Fifty Years of Fine and Wilf*

* Well, *almost* fifty years...

Jeffrey Shallit
School of Computer Science, University of Waterloo
Waterloo, Ontario N2L 3G1, Canada
shallit@cs.uwaterloo.ca
http://www.cs.uwaterloo.ca/~shallit
In this talk, I’ll be speaking about *words*.

A word is a (possibly) empty string of symbols chosen from a finite nonempty alphabet $\Sigma$.

$\Sigma^*$ is the set of all finite words.

$\epsilon$ is the empty word.

$|x|$ denotes the length of the word $x$, and $|x|_a$ is the number of occurrences of the symbol $a$ in $x$.

$x^k$ denotes the product $\underbrace{xxx \cdots x}_k$.

$w^\omega$ is the infinite word $www \cdots$.

If $S$ is a set of words, then $S^\omega$ is the set of all infinite words constructed by concatenating elements of $S$. 
Theorem

Let $x, y$ be nonempty words. Then the following three conditions are equivalent:

(1) $xy = yx$;

(2) There exist a nonempty word $z$ and integers $k, \ell > 0$ such that $x = z^k$ and $y = z^\ell$;

(3) There exist integers $i, j > 0$ such that $x^i = y^j$.

However, note that in the implication (1) $\implies$ (2), an even weaker hypothesis suffices: we only need that $xy$ agrees with $yx$ on the first $|x| + |y| - \gcd(|x|, |y|)$ symbols.
We say an infinite sequence \((f_n)_{n \geq 0}\) is periodic with period length \(h \geq 1\) if \(f_n = f_{n+h}\) for all \(n \geq 0\). The following is a classical “folk theorem”:

**Theorem.** If \((f_n)_{n \geq 0}\) is an infinite sequence that is periodic with period lengths \(h\) and \(k\), then it is periodic with period length \(\gcd(h, k)\).

**Proof.** By the extended Euclidean algorithm, there exist integers \(r, s \geq 0\) such that \(rh - sk = \gcd(h, k)\). Then we have

\[
f_n = f_{n+rh} = f_{n+rh-sk} = f_{n+\gcd(h,k)}
\]

for all \(n \geq 0\). □
The Fine-Wilf Paper

- N. J. Fine and H. S. Wilf, “Uniqueness theorems for periodic functions”


- The Fine-Wilf theorem: a version of the periodicity theorem for finite sequences.

- Answers the question: how long must the finite sequence \((f_n)_{0 \leq n < D}\) be for period lengths \(h\) and \(k\) to imply a period of length \(\gcd(h, k)\)?

- \(D = \lcm(h, k)\) works (of course!), but Fine and Wilf proved we can take \(D = h + k - \gcd(h, k)\).
Figure: Citations of Fine-Wilf, according to Web of Science
The Fine-Wilf Theorems

**Theorem 1.** Let \((f_n)_{n \geq 0}\) and \((g_n)_{n \geq 0}\) be two periodic sequences of period \(h\) and \(k\), respectively. If \(f_n = g_n\) for \(h + k - \text{gcd}(h, k)\) consecutive integers \(n\), then \(f_n = g_n\) for all \(n\). The result would be false if \(h + k - \text{gcd}(h, k)\) were replaced by any smaller number.

**Theorem 2.** Let \(f(x), g(x)\) be continuous periodic functions of periods \(\alpha\) and \(\beta\), respectively, where \(\alpha/\beta = p/q\), \(\text{gcd}(p, q) = 1\). If \(f(x) = g(x)\) on an interval of length \(\alpha + \beta - \beta/q\), then \(f = g\). The result would be false if \(\alpha + \beta - \beta/q\) were replaced by any smaller number.

**Theorem 3.** Let \(f(x), g(x)\) be continuous periodic functions of periods \(\alpha\) and \(\beta\), respectively, where \(\alpha/\beta\) is irrational. If \(f(x) = g(x)\) on an interval of length \(\alpha + \beta\), then \(f = g\). The result would be false if \(\alpha + \beta\) were replaced by any smaller number.
Theorem

Let $w$ and $x$ be nonempty words. Let $y \in w \{w, x\}^\omega$ and $z \in x \{w, x\}^\omega$. Then the following conditions are equivalent:

(a) $y$ and $z$ agree on a prefix of length at least $|w| + |x| - \gcd(|w|, |x|);
(b) $wx = xw$;
(c) $y = z$.

