$N = \{0, 1, 2, ...\}$

Σ Alphabet, elemente Buchst., Wort Abbildungen $[1..n] \to \Sigma$.
Σ^n Worte d. Länge n, Σ^* Wort über Sigma, Σ^+ mit Länge >= 1.
$|u|$ Wortlänge, $|u|_a$ Anzahl a in U, $alph(u)$ Menge der Buchst. in u.
conc $(\Sigma^*, conc, \epsilon)$ Monoid, kommutativ nur bei Einbuchstabigen Alph., Kürzen: $uv = uw \to v = w; uv = wv \to u = w$

Alphabet so das Faktorisierung eindeutung, $\{a, b, ab\}$ nicht zulässig.

u=vxw v Präfix, w Suffix, x Faktor. Falls $\neq \epsilon$, dann echt.

uv=wx, dann (1)$|u| > |w|$ und $u = wy, x = yv$; (2) = und $u = w, v = x$; (3) $<$ und $w = uy, v = yx$; y eindeutig.

(im)primitiv imprimitv: $u = z^k, k \geq 2$; primitiv: sonst. **Wurzel** $\varrho(u)$ KÜRZESTE z, so dass $u = z^k$. EINDEUTIG f.a. u

kommutieren u,v komm., falls $uv = vu$

$uv = vu$ **gdw** $\varrho(u) = \varrho(v)$ folgt aus: $uv = vu \to u = z^p, v = z^q$
u,v primitiv, $u^p = v^q \to u = v$
$u^p = v^q \to \varrho(u) = \varrho(v)$
$\varrho(x^k) = \varrho(x)$

Fine und Wilf $u, v \in \sigma^+$, $u^p =_k v^q$ mit $k = |u| + |v| - ggT(|u|, |v|)$ und $=_k$ übereinstimmung auf ersten k Buchst. DANN: $\varrho(u) = \varrho(v)$. Grenze k scharf.
(Beweis: $r = |u|/ggt, s = |v|/ggt$, $u = x_1...x_r$, $v = x_1..x_s$; $l = h * r \ mod \ s$ zeige $x_l = x_r$ mit Ind. über h.)

$uv =_k vu \to uv = vu (\to \varrho(u) = \varrho(v))$

Zusammenfassung: $u, v \in \Sigma^+$,$k = |u| + |v| - ggT(|u|, |v|)$ dann:
$\exists_{p,q} u^p =_k v^q$ gdw $uv =_k vu$ gdw $uv = vu$ gdw $\varrho(u) = \varrho(v)$.

konjugiert $u \approx v$, falls $\exists_{x,y} u = xy$ und $v = yx$.
offensichtlich: zwei Wort konjugiert, falls zyklische Permutation voneinander.Äquivalenzrelation.

$|u| = |v| \to u \approx v$ gdw $\varrho(u) \approx \varrho(v)$

$|konj(u)| = |\varrho(u)|$

$u \approx v$ gdw $\exists_z zu = vz$.
zeige mit: $uv = vw$ gdw $\exists_{x,y,k} u = xy, w = yx, v = (xy)^k x$ (zeige erst: $uv = vw \to u^n v = vw^n$)

KMP Fehlerfkt.$f(k) := max(\{l | 0 < l < k$ mit $m_1..m_{l-1}$ ist Suffix von $m_1..m_{k-1}$ und $m_l \neq m_k\} \cup \{0\})$
Vergleiche der Reihe nach Text mit Muster (beginnend beim 1. Buchst.), falls Unterschied an Stelle i des Muster, verschiebe Muster so, dass der unterschieldiche Buchst. im Text mit dem f(i). Buchst. des Muster verglichen wird.

Fehlerfkt(i). berechnen: Pattern an Pos. i von Pattern vorbeischieben, so dass möglichst viele Positionen vor i übereinstimmen. j der Index des Buchst. des vorbeigeschobenen Patterns, der neben Buchst. i liegt.
$f(i) = f(j), falls m_i = m_j$; sonst$f(i) = j$.

Wurzel $|\varrho(u)| = |u| + 1 - f_{u\$}(n + 1)$, falls u imprimitiv. (Beweis: f >=: $\varrho(u)^{k-1}$ Präfix und Suffix von u., f<: Angenommen das, dann ... $\varrho(u) = z^{p+q}$, also nicht primitiv.)

Überlappungen u überlappend falls u=xy und u=yz bzw. uv=wu, also u hat Suffix das auch Präfix ist. |x|>0. OVL(u) Menge d. überlapp. OV(u) längste Überlapp.

längste Überlapp. best. $|OV(u)| f_{u\$}(n + 1) - 1$

linksüberlappung u ü. von links mit v, falls u=yx und v=xy, $|x| > 0$. LOVL(u,v) Menge der linksüberlappungen, lov(u,v) längste.

längste links-überlappung$lov(u) = f_{v\$u\#}(n + m + 2) - 1$.

unvermeidbare Eigenschaften P unvermeidbar, falls jedes hinreichend lange Wort Eigenschaft P hat.

unendl. Worte$u : N \to \Sigma$. $u_0 u_1 ... u_i ....$Präfix der Länge k $u^{[k]}$.

