet \beta\gamma)\delta^2(01)^{-1}\delta^2(\boldsymbol{x}) \in G \iff$$
$$00101000100\ 01000101(01000101)^{-1}\delta^2(\boldsymbol{x}) \in G \iff 0\delta(0\delta(00\boldsymbol{x}))$$
$$\in G \iff 10\delta(00\boldsymbol{x}) \in G \text{ 或 } 0\delta(00\boldsymbol{x}) \in G_{001} \iff 0\delta(00\boldsymbol{x}) \in G$$
$$\iff 100\boldsymbol{x} \in G \text{ 或 } 00\boldsymbol{x} \in G_{001} \iff 100\boldsymbol{x} \in G$$

$$\beta\gamma\gamma\boldsymbol{f} \in W_{001} \iff 001\bullet\beta\gamma\gamma\boldsymbol{f} \in G \iff (001\bullet\beta\gamma\gamma)\delta^3(01)^{-1}\delta^3(\boldsymbol{x}) \in G \iff$$

$$00101000100010001010100010\cancel{1\ 0100010101000100(0100010101000100)}^{-1}\delta^3(\boldsymbol{x}) \in G \iff$$

$$0\delta(0\delta^2(100\boldsymbol{x})) \in G \iff 0\delta^2(100\boldsymbol{x}) \in G \text{ or } \delta^2(100\boldsymbol{x}) \in G_{001} \iff$$

$$0\delta^2(100\boldsymbol{x}) \in G1\delta^(100\boldsymbol{x}) \in G \text{ or } \delta^(100\boldsymbol{x}) \in G_{001}$$

$$\iff 1\delta^(100\boldsymbol{x}) \in G \iff 0100\boldsymbol{x} \in G \iff 100\boldsymbol{x} \in G$$

$$\beta\gamma\alpha\beta\boldsymbol{f} \in W_{001} \iff 001\bullet\beta\gamma\alpha\beta\boldsymbol{f} \in G \iff (001\bullet\beta\gamma\alpha\beta)\delta^4(01)^{-1}\delta^4(\boldsymbol{x}) \in G \iff$$

$$0010100010001000101010001000\delta^4(\boldsymbol{x}) \in G \iff 0\delta(0\delta^2(1\delta(0\boldsymbol{x}))) \in G$$
$$\iff 10\delta^2(1\delta(0\boldsymbol{x})) \in G \text{ or } 0\delta^2(1\delta(0\boldsymbol{x})) \in G_{001} \iff 0\delta^2(1\delta(0\boldsymbol{x})) \in G$$
$$\iff 1\delta(1\delta(0\boldsymbol{x})) \in G \text{ or } \delta(1\delta(0\boldsymbol{x})) \in G_{001} \iff 1\delta(1\delta(0\boldsymbol{x})) \in G$$
$$\iff 01\delta(0\boldsymbol{x}) \in G \iff \delta(00\boldsymbol{x}) \in G \iff 00\boldsymbol{x} \in G \iff 100\boldsymbol{x} \in G$$

(d)

$\gamma \boldsymbol{f} \in W_{001} \iff 001\bullet\gamma\boldsymbol{f} \in G \iff (001\bullet\gamma)\delta(01)^{-1}\delta(\boldsymbol{x}) \in G \iff 00101\,0\cancel{100}\cancel{(0100)^{-1}}\delta(\boldsymbol{x}) \in G$

$$\iff 0\delta(00\boldsymbol{x})) \in G$$

但 $00\boldsymbol{x}$ 以前缀 0000 开头。因此 $\gamma^{-1}W_{001} = \emptyset$。

$\beta\gamma\alpha\gamma\boldsymbol{f} \in W_{001} \iff 001\bullet\beta\gamma\alpha\gamma\boldsymbol{f} \in G \iff (001\bullet\beta\gamma\alpha\gamma)\delta^4(01)^{-1}\delta^4(\boldsymbol{x}) \in G$

$$\iff 0010100010001000101010001000100010101000100\delta^4(\boldsymbol{x}) \in G \iff$$
$$0\delta(0\delta(^2(1\delta(00\boldsymbol{x}))) \in G$$

我们看到，$001\bullet\beta\gamma\alpha\gamma\boldsymbol{f}$包含了 $00\boldsymbol{x}$ 的前缀 0000，因此 $(\beta\gamma\alpha\gamma)^{-1}W_{001} = \emptyset$。 □

4 字典序极大的优良词

若 u 是一个长度为正的字，则用 u^- 表示从 u 中删去最后一个字母后得到的字。二进制词上的字典序递归定义如下：

$u < v \iff v \neq \epsilon$ 且 $((u = \epsilon)$ 或 $(u^- < v^-)$ 或 $((u = u^-0)$ 且 $(v = u^-1)))$.

注意，态射 δ 是反序的：设 u 和 v 是非空二进制词，且满足 $u < v$。记 $u = u'0u''$, $v = u'1v''$，其中 u' 是 u 和 v的最长公共前缀。于是 $\delta(u')01$ 是 $\delta(u)$的前缀，而$\delta(u')00$ 是 $\delta(v)$的前缀，因此 $\delta(u) > \delta(v)$.