Proof.

(c) $\implies$ (a): Trivial.

(b) $\implies$ (c): By Lyndon-Schützenberger.

We'll prove (a) $\implies$ (b).
Proof.

(a) \( y \in w\{w, x\}^\omega \) and \( z \in x\{w, x\}^\omega \) agree on a prefix of length at least \(|w| + |x| - \gcd(|w|, |x|) \) \( \implies \) (b) \( wx = xw \):

We prove the contrapositive. Suppose \( wx \neq xw \).

Then we prove that \( y \) and \( z \) differ at a position \( \leq |w| + |x| - \gcd(|w|, |x|) \).

The proof is by induction on \(|w| + |x|\).

Case 1: \(|w| = |x|\) (which includes the base case \(|w| + |x| = 2\)). Then \( y \) and \( z \) must disagree at the \(|w|'th\) position or earlier, for otherwise \( w = x \) and \( wx = xw \); since \(|w| \leq |w| + |x| - \gcd(|w|, |x|) = |w|\), the result follows.
Case 2: $|w| < |x|$.  

If $w$ is not a prefix of $x$, then $y$ and $z$ disagree on the $|w|$'th position or earlier, and again $|w| \leq |w| + |x| - \gcd(|w|, |x|)$.  

So $w$ is a proper prefix of $x$.  

Write $x = wt$ for some nonempty word $t$.  

Now any common divisor of $|w|$ and $|x|$ must also divide $|x| - |w| = |t|$, and similarly any common divisor of both $|w|$ and $|t|$ must also divide $|w| + |t| = |x|$. So $\gcd(|w|, |x|) = \gcd(|w|, |t|)$.  

Now $wt \neq tw$, for otherwise we have $wx = wwt = wtw = xw$, a contradiction.

Then $y = ww \cdots$ and $z = wt \cdots$. By induction (since $|wt| < |wx|$), $w^{-1}y$ and $w^{-1}z$ disagree at position $|w| + |t| - \gcd(|w|, |t|)$ or earlier.

Hence $y$ and $z$ disagree at position $2|w| + |t| - \gcd(|w|, |t|) = |w| + |x| - \gcd(|w|, |x|)$ or earlier. ■
The proof also implies a way to get words that optimally “almost commute”, in the sense that $xw$ and $wx$ should agree on as long a segment as possible.

**Theorem**

*For each $m, n \geq 1$ there exist binary words $x, w$ of length $m, n$, respectively, such that $xw$ and $wx$ agree on a prefix of length $m + n - \gcd(m, n) - 1$ but differ at position $m + n - \gcd(m, n)$.*

These words are the finite *Sturmian words*.

Indeed, our proof even provides an algorithm for computing these words:

$$S(h, k) = \begin{cases} 
(0^h, 0^{h-1}1), & \text{if } h = k; \\
(x, w), & \text{if } h > k \text{ and } S(k, h) = (w, x); \\
(w, wt), & \text{if } h < k \text{ and } S(h, k-h) = (w, t). 
\end{cases}$$
Since 1965, research on Fine-Wilf has been in three areas:

- applications (esp. to string-searching algorithms such as Knuth-Morris-Pratt)
- generalizations (esp. to more than 2 numbers; partial words)
- variations (e.g., to abelian periods; to inequalities)
The famous linear-time string searching algorithm of Knuth-Morris-Pratt finds all occurrences of a pattern $p$ in a text $t$ in time bounded by $O(|p| + |t|)$.

It compares the pattern to a portion of the text beginning at position $i$, and, when a mismatch is found, shifts the pattern to the right based on the position of the mismatch.

The worst-case in their algorithm comes from “almost-periodic” words, where long sequences of matching characters occur without a complete match.

