Sieve theory and small gaps between primes: A variational problem

Andrew V. Sutherland
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

(on behalf of D.H.J. Polymath)

Explicit Methods in Number Theory
MATHEMATISCHES FORSCHUNGSINSTITUT OBERWOLFACH

July 9, 2015

Explicitly proving bounded gaps

Recall that our goal is to prove upper bounds on

$$H_m := \liminf_{n\to\infty} p_{n+m} - p_n.$$

We do this by establishing DHL$[k, m+1]$:= "*every admissible k-tuple $\mathcal{H}$ has infinitely many translates $n+\mathcal{H}$ that contain at least $m+1$ primes*."

The diameter $h_k - h_1$ of any admissible k-tuple $\mathcal{H} = \{h_1, \ldots, h_k\}$ is then an upper bound on H_m, and we can take the minimal such diameter $H(k)$.

To prove DHL$[k, m+1]$ it suffices to find weights $w_n \in \mathbb{R}_{\geq 0}$ such that

$$\sum_{x<n\leq 2x} w_n \Big(\Theta(n+\mathcal{H}) - m\log(3x)\Big) > 0$$

for all sufficiently large x. Here $\Theta(n+\mathcal{H}) := \sum_{p=n+h_i \text{ prime}} \log p$.

Picking the weights w_n

In the Maynard-Tao approach, for $n \in (x, 2x]$ we use weights of the form

$$w_n := \Big(\sum_{\substack{d_i | n+h_i \\ \prod d_i < R}} \lambda_{d_1,\ldots,d_k} \Big)^2,$$

where $R := x^{\vartheta/2-\epsilon}$ depends on the level of distribution ϑ; any $\vartheta < \frac{1}{2}$ works, and we may take $\vartheta = \frac{1}{2} + \frac{\varpi}{2}$ if we can prove MPZ$[\varpi, \delta]$.

Let W_n be the product of the primes $p < \log\log\log x$, and define

$$\lambda_{d_1,\ldots,d_k} := \Big(\prod_i \mu(d_i) d_i\Big) \sum_{d_i | r_i, r_i \perp W_n} \frac{\mu(\prod_i r_i)^2}{\prod_i \phi(r_i)} F\left(\frac{\log r_1}{\log R}, \ldots, \frac{\log r_k}{\log R}\right),$$

for any nonzero square-integrable function $F\colon [0,1]^k \to \mathbb{R}$ with support in

$$\mathcal{R}_k := \{x_1, \ldots, x_k) \in [0,1]^k : \sum_i x_i \leq 1\}.$$

Maynard's theorem

Define

$$I(F) := \int_0^1 \cdots \int_0^1 F(t_1, \ldots, t_k)^2 dt_1 \ldots dt_k,$$

$$J(F) := \sum_{i=1}^{k} \int_0^1 \cdots \int_0^1 \left(\int_0^1 F(t_1, \ldots, t_k) dt_i \right)^2 dt_1 \ldots dt_{i-1} dt_{i+1} \ldots dt_k,$$

$$\rho(F) := \frac{J(F)}{I(F)}, \qquad M_k := \sup_F \rho(F)$$

Theorem (Maynard 2013)

For any $0 < \vartheta < 1$, *if* EH$[\vartheta]$ *and* $M_k > \frac{2m}{\vartheta}$, *then* DHL$[k, m+1]$.

We thus seek explicit bounds on M_k (and $H(k)$).
To prove DHL$[k, m+1]$ we need $M_k > 4m$ (or $M_k > 2m$ under EH).

Polymath Theorems

For $\alpha > 0$, define $M_k^{[\alpha]} := \sup_F \rho(F)$, with the supremum over nonzero square-integrable real-valued functions with support in $[0,\alpha]^k \cap \mathcal{R}_k$.