对于每个非负整数 n，令 ℓ_n（分别为 m_n）表示长度为 n 且满足 ℓ_n（分别为 m_n）是某个单侧无限良词前缀的字典序最小（分别为字典序最大）字。

引理 4.1. 令 n 为非负整数。字 ℓ_n 是 ℓ_{n+1}的前缀。字 m_n 是 m_{n+1}的前缀。

证明。我们证明 ℓ_n 的结论； m_n 的证明类似。设 ℓ_nr 是一个单侧无限良词。令 p 为 ℓ_nr 的长度为 $n+1$ 的前缀，并令 q 为 ℓ_{n+1} 的长度为 n 的前缀。我们需要证明 $q = \ell_n$。p 和 q 都是某个单侧无限良词的前缀。根据定义，我们有 $\ell_{n+1} \leq p$ 和 $\ell_n \leq q$。如果 $\ell_n < q$，那么 $p^- = \ell_n < q = \ell^-_{n+1}$，因此 $p < \ell_{n+1}$。这就产生了矛盾。因此 $\ell_n = q$，证毕。 □

令 $\boldsymbol{\ell} = lim_{n\to\infty}\ell_n$， $\mathrm{m} = lim_{n\to\infty}m_n$。

引理 4.2. 字 $\boldsymbol{\ell}$ 是字典序最小的单侧无限良词。字 m 是字典序最大的单侧无限良词。

证明. 我们证明 $\boldsymbol{\ell}$ 是字典序最小的。证明 m 是字典序最大的过程类似。设 w 是一个单侧无限良词。对每个 n，令 w_n 为 w 的长度为 n 的前缀，使得 $\mathrm{w} = lim_{n\to\infty}w_n$。

若对某个 n 有 $w_n > \ell_n$，则 $\mathrm{w} > \boldsymbol{\ell}$。

否则，对所有 n都有 $w_n \leq \ell_n$。根据 ℓ_n 的定义，我们有 $w_n \geq \ell_n$，因此对所有 n都有 $w_n = \ell_n$。于是 $\mathrm{w} = \lim_{n\to\infty} w_n = \lim_{n\to\infty} \ell_n = \boldsymbol{\ell}$。

在所有情况下，我们都有 $\mathrm{w} \geq \boldsymbol{\ell}$。 □

引理 4.3。我们有 $\boldsymbol{\ell} = \delta(\mathrm{m})$。

证明。字 0001 是长度为 4 的最小优词。此外，0001 是 $\boldsymbol{d}$ 的一个因子，因此存在以前缀 0001 开头的单侧无限良词（它们是 $\boldsymbol{d}$ 的后缀）。因此， $\ell_4 = 0001$。由于 0001 是 $\boldsymbol{\ell}$ 的前缀，且 $\boldsymbol{\ell}$ 是优良的，根据定理 1.6，我们可将 $\boldsymbol{\ell} = \delta(\mathrm{m}')$ 写成某个 m' 的形式。由于 $\boldsymbol{\ell}$ 是优良的，根据引理 2.1，m' 是优良的。由此可得 $\mathrm{m}' \leq \mathrm{m}$。然而，如果 $\mathrm{m}' < \mathrm{m}$，那么$\delta(\mathrm{m}) < \delta(\mathrm{m}') = \boldsymbol{\ell}$，因为 δ 是反序的。这不可能，因为 $\boldsymbol{\ell}$ 是最小的。因此 $\mathrm{m}' = \mathrm{m}$，且 $\boldsymbol{\ell} = \delta(\mathrm{m})$。 □

引理 4.4。我们有 $0\mathrm{m} = \delta(\boldsymbol{\ell})$。

证明。由于 1 是长度为 1 的最大优词，且 $\boldsymbol{d}$ 的后缀以 1 开头，因此有 $m_1 = 1$。由引理 3.1，0m 是优词。由于 01 是 0m 的前缀，根据定理 1.6，我们可以将 0m 写作 $= \delta(\boldsymbol{\ell}')$，其中 $\boldsymbol{\ell}'$。由于 0m 是优词， $\boldsymbol{\ell}'$ 是优词，因此 $\boldsymbol{\ell}' \geq \boldsymbol{\ell}$。然而， δ 是保序的，所以如果 $\boldsymbol{\ell}' > \boldsymbol{\ell}$，那么 $\delta(\boldsymbol{\ell}) > \delta(\boldsymbol{\ell}') = \boldsymbol{m}$，这与 $\boldsymbol{m}$的极大性矛盾。因此 $\boldsymbol{\ell}' = \boldsymbol{\ell}$，且 $0\mathrm{m} = \delta(\boldsymbol{\ell})$。 □

定理 4.5. 字 $\boldsymbol{m}$是不动点

$$\boldsymbol{m} = h_1(\boldsymbol{m}) \tag{3}$$

其中 $h_1 = [1000, 1010]$。字 $\boldsymbol{\ell}$ 是不动点

$$\boldsymbol{\ell} = h_2(\boldsymbol{\ell}). \tag{4}$$

其中 $h_2 = [0001, 0101]$。

证明。由引理 4.3 和引理 4.4 可得 $\boldsymbol{m} = 0^{-1}\delta^2(\boldsymbol{m})$，因此对于每个 n，都有 $m_{4n-2} = 0^{-1}\delta^2(m_n)$。然而，对于 $a \in \{0,1\}$，$\delta^2(a) = 0h_1(a)0^{-1}$，因此对于 $w \in \{0,1\}^*$，$\delta^2(w) = 0h_1(w)0^{-1}$。由此可得 $m_{4n-2} = h_1(m_n)0^{-1}$。然后

$$\begin{aligned}
\boldsymbol{m} &= \lim_{n\to\infty} m_n \\
&= \lim_{n\to\infty} m_{4n-2} \\
&= \lim_{n\to\infty} h_1(m_n)0^{-1} \\
&= \lim_{n\to\infty} h_1(m_n) \\
&= h_1(\lim_{n\to\infty} m_n) \\
&= h_1(\boldsymbol{m})
\end{aligned}$$

$\boldsymbol{\ell}$ 的证明类似，利用 $h_2 = (01)^{-1}\delta^2 01$ 这一事实。 □

推论 4.6。$\boldsymbol{\ell}$ 或 $\boldsymbol{m}$ 的每个有限因子都是 $\boldsymbol{d}$ 的因子，反之亦然。然而，字 $\boldsymbol{\ell}$、$\boldsymbol{m}$ 和 $\boldsymbol{d}$ 没有公共后缀。