It turns out that such words are precisely the maximal “counterexamples” in the Fine-Wilf theorem (the Sturmian pairs).
Many authors have worked on generalizations to multiple periods: Castelli, Justin, Mignosi, Restivo, Holub, Simpson & Tijdeman, Constantinescu & Ilie, Tijdeman & Zamboni, ...

For example, Castelli, Mignosi, and Restivo (1999) proved that for three periods $p_1 \leq p_2 \leq p_3$ the appropriate bound is

$$\frac{1}{2}(p_1 + p_2 + p_3 - 2 \gcd(p_1, p_2, p_3) + h(p_1, p_2, p_3))$$

where $h$ is a function related to the Euclidean algorithm on three inputs.
Partial words: words together with “don’t care” symbols called “holes”. Holes match each other and all other symbols.

**Theorem**

There exists a computable function $L(h, p, q)$ such that if a word $w$ with $h$ holes with periods $p$ and $q$ is of length $\geq L(h, p, q)$, then $w$ also has period $\gcd(p, q)$.

Berstel and Boasson (1999) proved we can take $L(1, p, q) = p + q$.

Shur and Konovalova (2004) proved we can take $L(2, p, q) = 2p + q - \gcd(p, q)$.

Many results by Blanchet-Sadri and co-authors.
Fine & Wilf works for equalities. How about inequalities?

For example, suppose \( f = (f_n)_{n \geq 0}, g = (g_n)_{n \geq 0} \) are two periodic sequences of period \( h \) and \( k \), respectively. Suppose \( f_n \leq g_n \) for a prefix of length \( D \). We want to conclude that \( f_n \leq g_n \) everywhere.

Here the correct bound is \( D = \text{lcm}(h, k) \). Example: take

\[
\begin{align*}
f &= (1^{h-1}2)\omega \\
g &= (2^{k-1}1)\omega
\end{align*}
\]

Then \( f_n \leq g_n \) for \( 0 \leq n < \text{lcm}(h, k) - 1 \), but the inequality fails at \( n = \text{lcm}(h, k) - 1 \).

So we need some additional hypothesis.
Theorem. Let \( f = (f_n)_{n \geq 0}, \ g = (g_n)_{n \geq 0} \) be two periodic sequences of real numbers, of period lengths \( h \) and \( k \), respectively, such that

\[
\sum_{0 \leq i < h} f_i \geq 0
\]

and

\[
\sum_{0 \leq j < k} g_j \leq 0.
\]

Let \( d = \gcd(h, k) \).

(a) If

\[
f_n \leq g_n \quad \text{for } 0 \leq n < h + k - d
\]

then \( f_n = g_n \) for all \( n \geq 0 \).

(b) The conclusion (a) would be false if in the hypothesis \( h + k - d \) were replaced by any smaller integer.
Define

\[ P(z) = 1 + z + \cdots + z^{h-1} = \frac{(z^h - 1)}{(z - 1)}; \]
\[ Q(z) = 1 + z + \cdots + z^{k-1} = \frac{(z^k - 1)}{(z - 1)}; \]
\[ R(z) = \frac{(z^k - 1)}{(z^d - 1)}; \quad d = \gcd(h, k) \]
\[ S(z) = \frac{(z^h - 1)}{(z^d - 1)}. \]

By hypothesis \( P \circ f \geq 0 \), where by \( \circ \) we mean take the dot product of the coefficients of \( P \) with consecutive overlapping windows of \( f \). Then \( R \circ (P \circ f) \geq 0 \). But then \( RP \circ f \geq 0 \).
Similarly, the hypothesis
\[ \sum_{0 \leq j < k} g_j \leq 0 \]
means \( Q \circ (-g) \geq 0 \).
Then \( SQ \circ (-g) \geq 0 \).
But \( RP = SQ \), so
\[ \sum_{0 \leq i < h+k-d} e_i(f_i - g_i) \geq 0. \quad (4) \]
where \( R(z)P(z) = \sum_{0 \leq i < h+k-d} e_i z^i \).

It can be shown that the \( e_i \) are strictly positive, so since \( f_n \leq g_n \)
for \( 0 \leq n < h + k - d \), we get \( f_n = g_n \) for \( 0 \leq n < h + k - d \).