EigenschaftenP Eig., dann $L_P \subset \Sigma^*$Menge der Worte die Eig. P haben. faktorvererblich, falls für $x \in L_P$ alle Teilworte von x in L_P
P faktorvererblich $\to L_P$ unendlich gdw $\exists u \in \Sigma^*$, so dass jeder endliche Faktor in L_P (Konstruktion von u: Buchstaben der Reihe nach wählen, so das es immer noch unendl. viele Worte in L_p gibt die mit diesen Buchst. beginnen.)

Limes$w_0, w_1 ... w_i ...$ Folge von Worten mit w_i Präfix von w_{i+1}. Dann $u = lim_{i \to \infty} w_i$, falls jeder Präfix der Länge k von $u = w_k$.

$h^\omega(a_0)$, falls $h(a) \neq \epsilon$ und $\exists_{a_0} : h(a_0) = a_0 w$.
(i) $h^{n+1} = h^n(a_0) h^n(w) = a_0 h^0(w) h^1(w) h^2(w)..h^n(w)$.
(iii):$h^\omega(a_0) = lim_{n \to \infty} h^n(a_0) = a_0 w h(w) h^2(w)..h^i(w)..$ Fixpunkt von h.

Fibonacci-Wort$h(a) = ab, h(b) = a$. Dann $h^\omega(a) = lim_{n \to \infty} f_n$ mit $f_0 = a, f_1 = ab, f_{n+2} = f_{n+1} f_n$ Fibonacci-Wort.

unendl. wort von Thue-Morse$h(a) = ab, h(b) = ba$. DANN $t = h^\omega(a)$ unendl. W. v. Thue-Morse.
(i)$|h^n(a)| = |h^n(b)| = 2^n$ (ii) $h^{n+1}(a) = h^n(a) h^n(b)$ (iii) $h^n(a) = \bar{h^n}(b)$
(iv) u_{2n}, v_{2n} Palindrome und $u_{2n+1} = mirror(v_{2n+1})$

$t_n = a$, **falls n binär gerade Anzahl 1en enth.** (Beweis: $t_{2i} = t_i, \hat{t_i} = t_{2i+1}$ mit $t = h(t)$ (Fixpunkt). Dann Induktion, und $d_2(2m) = d_2(m)$ nutzen.)

kubikfrei w stark kubikfrei, falls w kein Faktor axaxa hat.
w quadratfrei, falls w keinen Faktor xx hat.

t stark kubikfrei(1)h vererbt kubikfreiheit (w kubikfrei, so h(w) kubikfrei). (Beweis: Widerspruch, h(w)=u cxcxc v, Fall u gerade -> cxcx gerade -> x ungerade (sonst x und cxc in $\{ab, ba\}^*$) -> w=rsst, h(r)=u,h(s)=cx,h(t)=cv...-> s und t beginnen mit c).

quadrat unverm. falls |Σ|=2, für alle Wort mit 4 Buchst.

quadrat vermeidbar falls |Σ|=3: Wort $m := f^\omega(a)$, f(a)=abc, f(b)=ac, f(c)=b.
g(a):=abb, g(b):=ab, g(c):=a. Dann h(g(m)) = g(f(m))=g(m). Also g(m) Fixp. von h der mit a beg., also g(m)=t.
m quadratfrei (Beweis: indirekt, m=uxxv, g(m)=g(u)g(x)g(x)g(v), g(w) beginnt mit a -> t nicht kubikfrei).

quadratfreie HMP h quadratfrei, falls (1)h nicht trivial (nicht immer ϵ).(2)w quadratfrei $\to$ h(w) quadratfrei.
Gilt, falls (1)h(w) quadratfrei für alle quadratfreien w mit $|w| <= 3$; (2) $\forall_{a,b}$ h(a) kein Faktor von h(b). (nicht notwendig!)

PatternV Variablenmenge, $p \in V^*$ Pattern,
$L(p) = \{h(p) | hnicht - loeschenderHMP\}$.
w enthält p, falls Faktor von w in L(p).

vermeidbarkeit vermeidbar auf Σ, falls unendlich. viel w nicht p enthalten.
unvermeidbar auf Σ, falls jedes hinreichend lange w p enthält.
vermeidbar: falls ∃Σ mit p vermeidbar auf Σ.
unvermeidbar: ∀Σ p unvermeidbar auf Σ
$\mu(p)$ vermeidbarkeitsindex: kleinste Zahl, mit p k-vermeidbar (k Betrag des kleinsten Alphabet, auf dem p vermeidbar ist).

xyx unvermeidbar(Beweis: jedes Wort w der länge $2 * |\Sigma| + 1$ enth. xyx, denn dann kommt mind. ein Buchst 3 mal vor.)