证明。由于 δ^2、h_1 和 h_2 彼此共轭，它们的不动点具有相同的因子。例如，设 u 是 $\boldsymbol{d} = \lim_{n\to\infty} \delta^n(0)$ 的一个因子。那么对于某个 n，u 是 $\delta^{2n}(0) = 0h_1^{2n}(0)0^{-1}$ 的一个因子，而后者是 $h_1^{2n}(1010) = h_1^{2n+1}(1)$ 的一个因子，也是 $\boldsymbol{m} = \lim_{n\to\infty} h_1^n(1)$ 的一个前缀。因此，u 是 $\boldsymbol{m}$ 的一个因子。

现在假设 $\boldsymbol{d}$ 和 $\boldsymbol{\ell}$ 有一个公共后缀 $\boldsymbol{s}$。（其他不动点对的证明类似。）设 $\boldsymbol{d} = d_i\boldsymbol{s}$ 和 $\boldsymbol{\ell} = \ell_j\boldsymbol{s}$ 分别为某个 i 和j，其中 d_n 是 $\boldsymbol{d}$ 中长度为 n的前缀。于是

$$\begin{aligned}
d_i\boldsymbol{s} &= \boldsymbol{d} \\
&= \delta^2(\boldsymbol{d}) \\
&= \delta^2(d_i)\delta^2(\boldsymbol{s})
\end{aligned}$$

因此 $\boldsymbol{s} = s_{3i}\delta^2(\boldsymbol{s})$，其中 s_n 是 $\boldsymbol{s}$ 中长度为 n的前缀。类似地，

$$\begin{aligned} \ell_j \boldsymbol{s} &= \boldsymbol{\ell} \\ &= h_2(\boldsymbol{\ell}) \\ &= h_2(\ell_j) h_2(\boldsymbol{s}) \end{aligned}$$

因此 $\boldsymbol{s} = s_{3j} h_2(\boldsymbol{s})$。于是

$$\begin{aligned} s_{3j} h_2(\boldsymbol{s}) &= \boldsymbol{s} \\ &= s_{3i}\delta^2(\boldsymbol{s}) \\ &= s_{3i} 01 h_2(\boldsymbol{s}) \\ &= s_{3i+2} h_2(\boldsymbol{s}) \end{aligned}$$

由此可知，$h_2(\boldsymbol{s})$ 的周期为 $|3(i-j)+2|$。然而，$h_2(\boldsymbol{s})$是优良的，且不包含四次幂。因此，$3(i-j) + 2 = 0$。这是不可能的，因为 i 和 j 是整数。 □

5 $\boldsymbol{d}$中的二进制模式

每个字都会将其因子作为模式遇到。如果一个字遇到某个二进制模式 p，它必然也会遇到补 $\overline{p}$。Shur [12] 已证明，除取补外，t 遇到的唯一二进制模式是它的因子和 00100。$\boldsymbol{d}$ 的情况则更为复杂。

引理 5.1. $\boldsymbol{d}$ 的任意因子 $0u$ 都可写成 $h_1(p)$，其中 p是满足 $h_1 = [0, 01]$ 的某个字。$\boldsymbol{d}$ 的任意因子 $u0$ 都可写成 $h_2(p)$，其中p 是满足 $h_1 = [0, 10]$ 的某个字。因此，字 $\boldsymbol{d}$ 在 h_1和 h_2各自之下都有一个原像。

证明。$\boldsymbol{d}$ 中的每个 1 前后都各有一个 0。 □

推论 5.2. 设 $0u$ 是 $\boldsymbol{d}$ 的一个因子，且满足 $|0u| \geq 13$。那么 $0u$ 可写成 $h_1(p)$，其中 p 既不是 $\boldsymbol{d}$ 的因子，也不是 $\boldsymbol{d}$ 的某个因子的补。因此，$\boldsymbol{d}$ 会遇到无穷多个模式 p，使得 p 和 $\overline{p}$ 都不是 $\boldsymbol{d}$ 的因子。

证明。 $\boldsymbol{d}$ 中不含 010001 的最长因子是 100010101000，其长度为 12。因此， $0u = h_1(p)$ 含有因子 $010001 = h_1(1001)$，而 p 含有因子 1001。1001 和 0110 都不是 $\boldsymbol{d}$ 的因子，因此 p 和 $\overline{p}$ 都不是 d 的因子。□

对于任意给定的模式 p，自动证明系统 Walnut [9] 理论上只要有足够的计算能力和时间，就能判断 $\boldsymbol{d}$ 是否遇到 p。然而，在下一个定理中，我们有效地刻画了 $\boldsymbol{d}$ 遇到的所有二进制模式 p。本节余下部分将用于证明该定理。

定理 5.3. 字 p 是 $\boldsymbol{d}$ 遇到的二进制模式，当且仅当以下条件之一成立

1. p 和 $\overline{p}$ 中有一个是 $\boldsymbol{d}$、$h_1(\boldsymbol{d})$ 或 $h_2(\boldsymbol{d})$ 的因子。 2. p 和 $\overline{p}$ 中有一个属于 0010100、01001001000、00100100100、001001001000、00010010010、000100100100、0010001000100、00100010001000、00010001000100 和 000100010001000。

这两种可能性互不相同。

下面这个引理 2.5 的类似结论将有助于分析 w 中出现的模式：

引理 5.4. 设 u、v 是二进制词，且 uvu 是 $\boldsymbol{d}$、$|u| \geq 3$ 的因子，并且满足以下条件之一

- $|u|_{00} \geq 1$，或
- $|u|_{10101} \geq 1$。

那么 $|uv| \equiv 0$ （模 4）。

证明。假设 00 是 u 的因子，且 $|u| \geq 3$。那么字 u 含有因子 000、001 和 100 之一。这些因子在 $\boldsymbol{d}$ 中只能分别作为 $\boldsymbol{d}$ 的某个形如 $\delta^2(p0)0$、$\delta^2(p0)01$ 和 $\delta^2(p0)$ 的前缀的后缀出现。因此，因子 000、001 或 100 的任意一次出现的位置，以及由此 u 的任意一次出现的位置，在 $\boldsymbol{d}$ 中模 4 固定，结论得证。