By the Fine & Wilf theorem, \( f_n = g_n \) for \( n \geq 0 \). ■
Maximal counter-examples in (b) can be deduced as the *first differences* of the maximal counter-examples to Fine & Wilf (the Sturmian pairs).

For example, for $h = 5$, $k = 8$ we have $w = (-1, 1, -1, 0, 1)$ and $x = (0, 1, -1, 0, 1, -1, 1, -1)$. Then

<table>
<thead>
<tr>
<th>$n$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_n$</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$g_n$</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Suppose we have two periodic sequences of integers, say \((f_n)_{n \geq 0}\) of period \(h\) and \((g_n)_{n \geq 0}\) of period \(k\). For how many consecutive terms can \(f_n + g_n\) strictly decrease?

The answer, once again, is

\[ h + k - \gcd(h, k). \]

Here is an example achieving \(h + k - 1\) for \(h = 5\), \(k = 8\):

<table>
<thead>
<tr>
<th>(n)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>(f(n))</td>
<td>0</td>
<td>-16</td>
<td>8</td>
<td>-8</td>
<td>-24</td>
<td>0</td>
<td>-16</td>
<td>8</td>
<td>-8</td>
<td>-24</td>
<td>0</td>
<td>-16</td>
<td>8</td>
</tr>
<tr>
<td>(g(n))</td>
<td>0</td>
<td>15</td>
<td>-10</td>
<td>5</td>
<td>20</td>
<td>-5</td>
<td>10</td>
<td>-15</td>
<td>0</td>
<td>15</td>
<td>-10</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>(f + g)</td>
<td>0</td>
<td>-1</td>
<td>-2</td>
<td>-3</td>
<td>-4</td>
<td>-5</td>
<td>-6</td>
<td>-7</td>
<td>-8</td>
<td>-9</td>
<td>-10</td>
<td>-11</td>
<td>28</td>
</tr>
</tbody>
</table>
A morphism is a map $h$ from $\Sigma^*$ to $\Delta^*$ such that

$$h(xy) = h(x)h(y)$$

for all words $x, y$.

It follows that $h$ can be uniquely specified by providing its image on each letter of $\Sigma$.

For example, let

$$h(0) = r$$
$$h(1) = em$$
$$h(2) = b$$
$$h(3) = er$$

Then

$$h(011233) = rememberer.$$
If $\Sigma = \Delta$ we can iterate $h$. We write

\[
\begin{align*}
    h^2(x) & \quad \text{for} \quad h(h(x)), \\
    h^3(x) & \quad \text{for} \quad h(h(h(x))), \\
\text{etc.}
\end{align*}
\]
Iterated morphisms appear in many different areas (often under the name L-systems), including:

- models of plant growth in mathematical biology
- computer graphics
- infinite words avoiding certain patterns
For example, consider the map $\varphi$ defined by

\[
\begin{align*}
\varphi(a_r) &= a_l b_r \\
\varphi(a_l) &= b_l a_r \\
\varphi(b_r) &= a_r \\
\varphi(b_l) &= a_l
\end{align*}
\]

Iterating $\varphi$ on $a_r$ gives

\[
\begin{align*}
\varphi^0(a_r) &= a_r \\
\varphi^1(a_r) &= a_l b_r \\
\varphi^2(a_r) &= b_l a_r a_r \\
\varphi^3(a_r) &= a_l a_l b_r a_l b_r \\
\vdots
\end{align*}
\]

Here the $a$'s represent fat cells and the $b$'s represent thin cells.

This models the development of the blue-green bacterium *Anabaena catenula*. 
Szilard and Quinton [1979] observed that many interesting pictures, including approximations to fractals, could be coded using iterated morphisms.

A beautiful book by Prusinkiewicz and Lindenmayer provides many examples.
Example: code a picture using “turtle graphics” where $R$ codes a move followed by a right turn, $L$ codes a move followed by a left turn, and $S$ codes a move straight ahead with no turn.