pyp unvermeidbar falls p unvermeidbar. (Beweis: l, so dass alle Worte mit dieser Länge p enthalten. Dann $|\Sigma|^l$ Worte in Σ^l. Also enthält ein Wort der Länge $(|\Sigma|^l + 1) * l$ p doppelt. Nun noch garantieren das y nicht leer)

Zimin-Worte $z_0 = \epsilon$. $z_{n+1} = z_n x_n z_n$, x_i Variablen. **alle Zimin-Worte sind unvermeidbar** (da pyp unvermeidbar.)

binäre Pattern $V = \{x, y\}; p \in V^*$ Dann: $\mu(p) \in \{2, 3, \infty\}$

Codes$A \circ B = \{xy | x \in A, y \in B\}, A^0 = \epsilon, A^{n+1} = A^n \circ A, A^+, A^*$.
$\subset \Sigma^*$ Code, wenn $\forall_{w \in A^*}$ w eindeutiges Produkt aus Worten in A. (Untermonoid $A^* \subset \Sigma^*$ frei von A erzeugt.)

Codierungsgraph$C(A) = (V, E, \Sigma, f)$Für jedes Wort einen Zykel von S nach S über Knoten, so dass jeder Buchst. eine Kante hat.
$w \in A^*$ gdw ∃ Kantenfolge von S nach S die mit w markiert ist. w eindeutig als Produkt darst. GDW es gibt genau eine Kantenfolge.

Testgraph$T(A) = (V_1 = V \times V, E_1, \Sigma, f_1)$ hat C(A) Knoten $p \neq q$ und Kanten $S \to^a p, S \to^a q$, dann füge Kante $S_0 \to^a (p, q)$ hinzu. Oder ist Kante zu (p,q) in T(A) vorhanden und in C(A) $p \to^a r, q \to^a t$, dann füge Kante $(p, q) \to^a (r, t)$ hinzu. (Es werden also je zwei Pfade in C(A) die gleich markiert sind, aber über unterschiedliche Knoten führen, in einen Pfad umgewandelt. Landet man irgendwann in dem Pfad in T(A) bei (S,S), so gibt es also zwei Pfade in C(A) die mit der gleichen Markierung von S nach S gehen).

A ist ein Code gdw (S,S) in T(A) nicht erreichbar (gdw keine Kante enthält S,S) $(O(\sum_{i=1}^n |x_i|)^2)$

Theorem HMP$A \subset \Sigma^+, \Gamma$ Alphabet. A Code gdw
$\forall_{h : \Gamma \to A, \text{bijektion}} h : \Gamma^* \to \Sigma^*$ injektiv.

A hat Verzögerung k, falls k kleinste Zahl mit: $\forall x_1, .., x_k, y_1, .., y_n \in A$ und $x_1..x_n$ Präfix von $y_1..y_n$ gilt $x_1 = y_1$.

unbeschränkte Verzögerung gdw T(A) enth. Zykel (denn dann gibt es $p \neq q$ in C(A), so dass von (p,q) in T(A) aus beliebig lange Pfade erzeugt werden können, bevor man zurück zu S gelangt.)

Abschätzung $A = \{x, y\}$ Code. Verz. k $\leq |x| + |y| - ggt(|x|, |y|)$.
(Beweis: falls kein Präfix-Code: $y = x^l u$, so das x kein Präfix von u. $w \in A^*$. $w = x \circ A^+$..xy Präfix v. w $w = y \circ A^+$, dann $w = x \circ x^l u | x \circ$ ..also yx Präfix von w. $xy \neq yx$. also auf ersten k stellen untersch.)

Präfix-Code$A = \{x_i\} \subset \Sigma^+$ Präfix-Code, falls kein echter Präfix eines Wortes x_i wieder in A. Hat Verzögerung 1

McMillan $A = \{x_1, .., x_n\}$ Code $\lambda_i = |x_i|$. Dann gilt $\sum_{i=1}^{n} |\Sigma|^{-\lambda_i} \leq 1$.
und dann gilt (auch für beliebige $\lambda_i, i = 1..n$): es gibt einen Präfix-Code $B = y_1..y_n$ mit $|y_i| = \lambda_i$.

Konstruktion des Präfix-Code: Wähle $\lambda = max_{\lambda_i}$. μ_j Anzahl λ_i mit Wert j. Dann $\mu_\lambda \leq r^\lambda - \mu_1 * r^{\lambda-1} - ... - \mu_{\lambda-1} * r^1$
...
$\mu_2 \leq r^2 - \mu_1 * r$
$\mu_1 \leq r$
Wähle also $B_2 \subset \Sigma$ mit $|B_1| = \mu_1$, $B_2 \subset \Sigma^2 - (B_1 \circ \Sigma)$ mit $|B_2| = \mu_2 ... B_\lambda$. Setze $B := \cup_{i=1}^{\lambda} B_u$