假设 $|u|_{10101} \geq 1$。因子 10101 只在 $\boldsymbol{d}$ 中出现，且是 $\boldsymbol{d}$ 的某个形如 $\delta^2(p10)(00)^{-1}$ 的前缀的后缀； $\boldsymbol{d}$ 中 u 的任意一次出现的位置也再次固定，模 4 同余。□

定理 5.5。假设 $\boldsymbol{d}$ 遇到二进制模式 p。那么 p和 $\overline{p}$ 之一要么

1. 是 $\boldsymbol{d}$, $h_1^{-1}(\boldsymbol{d})$ 的因子，或是 $h_2^{-1}(\boldsymbol{d})$ 的因子；或者 2. 具有这样的性质：它所有形如 10^k1 的因子长度都相同。

证明。不失一般性，如有必要，用 $\overline{p}$ 替换 p，假设 0 是 p 的首字母。假设 p 和 $\overline{p}$ 都不是 $\boldsymbol{d}$, $h_1(\boldsymbol{d})$ 或 $h_2(\boldsymbol{d})$ 的因子。由于 000 是 $\boldsymbol{d}$ 的因子，而 $\boldsymbol{d}$ 中不出现 0000，因此必有 $|p|_1 > 0$。设 $g(p)$ 是 $\delta^n(0)$ 的一个因子，其中 $g = [X, Y]$ 是非擦除态射，且 n 尽可能小。由引理 2.4，$|X|$ 和 $|Y|$ 不可能同时为偶数。我们根据 $|X|_1, |Y|_1$ 分情况讨论。

情况 1 我们有 $|X|_1, |Y|_1 > 0$：如果 XX（分别地，YY, XYX, YXY）是 $\delta^n(0)$ 的因子，那么由引理 2.5，$|X|$（分别地，$|Y|$, $|XY|$, $|YXY|$）为偶数。因此，XX 和 YY 不可能同时是 $g(p)$ 的因子。同样，XX 和 XYX 不可能同时是 $g(p)$ 的因子，否则 $|X|$ 和 $|XY|$ 都为偶数，从而迫使 $|Y|$ 为偶数。类似地，YY 和 YXY 不可能同时是 $g(p)$ 的因子。

> 如果 XX 和 YY 都不是 $g(p)$ 的因子，那么 $g(p)$ 就是由 X 和 Y 组成的交替串，因此是 $XYXYXYX$ 的前缀。（由于 $\boldsymbol{d}$ 是优良的，四次幂 $(XY)^4$ 不是 $\boldsymbol{d}$ 的因子。）但这样一来，p 就是 0101010 的前缀，而 0101010 是 $\boldsymbol{d}$ 的因子。这就产生了矛盾。
>
> 那么，假设 XX 是 $g(p)$ 的因子。于是 YY 和XYX 都不是 $g(p)$ 的因子。若 $|p|_1 \geq 2$，则 $g(p)$ 会有一个形如 X^rYX^sY, $r, s \geq 1$ 的前缀。这包含因子 XYX，这是不可能的。由此可得 $|p|_1 \leq 1$。由于 $\boldsymbol{d}$ 是优良的，$g(p)$ 不可能含有因子 $XXXX$，因此我们得出 p 是 0001000 的因子。而显然，0001000 是 $\boldsymbol{d}$ 的因子。
>
> 最后，假设 YY是 $g(p)$ 的因子。那么 XX和 YXY都不是 $g(p)$ 的因子。因此，p以 0 开头，且含有因子 11，但不含 00、101 或 1111。（若 p 中有因子 1111，则优词 $\boldsymbol{d}$ 中会出现四次幂。）由此可知，p 是 011、0110、0111~~1~~和 01110 之一。~~然~~而，$h_1(011) = 0100$、$h_1(0110) = 010001$、$\overline{0111} = 1000$以及$\overline{0111} = 1000$ 都是 $\boldsymbol{d}$ 的因子。

情况 2 我们有 $|X|_1 = |Y|_1 = 0$。这迫使 $g(p)$ 成为 000 的因子，因此 p 是长度不超过 3 的二进制词。每个这样的字或其补都是 $\boldsymbol{d}$ 的因子。

情况 3 我们有 $|X|_1 > 0$，但 $|Y|_1 = 0$。写作 $Y = 0^n$，其中 $1 \leq n \leq 3$。

情形 3(a) 我们有 $|X| \geq 3$，且 $|X|_{00} \geq 1$ 或 $|X|_{10101} \geq 1$。如果 p只含有一个 0（即它的第一个字母），那么由于四次幂不会出现在 $\boldsymbol{d}$中，p 必须是 0111 的前缀，而 0111 是 $\boldsymbol{d}$的一个因子的补。否则，p有一个形如 01^k0 的因子，其中 k、$0 \leq k \leq 3$。因此，$XY^kX = X0^{nk}X$ 是$\boldsymbol{d}$的一个因子。由引理 5.4，取 $a = X$、$b = 0^{nk}$，可得 $|X| \equiv -nk$（模 4）。如果 $n = 2$，那么由于 0^4 不是 $\boldsymbol{d}$的因子，可得 $0 \leq k \leq 1$。此外，$|X| \equiv -nk$（模 4），所以 $|X|$是偶数，从而得到 $k \equiv -|X|/2$（模 2），并且 k 已确定；p 中每一对 0 之间都由相同数量的 1 隔开。于是 $\overline{p}$具有这样的性质：它所有形如 10^k1 的因子长度都相同。如果 $n = 1$ 或 $n = 3$，那么由于 $0 \leq k \leq 3$，k 由同余关系 $|X| \equiv -nk$（模 4）确定，并且同样地，$\overline{p}$ 具有这样的性质：它所有形如 10^k1的因子长度都相同。情形 3(b) 要么 $|X| < 3$，要么 $|X|_{00} = |X|_{10101} = 0$。如果 $|X|_{00} = 0$，那么 X不含因子 00 或 11，因此是由 0 和 1 组成的交替串。因此，给定条件意味着 $X = 00$，或者 X 是 01010 的一个因子。如果 $X = 00$，那么 $g(p)$ 必须是 000 的一个因子，而 000 是$\boldsymbol{d}$的一个因子。于是，假设 X 是 01010 的一个因子。如果 p 至多只含有一个 0（即它的第一个字母），那么 p 就会是 0111 的前缀，而 0111 是 $\boldsymbol{d}$的一个因子的补。这是不可能的。因此，假设 p 有一个形如 01^k0的因子。如果 $|X| \geq 3$，那么 X 是 010、101、0101、1010 或 01010 中的一个。可以验证，对于 X 的每一种可能，恰有一个 k 的值使得 $X0^kX$是 $\boldsymbol{d}$的一个因子：