Consider the map $g$ defined as follows:

\[
\begin{align*}
    g(R) &= RLLSRRRLR \\
    g(L) &= RLLSRRLL \\
    g(S) &= RLLSRRLS
\end{align*}
\]

By iterating $g$ on $RRRR$ we get

\[
\begin{align*}
    g^0(R) &= RRRR \\
    g^1(R) &= RLLSRRLLSRRRLRRLLS \ldots
\end{align*}
\]

These words code successive approximations to a von Koch fractal curve.
Figure: Four iterations in the construction of the von Koch curve
The Matrix Associated with a Morphism

Given a morphism \( \varphi : \Sigma^* \to \Sigma^* \) for some finite set \( \Sigma = \{a_1, a_2, \ldots, a_d\} \), we define the \textit{incidence matrix} \( M = M(\varphi) \) as follows:

\[
M = (m_{i,j})_{1 \leq i, j \leq d}
\]

where \( m_{i,j} \) is the number of occurrences of \( a_i \) in \( \varphi(a_j) \), i.e.,

\[
m_{i,j} = |\varphi(a_j)|_{a_i}.
\]

Example. Consider the morphism \( \varphi \) defined by

\[
\varphi : a \to ab, \quad b \to cc \quad c \to bb.
\]

Then

\[
M(\varphi) = \begin{pmatrix}
a & b & c \\
a & 1 & 0 & 0 \\
b & 1 & 0 & 2 \\
c & 0 & 2 & 0 \\
\end{pmatrix}
\]
The matrix $M(\varphi)$ is useful because of the following proposition.

**Proposition.** We have

$$
\begin{bmatrix}
|\varphi(w)|_{a_1} \\
|\varphi(w)|_{a_2} \\
\vdots \\
|\varphi(w)|_{a_d}
\end{bmatrix} =
M(\varphi)
\begin{bmatrix}
|w|_{a_1} \\
|w|_{a_2} \\
\vdots \\
|w|_{a_d}
\end{bmatrix}.
$$

**Proof.** We have

$$
|\varphi(w)|_{a_i} = \sum_{1 \leq j \leq d} |\varphi(a_j)|_{a_i} |w|_{a_j}.
$$
The Matrix Associated with a Morphism

Corollary.

\[
\begin{bmatrix}
\varphi^n(w)|_{a_1} \\
\varphi^n(w)|_{a_2} \\
\vdots \\
\varphi^n(w)|_{a_d}
\end{bmatrix}
\begin{bmatrix}
w|_{a_1} \\
w|_{a_2} \\
\vdots \\
w|_{a_d}
\end{bmatrix}
= (M(\varphi))^n
\begin{bmatrix}
w|_{a_1} \\
w|_{a_2} \\
\vdots \\
w|_{a_d}
\end{bmatrix}
\]
Hence we find

**Corollary.**

\[
|\varphi^n(w)| = \left[ \begin{array}{ccc} 1 & 1 & 1 \\ & \ddots & \end{array} \right] M(\varphi)^n \left[ \begin{array}{c} \vert w \vert_{a_1} \\ \vert w \vert_{a_2} \\ \vdots \\ \vert w \vert_{a_d} \end{array} \right].
\]
We can now ask questions about the sequence of lengths

\[ |x|, \ |h(x)|, \ |h^2(x)|, \ldots \]

These questions were very popular in mathematical biology (L-systems) in the 1980’s.

For example, here is a classical result:

**Theorem.** Suppose \( h : \Sigma^* \rightarrow \Sigma^* \) is a morphism, and suppose there exist a word \( w \in \Sigma^* \) and a constant \( c \) such that

\[ c = |w| = |h(w)| = \cdots = |h^n(w)|, \]

where \( n = |\Sigma| \). Then \( c = |h^i(w)| \) for all \( i \geq 0 \).
Proof of the Theorem

It suffices to show \(|h^{n+1}(w)| = c\), because then the theorem follows by induction on \(n\).

Let \(M\) be the incidence matrix of \(h\). By the Cayley-Hamilton theorem,

\[ M^n = c_0 M^0 + \cdots + c_{n-1} M^{n-1} \]

for some constants \(c_0, c_1, \ldots, c_{n-1}\).