- 如果 X 是 010 或 01010，$k = 1$。
- 如果 $X = 101$，那么因子 XX 将意味着 $\boldsymbol{d}$ 含有因子 11；$\boldsymbol{d}$ 中的因子 $X0X$ 会向右扩展成一个因子

$X0X0 = (10)^4$ 属于 $\boldsymbol{d}$；因子 $X00X$ 将意味着 $\boldsymbol{d}$ 含有因子 1001；因此 $k = 3$。

- 如果 X 是 0101 或 1010，那么 XX 是一个四次幂，而 $X0X$ 含有因子 1001； $X000X$ 含有 0000；因此 $k = 2$。

因此，若 $|X| \geq 3$，则 $\overline{p}$ 具有这样的性质：它所有形如 10^k1 的因子长度都相同。剩下的情形是 $|X| \leq 2$ 且 $X \neq 00$，也就是说，X 是 1、01 和 10 之一。$X = 1$。由于 $\boldsymbol{d}$ 含有因子 XY^kX，所以 $Y = 00$ 不可能，否则 $\boldsymbol{d}$ 会含有 0000 或 1001。因此 $Y = 0$。于是 $g = [1, 0]$，而 p 是 $\boldsymbol{d}$ 的某个因子的补。X=01。由于 $|X|$ 是偶数，我们必须有 $|Y|$ 为奇数，因此 $n = 1$ 或 $n = 3$。若 $n = 3$，则 $YX = 0^41$ 和 $YY = 0^6$ 不是 $\boldsymbol{d}$ 的因子，这就迫使 $|p|_1 = 1$ 成立，而 p 是 00010000 的因子，它又是 $\boldsymbol{d}$ 的因子。于是设 $n = 1$。然而现在 $g = [01, 0]$，且 $h_1(\overline{p}) = g(p)$ 是 $\boldsymbol{d}$ 的因子，因此 $\overline{p}$ 是 $h_1^{-1}(\boldsymbol{d})$ 的因子。X=10。同样，$|X|$ 是偶数，因此可得 $n = 3$ 或 $n = 1$。在第一种情形下，XY 含有 0000，这是不可能的。因此 p 是 000 的因子，也是 $\boldsymbol{d}$ 的因子。在第二种情形下，$g = [10, 0]$，且 $\overline{p}$ 是 $h_2^{-1}(\boldsymbol{d})$ 的因子。

情况 4 我们有 $|X|_1 = 0$ 但 $|Y|_1 > 0$。写作 $X = 0^n$，其中 $1 \leq n \leq 3$。

情形 4(a) 我们有 $|Y| \geq 3$，且 $|Y|_{00} \geq 1$ 或 $|Y|_{10101} \geq 1$。如果 p 至多包含一个 1，那么由于 $\boldsymbol{d}$中不会出现四次幂，p 必定是 0001000 的因子，而 0001000 是 $\boldsymbol{d}$的因子。否则，p 含有某个形如 10^k1 的因子，其中 k，$0 \leq k \leq 3$。因此，$YX^kY = Y0^{nk}Y$是 $\boldsymbol{d}$的因子。由引理 5.4，取 $a = Y$、$b = 0^{nk}$，可得 $|Y| \equiv -nk \pmod 4$。如果 $n = 2$，那么由于 0^4 不是 $\boldsymbol{d}$ 的因子，可得 $0 \leq k \leq 1$。此外，$|Y| \equiv -nk$，因此 $|Y|$ 为偶数，从而 $k \equiv -|Y|/2 \pmod 2$，且k 已确定。因此，p具有这样的性质：其所有形如 10^k1 的因子长度都相同。如果 $n = 1$ 或 $n = 3$，那么由于 $0 \leq k \leq 3$，k 由同余关系 $|Y| \equiv -nk \pmod 4$ 确定；同样，p 具有这样的性质：其所有形如 10^k1 的因子长度都相同。

情形 4(b) 要么 $|Y| \leq 3$，要么 $|Y|_{00} = |Y|_{10101} = 0$。如果 $|Y|_{00} = 0$，那么 Y不包含因子 00 或 11，并且是由 0 和 1 组成的交替串。因此，给定条件意味着 $Y = 00$，或者 Y是 01010 的因子。如果 $Y = 00$，那么 $g(p)$ 必定是 000 的因子，而 000 是$\boldsymbol{d}$的因子。于是，假设 Y 是 01010 的因子。如果 p至多包含一个 1，那么 p 必定是 0001000 的因子，而 0001000 是 $\boldsymbol{d}$的因子。因此，假设 p 含有某个形如 10^k1的因子。如果 $|Y| > 2$，按照情形 3(a) 中的分析可知，p 具有这样的性质：其所有形如 10^k1 的因子长度都相同。通过类似于情形 3(a) 的分析可知，$|Y| \leq 2$ 的情形不可能发生。