Define \(f_i = |h^i(w)|\) and let

\[ v = [|w|_{a_1} |w|_{a_2} \cdots |w|_{a_n}]^T. \]

Then for \(0 \leq i < n\) we have

\[ f_{i+1} - f_i = [1 1 \cdots 1](M^{i+1} - M^i)v \]
\[ = [1 1 \cdots 1]M^i(M - I)v \]
\[ = [1 1 \cdots 1]M^i v' = 0, \]

where \(v' := (M - I)v\).
Proof of the Theorem

Now

\[ f_{n+1} - f_n = [1 1 \cdots 1] M^n v' \]

\[ = [1 1 \cdots 1] (c_0 + \cdots + c_{n-1} M^{n-1}) v' \]

\[ = \sum_{0 \leq i < n} c_i [1 1 \cdots 1] M^i v' \]

\[ = 0, \]

since each summand is 0.

Hence \( f_{n+1} = f_n \). \qed
Another Question

We might also ask, how long can the sequence of lengths

$$|x|, |h(x)|, |h^2(x)|, \ldots$$

strictly decrease?

This question arose naturally in a paper with Wang characterizing the two-sided infinite fixed points of morphisms, i.e., those two-sided infinite words $w$ such that $h(w) = w$. 
If $\Sigma$ has $n$ elements, we can easily find a decreasing sequence of length $n$. For example, for $n = 5$, define $h$ as follows:

\[
\begin{align*}
  h(a) &= b \\
  h(b) &= c \\
  h(c) &= d \\
  h(d) &= e \\
  h(e) &= \varepsilon
\end{align*}
\]

Then we have

\[
\begin{align*}
  h(abcde) &= bcde \\
  h^2(abcde) &= cde \\
  h^3(abcde) &= de \\
  h^4(abcde) &= e \\
  h^5(abcde) &= \varepsilon
\end{align*}
\]
So

\[|\text{abcde}| > |h(\text{abcde})| > |h^2(\text{abcde})| > |h^3(\text{abcde})| > |h^4(\text{abcde})| > |h^5(\text{abcde})| = 0.\]
**Conjecture.** If \( h : \Sigma^* \rightarrow \Sigma^* \), and \( \Sigma \) has \( n \) elements, then

\[
|w| > |h(w)| > \cdots > |h^k(w)|
\]

implies that \( k \leq n \).

Another way to state the Decreasing Length Conjecture is the following:

**Conjecture.** Let \( M \) be an \( n \times n \) matrix with non-negative integer entries. Let \( v \) be a column vector of non-negative integers, and let \( u \) be the row vector \([1 \ 1 \ 1 \ \cdots \ 1]\). If

\[
uv > uMv > uM^2v > \cdots > uM^k v
\]

then \( k \leq n \).
There is a nice correspondence between directed graphs and non-negative matrices, as follows:

If $G$ is a directed graph on $n$ vertices, we can construct a non-negative matrix

$$M(G) = (m_{i,j})_{1 \leq i,j \leq n}$$

as follows: let

$$m_{i,j} = \begin{cases} 
1, & \text{if there is a directed edge from vertex } i \text{ to vertex } j \text{ in } G; \\
0, & \text{otherwise.} 
\end{cases}$$

Then the number of distinct walks of length $n$ from vertex $i$ to vertex $j$ in $G$ is just the $i,j$'th entry of $M^n$. 

Path Algebra
Similarly, given a non-negative $n \times n$ matrix $M = (m_{i,j})_{1 \leq i, j \leq n}$ we may form its associated graph $G(M)$ on $n$ vertices, where we put a directed edge from vertex $i$ to vertex $j$ iff $m_{i,j} > 0$. 
Lemma. Let $r \geq 1$ be an integer, and suppose there exist $r$ sequences of real numbers $b_i = (b_i(n))_{n \geq 0}$, $1 \leq i \leq r$, and $r$ positive integers $h_1, h_2, \ldots, h_r$, such that the following conditions hold:

(a) $b_i(n + h_i) \geq b_i(n)$ for $1 \leq i \leq r$ and $n \geq 0$;

(b) There exists an integer $D \geq 1$ such that
\[
\sum_{1 \leq i \leq r} b_i(n) > \sum_{1 \leq i \leq r} b_i(n + 1)
\]
for $0 \leq n < D$.