□

引理 5.6。 $\boldsymbol{d}$ 的因子集合在反转下封闭。

证明。首先，我们注意到，如果 u 是一个二进制词，那么

$$(\delta(u))^R = 0^{-1}\delta(u^R)0.$$

假设 v 是 $\boldsymbol{d}$的一个因子。可以看出，$\boldsymbol{d}$ 的长度为 2 的因子集合为$\{00, 01, 10\}$，该集合在反转下封闭。因此，如果 $|v| \leq 2$，那么 v^R 是 $\boldsymbol{d}$的一个因子。假设 $|v| > 2$，并且对于 $\boldsymbol{d}$的每个更短因子 u，其反转 u^R 都是 $\boldsymbol{d}$的一个因子。然而，$|v| > 2$ 意味着 v 是δ（u）的一个因子，而 u 是 $\boldsymbol{d}$ 的一个因子，且其长度小于 v。因此，u 和 u^R 都是 $\boldsymbol{d}$的因子。于是，$(\delta(u))^R = 0^{-1}\delta(u^R)0$也是 $\boldsymbol{d}$的一个因子，因此 v^R 也如此。由归纳法可得结论。 □

推论 5.7。如果 $\boldsymbol{d}$ 遇到模式 p，那么 $\boldsymbol{d}$ 也会遇到模式 p^R。

定理 5.8。以下条件等价：

1. 字 p 是 $\boldsymbol{d}$中出现的二进制模式，但 p 和 $\overline{p}$ 都不是 $\boldsymbol{d}$、$h_1(\boldsymbol{d})$ 或 $h_2(\boldsymbol{d})$ 的因子。2. p 和 $\overline{p}$其中之一属于以下字符串：0010100、01001001000、00100100100、001001001000、00010010010、000100100100、0010001000100、00100010001000、00010001000100 和 000100010001000。

证明。（$1 \Rightarrow 2$）：假设字 p 是 $\boldsymbol{d}$中出现的二进制模式，但 p 和 $\overline{p}$ 都不是 $\boldsymbol{d}$、$h_1(\boldsymbol{d})$ 或 $h_2(\boldsymbol{d})$ 的因子。根据定理 5.5，必要时用 $\overline{p}$ 替换 p ，存在一个数 k ，使得在 p中相邻的 0 之间恰好有 k 1 个 1。因此，对于某些 k、r、s、t，有 $p = 0^r(10^k)^s10^t$。由于 $\boldsymbol{d}$不含四次幂， k、r、s、$t \leq 3$。出于同样的原因，如果 $s = 3$，那么 $r + s < k$。如果 $s = 0$，那么 p 是 0001000 的因子，而 0001000 是 $\boldsymbol{d}$的因子。因此，我们有 $s \geq 1$。下面根据 k的取值分类讨论：

$k = 0$

在这种情况下，为了避免出现四次幂，我们还必须有 $s \leq 2$，并且 p 是 00011000 或 000111000 的一个因子。注意到 $h_1(\overline{00011000}) = 01010100010101$ 是 $\boldsymbol{d}$的一个因子。因此， p 不能是 00011000 的因子，否则$\overline{p}$就会是 $h_1^{-1}(\boldsymbol{d})$ 的一个因子。由此得出， p 是 000111000 的一个因子。注意， $h_2(00011100) = h_1(00111000) = 00010101000$ 是 $\boldsymbol{d}$的一个因子，因此p不能是 000111000 的真因子。于是我们必有$p = 000111000$。为了导出矛盾，我们证明 $\boldsymbol{d}$ 不包含p。

假设 $g(p)$ 是 $\boldsymbol{d}$的一个因子， $g = [X, Y]$ 是一个非擦除态射。根据引理 2.4，不妨设 $|X|$ 和 $|Y|$ 中有一个是奇数。 XX 和 YY都是 $g(p) = XXXYYYXXX$ 的因子，因此根据引理 2.5，我们必有 $|X|_1 = 0$或 $|Y|_1 = 0$。

设 $|X|_1 = 0$。因此， XXX 以 000 结尾。由于 $XXXY$ 是 $g(p)$ 的因子，而后者又是 $\boldsymbol{d}$ 的因子， Y 的第一个字母是 1。由于 XXX 以 000 开头，且 $YXXX$ 是 $g(p)$ 的因子， Y 的最后一个字母是 1。但这样一来，YY 就包含因子 11，于是 11 是优词 $\boldsymbol{d}$ 的因子，这是不可能的。

设 $|Y|_1 = 0$。在前一段的论证中交换 X 和 Y，同样可以说明这是不可能的。

$k = 1$

根据 s 是 1、2 还是 3，我们分为以下子情形：

$s = 1$

如果 $r + t \geq 5$，那么 p 包含 00010100 或其反向词。然而， $\boldsymbol{d}$ 是优良的，因此不会遇到 00010100。根据推论 5.7， $\boldsymbol{d}$不会遇到 $\boldsymbol{d}$ 的反向词。由此可知 $r + t \leq 4$，因此 p 是 0101000 的因子，

0010100，或 0001010。然而，0101000 和 0001010 都是 $\boldsymbol{d}$ 的因子，因此 p不可能是其中之一的因子。因此， p 必定是 0010100 的因子。010100 和 010100 都是 $\boldsymbol{d}$ 的因子，因此 p 不可能是 0010100 的真因子，这就迫使 $p = 0010100$。

$s = 2$

在这种情况下， p 是 00010101000 的一个因子，而 00010101000 是 $\boldsymbol{d}$ 的一个因子。

$s = 3$

如果 $r = t = 0$，那么 $p = 1010101$ 是 $\boldsymbol{d}$ 的一个因子的补。然而，如果 $r > 0$ 或 $t > 0$，那么 p 包含四次幂 01010101 和 10101010 之一，因此 $\boldsymbol{d}$ 不会遇到 p，这与矛盾相悖。

$k = 2$

我们根据 s 是 1、2 还是 3，分为以下子情形:

$s = 1$

如果 $r + t \leq 5$ ，那么 p 是 000100100 或 001001000 的因子。然而，h_2(000100100) $= 00010001000 = h_1$(001001000) 是 $\boldsymbol{d}$ 的一个因子，因此 p 是 $h_1^{-1}(\boldsymbol{d})$ 或 $h_2^{-1}(\boldsymbol{d})$ 的一个因子，这与矛盾相悖。如果 $r = t = 3$，那么 p =0001001000 $= h_2$(00010100)，且 $\boldsymbol{d}$ 会遇到 00010100。由于 $\boldsymbol{d}$ 是优良的，这是不可能的。