Then $D \leq h_1 + h_2 + \cdots + h_r - r$. 

**Remark.** When $r = 2$ and $\gcd(h_1, h_2) = 1$, then it can be shown that the bound in this Lemma is tight.

For example, for $h_1 = 5$, $h_2 = 8$ we find

<table>
<thead>
<tr>
<th>$n$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1(n)$</td>
<td>0</td>
<td>-16</td>
<td>8</td>
<td>-8</td>
<td>-24</td>
<td>0</td>
<td>-16</td>
<td>8</td>
<td>-8</td>
<td>-24</td>
<td>0</td>
<td>-16</td>
<td>8</td>
</tr>
<tr>
<td>$b_2(n)$</td>
<td>0</td>
<td>15</td>
<td>-10</td>
<td>5</td>
<td>20</td>
<td>-5</td>
<td>10</td>
<td>-15</td>
<td>0</td>
<td>15</td>
<td>-10</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>$b_1(n) + b_2(n)$</td>
<td>0</td>
<td>-1</td>
<td>-2</td>
<td>-3</td>
<td>-4</td>
<td>-5</td>
<td>-6</td>
<td>-7</td>
<td>-8</td>
<td>-9</td>
<td>-10</td>
<td>-11</td>
<td>28</td>
</tr>
</tbody>
</table>
**Theorem.** Suppose $M$ is an $n \times n$ matrix with non-negative integer entries. If there exist a row vector $u$ and a column vector $v$ with non-negative integer entries such that

$$uv > uMv > uM^2v > \cdots > uM^k v,$$

then $k \leq n$. Also $k = n$ only if $M^n = 0$. 
Proof.

- Let $M$ be the matrix in the statement of the theorem and $G$ its associated graph.
- Let $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)^T$.
- Let $V$ be the set of vertices in $G$.
- Consider some maximal set of vertices forming disjoint cycles $\{C_1, C_2, \ldots, C_r\}$ in $G$.
- Then $V$ can be written as the disjoint union

$$V = C_1 \cup C_2 \cup \cdots \cup C_r \cup W,$$

where $W$ is the set of vertices which do not lie in any of the disjoint cycles.
Any directed walk in $G$ of length $|W|$ or greater must intersect some cycle $C_i$, for otherwise the walk would contain a cycle disjoint from $C_1, C_2, \ldots, C_r$.

Associate each walk of length $\geq |W|$ with the first cycle $C_i$ it intersects.

Define $P_{i,j,l}^s$ to be the number of directed walks of length $s$ from vertex $i$ to vertex $j$ associated with cycle $l$.

Also define

$$T_l^s := \sum_{1 \leq i, j \leq n} u_i v_j P_{i,j,l}^s.$$

Then for any $s \geq |W|$ we have

$$uM^s v = \sum_{1 \leq l \leq r} T_l^s. \quad (5)$$
Then

\[ T_i^s \le T_i^{s+|C_i|}, \]

since any walk of length \( s \) associated with cycle \( C_i \) can be extended to a walk of length \( s + |C_i| \) by traversing the cycle \( C_i \) once.

From the inequality \( uM^s v > uM^{s+1} v \) for \( 0 \le s \le k - 1 \) and Eq. (5) we have

\[
\sum_{1 \le l \le r} T_i^s > \sum_{1 \le l \le r} T_i^{s+1}
\]

for \( |W| \le s < k \).

Now for \( 1 \le i \le r \) and \( j \ge 0 \) define \( b_i(j) = T_i^{|W|+j} \) and \( h_i = |C_i| \).

Then the conditions of the previous Lemma are satisfied.
We conclude that

\[ k - |W| \leq |C_1| + |C_2| + \cdots + |C_r| - r. \]

Moreover

\[ |C_1| + |C_2| + \cdots + |C_r| + |W| = |V| = n \]

and so \( k \leq n - r \).

Finally \( k = n \) implies that \( r = 0 \), so \( G \) is acyclic and \( M^n = 0 \).

So the Decreasing Length Conjecture is proved.