$s = 2$

如果 $r + t \leq 3$，那么 p 是 1001001000、0100100100、0010010010 或 0001001001 的因子；然而， h_1(1001001000) $= 0100010001000$、 h_2(0100100100) =0100010001000、 h_1(0010010010) $= 0001000100010$以及 h_2(0001001001) =0001000100010 都是 $\boldsymbol{d}$ 的因子，因此 p 是 $h_1^{-1}(\boldsymbol{d})$ 或 $h_2^{-1}(\boldsymbol{d})$ 的因子。

如果 $r = t = 3$，那么 $p = 0001001001000$。为导出矛盾，我们证明 $\boldsymbol{d}$ 不包含 p。否则，不失一般性地假设 $\boldsymbol{d}$ 有一个因子 $g(p)$，其中 $g = [X, Y]$ 是非擦除态射，且 $|X|$和$|Y|$ 不全为奇数。字 $\boldsymbol{d}$ 有因子 $XXXYXXYXXYXXX$。我们不能有 $|X|_1 > 0$，否则由于 XX 和 XYX 是 $\boldsymbol{d}$的因子，引理 2.5 将迫使 $|X|$ 和 $|Y|$ 均为偶数。记 $X = 0^n$。由于 XX 是一个

$\boldsymbol{d}$的因子，我们必有 $n=1$ 和 $X=0$。由于 $XXXY$ 和 $YXXX$ 是 $\boldsymbol{d}$的因子， Y 的首字母和末字母都是 1。于是 $YXXY$ 包含因子 1001，而 1001 不可能是 $\boldsymbol{d}$的因子。这就产生了矛盾。

由此可得 $4 \leq r+t \leq 5$，因此 p 是 01001001001000、00100100100100、00010010010010、001001001001000 和 000100100100100 之一。

$s=3$

若 $r+t \leq 1$，则 p 是 01001001001 或 10010010010 的因子。然而h_2 (01001001001) $=010001000100010$ 和 h_1(010001000100010) 都是 $\boldsymbol{d}$ 的因子，因此 p 是 $h_1^{-1}(\boldsymbol{d})$ 或 $h_2^{-1}(\boldsymbol{d})$ 的因子，这就导致了矛盾。另一方面，若 $r+t \geq 2$，则四次幂 $(001)^4$、$(010)^4$ 和 $(100)^4$ 中有一个是 p 的因子，而 $\boldsymbol{d}$ 包含一个四次幂，这是不可能的。

$k=3$

我们根据 s 是 1、2 还是 3 来划分子情形:

$s=1$

在这种情况下， p 是 00010001000 的因子，而后者是 $\boldsymbol{d}$ 的因子。

$s=2$

如果 $r \leq 1$，那么 p 是 0100010001000 的因子，而后者是 $\boldsymbol{d}$ 的因子。这是不可能的。同理， $t \leq 1$ 也不可能。因此， p 必须是 0010001000100、00010001000100、00100010001000 和 000100010001000 之一。

$s=3$

如果 r， $t \leq 1$，那么 p 是 010001000100010 的因子，而后者是 $\boldsymbol{d}$ 的因子，这就产生了矛盾。假设 $r \geq 2$ 或 $t \geq 2$。那么 p 以 $q=001000100010001$ 或其反向词为因子。根据推论 5.7，只需证明 $\boldsymbol{d}$ 不会遇到 q。

为导出矛盾，假设 $g(q)$ 是 $\boldsymbol{d}$ 的因子，其中 $g=[X,Y]$ 是一个非擦除态射，且 $|X|$ 和 Y 中有一个是奇数。因此，字 $\boldsymbol{d}$ 含有因子 $XXYXXXYXXXYXXXY$。由于 XX 和 XYX 是 $\boldsymbol{d}$ 的因子，我们必有 $X=0^n$，其中某个 n；否则，引理 2.5 表明 $|X|$ 和 $|Y|$ 都是偶数，这就产生了矛盾。由于 XX 是 $\boldsymbol{d}$ 的因子， $n=1$

以及 $X = 0$。由于 $YXXX = Y000$ 是 $\boldsymbol{d}$ 的因子， Y 的最后一个字母是 1。因此，在 $\boldsymbol{d}$ 中， Y 后面总是跟着 0，所以 $XXYXXXYXXXYXXXY0$ 是 $\boldsymbol{d}$ 的因子。然而， $XXYXXXYXXXYXXXY0 = (00Y0)^4$, 和 $\boldsymbol{d}$ 含有一个四次幂。这是不可能的。

$(1 \Leftarrow 2)$: 我们证明， $\boldsymbol{d}$ 包含所列出的每个模式，但这些模式及其补都不是 $\boldsymbol{d}$、 $h_1(\boldsymbol{d})$ 或 $h_1(\boldsymbol{d})$ 的因子。

首先，我们证明 $\boldsymbol{d}$ 包含所列出的模式。注意，这些模式分别是 0010100、001001001000、000100100100 或 000100010001000 的因子。由于 001001001000 是 000100100100 的反向词，根据引理 5.7，只需证明 $\boldsymbol{d}$ 包含 0010100、001001001000 或 000100010001000。

- $p = 0010100$: Word $\boldsymbol{d}$ has the factor

$$00010001000 = g(0010100)$$

  where $g = [0, 010]$.

- $p = 001001001000$: Word $\boldsymbol{d}$ has the factor

$$000101010001010100010101000 = g(001001001000)$$

  where $g = [0, 010101]$.

- $p = 000100010001000$: Word $\boldsymbol{d}$ has the factor

$$000101010001010100010101000 = g(000100010001000)$$

  where $g = [0, 10101]$.

接下来，我们证明这些模式及其补都不是 $\boldsymbol{d}$、 $h_1^{-1}(\boldsymbol{d})$ 或 $h_2^{-1}(\boldsymbol{d})$ 的因子。注意， $h_1(101) = 01001$ 和 $h_2(101) = 10010$，且 1001 不是 $\boldsymbol{d}$ 的因子，因为 $\boldsymbol{d}$ 是优良的。所有所列模式的补都包含因子 101，因此这些补都不是 $h_1^{-1}(\boldsymbol{d})$ 或 $h_2^{-1}(\boldsymbol{d})$ 的因子。此外，所有所列因子的补都包含因子 11，因此这些补都不是 $\boldsymbol{d}$. 的因子。因此，只需证明所列模式都不是 $\boldsymbol{d}$、$h_1^{-1}(\boldsymbol{d})$ 或 $h_2^{-1}(\boldsymbol{d})$ 的因子。

列出的每种模式都以 0010100、01001001000、00100100100、00010010010 和 000100010001000 中的一个为因子。因此，只需证明这五种模式都不是 $\boldsymbol{d}$、 $h_1^{-1}(\boldsymbol{d})$ 或 $h_2^{-1}(\boldsymbol{d})$ 的因子。

情况 1：0010100。

如果这是 $\boldsymbol{d}$ 的因子，那么其中的 1 就会迫使它出现在上下文 00010100 中；但这是不可能的，因为 $\boldsymbol{d}$ 是优良的。由于 101 是 0010100 的因子，因此它也不是 $h_1^{-1}(\boldsymbol{d})$ 或 $h_2^{-1}(\boldsymbol{d})$ 的因子。

情况 2：01001001000。

其中包含因子 1001，因此它不是 $\boldsymbol{d}$ 的因子。

字 $00100010001000 = h_1$(01001001000) 不能是 $\boldsymbol{d}$ 的因子；如果其后跟着 0，那么 0000 将是 $\boldsymbol{d}$ 的因子；如果其后跟着 10，那么 $(0100)^4$ 是 $\boldsymbol{d}$ 的因子，这不可能。

字 h_2(01001001000) 以四次幂 0000 结尾，因此不能是 $\boldsymbol{d}$ 的因子

情况 3：00100100100。

其中包含因子 1001，因此不是 $\boldsymbol{d}$ 的因子。

字 h_1(00100100100) $= 00010001000100$ 不可能是 $\boldsymbol{d}$ 的因子；如果后面接 1，那么 1001 就会是 $\boldsymbol{d}$ 的因子；如果后面接 00，那么 0000 就是 $\boldsymbol{d}$ 的因子；如果后面接 01，那么 $(0001)^4$ 就是 $\boldsymbol{d}$ 的因子，这是不可能的。

字 $00100010001000 = h_2$(00100100100) 不可能是 $\boldsymbol{d}$ 的因子，理由如情况 2 所述。

情况 4：00010010010。

其中包含因子 1001，因此不是 $\boldsymbol{d}$ 的因子。

字 h_1(00010010010) 以 0000 开头，因此不可能是 $\boldsymbol{d}$ 的因子。

字 $00010001000100 = h_2$(00010010010) 不可能是 $\boldsymbol{d}$ 的因子，正如情况 3 中所论证的。

情况 5：000100010001000。

这不可能是 $\boldsymbol{d}$ 的因子；如果后面跟着 0，就会得到因子 0000；如果后面跟着 1，就会得到因子 $(0001)^4$, ，这是不可能的。

字 h_1(000100010001000) 以 0000 开头，因此不可能是$\boldsymbol{d}$的因子。

字 h_2(000100010001000) 以 0000 结尾，因此不可能是$\boldsymbol{d}$的因子。 □

参考文献

[1] J. Berstel. 关于无重叠词的 Fife 定理的一种改写。载于 J. Karhum¨aki、H. Maurer 和 G. Rozenberg 编，《理论计算机科学的成果与趋势》，《计算机科学讲义》第 812 卷，第 19–29 页。施普林格出版社，1994[2] J.-P. Allouche, J. Currie, J. Shallit. 极值无限无重叠二进制词。《电子组合学杂志》5，（1998）#R27。[3]J.-P. Allouche 和 J. Shallit，《无处不在的 Prouhet-Thue-Morse 序列》，载于《序列及其应用》（新加坡，1998），施普林格出版社，伦敦，1999，第 1–16 页。[4] A. Carpi. 无重叠词与有限自动机。《理论计算机科学》115（1993），243–260。[5] J. Cassaigne. 无重叠二进制词计数。载于 STACS 93，第十届计算机科学理论方面研讨会论文集，《计算机科学讲义》第 665 卷，第 216–225 页。施普林格出版社，1993。[6] D. Damanik. 倍周期序列中的局部对称性。《离散应用数学》100（2000），115–121。[7] E. Fife. 不含 BBb 的二进制序列。《美国数学会会刊》261（1980），115–136。[8] Y. Kobayashi. 不可约二进制词的枚举。《离散应用数学》20（1988），221–232。[9]H. Mousavi. Walnut 中的自动定理证明。文档（2016–2021），见 `https://arxiv.org/abs/1603.06017` 。[10] N. Rampersad，《无重叠词及其推广》。博士论文，滑铁卢大学，2007。

[11] A. Restivo 和 S. Salemi. 双符号上的无重叠词。载于 M. Nivat 和 D. Perrin 编，《无限词上的自动机》，《计算机科学讲义》第 192 卷，第 198–206 页。施普林格出版社，1984。[12] A. M. Shur. 图厄-莫尔斯序列所避开的二进制词。《半群论坛》53，第 212–219 页，1996a。.[13] A. Thue. 关于无限符号序列。Kra. Vidensk. Selsk. Skrifter. I. Mat. Nat. Kl. 7（1906），1–22。重印于《阿克塞尔·图埃数学论文选集》，T. Nagell 等编，Universitetsforlaget，奥斯陆，1977，第 139–158 页。..[14] A. Thue. 关于某些符号序列中相同部分的相对位置。Kra. Vidensk. Selsk. Skrifter. I. Mat. Nat. Kl. 1（1912），1–67。重印于《阿克塞尔·图埃数学论文选集》，T. Nagell 等编，Universitetsforlaget，奥斯陆，1977，第 413–477 页。