Theorem (D.H.J. Polymath 2014)

If $\mathsf{MPZ}[\varpi,\delta]$ *and* $M_k^{\left[\frac{\delta}{1/4+\varpi}\right]} > \frac{m}{1/4+\varpi}$ *then* $\mathsf{DHL}[k,m+1]$*.*

For $\epsilon \in (0,1)$ and $F\colon [0,1+\epsilon)^k \to \mathbb{R}$ with support in $(1+\epsilon)\mathcal{R}_k$, define

$$J_{1-\epsilon}(F) := \sum_{i=1}^{k} \int_{(1-\epsilon)\mathcal{R}_{k-1}^{(i)}} \left(\int_0^{1+\epsilon} F^2 dt_i \right)^2, \qquad M_{k,\epsilon} := \sup_F \frac{J_{1-\epsilon}(F)}{I(F)}.$$

Theorem (D.H.J. Polymath 2014)

Assume either $\mathsf{EH}[\vartheta]$ *with* $1+\epsilon < \frac{1}{\theta}$ *or* $\mathsf{GEH}[\vartheta]$ *with* $\epsilon < \frac{1}{k-1}$*.*
Then $M_{k,\epsilon} > \frac{2m}{\theta}$ *implies* $\mathsf{DHL}[k,m+1]$*.*

Cauchy-Schwarz bound

Suppose we can construct functions $G_i \colon \mathcal{R}_k \to \mathbb{R}_{>0}$, for $1 \le i \le k$, such that

$$\int_0^1 G_i(t_i, \ldots, t_k) dt_i \le 1$$

for all $(t_1, \ldots, t_k) \in [0,1]^k$ (extend G_i to $[0,1]^k$ by zero).

By Cauchy-Schwarz, for any $F \in L^2(\mathcal{R}^k)$ and each i, we have

$$\left(\int_0^1 F(t_1, \ldots, t_k) dt_i \right)^2 \le \int_0^1 F(t_1, \ldots, t_k)^2 dt_i \le \int_0^1 \frac{F(t_1, \ldots, t_k)^2}{G_i(t_1, \ldots, t_k)} dt_i.$$

Thus for $F \ne 0$ we have

$$\rho(F) = \frac{J(F)}{I(F)} \le \frac{\sum_i \int (F^2 / G_i)}{\int F^2} \le \sup_{\mathcal{R}_k} \sum \frac{1}{G_i(t_i, \ldots, t_k)}.$$

The RHS is an upper bound on $M_k = \sup \rho(F)$.

Computing M_k with eigenfunctions

Lemma

*If there exists a **strictly positive** $F \in L^2(\mathcal{R}_k)$ satisfying*

$$\lambda F(t_1, \ldots, t_k) = \sum_{i=1}^{k} \int_0^1 F(t_1, \ldots, t_{i-1}, t, t_{i+1}, \ldots t_k) dt$$

for some fixed $\lambda > 0$ and all $(t_1, \ldots, t_k)$ in $\mathcal{R}_k$, then $M_k = \lambda$.

Proof: Integrating both sides against F yields

$$\lambda I(F) = J(F),$$

so $M_k = \sup J(F)/I(F) \geq \lambda$. On the other hand, if we put

$$G_i(t_1, \ldots, t_k) := \frac{F(t_1, \ldots, t_k)}{\int_0^1 F(t_1, \ldots, t_{i-1}, t, t_{i+1}, \ldots, t_k) dt},$$

then $\sup_{\mathcal{R}_k} \sum_i \frac{1}{G_i(t_1, \ldots, t_k)} = \lambda \geq M_k$. $\square$

Computation of M_2

Recall the Lambert function $W\colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, defined by $W(x)e^{W(x)} = x$.
Let $\lambda := \frac{1}{1-W(1/e)}$ and define $f\colon [0,1] \to \mathbb{R}_{\geq 0}$ by

$$f(x) := \frac{1}{\lambda - 1 + x} + \frac{1}{2\lambda - 1} \log \frac{\lambda - x}{\lambda - 1 + x}.$$

One finds that for any $x \in [0,1]$ we have

$$\int_0^{1-x} f(t)dt = (\lambda - 1 + x)f(x).$$

Now define $F\colon \mathcal{R}_2 \to \mathbb{R}_{>0}$ by $F(x,y) := f(x) + f(y)$. For all $(x,y) \in \mathcal{R}_2$,

$$\begin{aligned} \int_0^1 F(t,y)dt + \int_0^1 F(x,t)dt &= \int_0^{1-y} F(t,y)dt + \int_0^{1-x} F(x,t)dt \\ &= \lambda f(y) + \lambda f(x) = \lambda F(x,y). \end{aligned}$$

Therefore $M_2 = \lambda = 1.38593\ldots$, by the lemma.

An upper bound on M_k

Lemma

$M_k \leq \frac{k}{k-1} \log k$ *for all* $k \geq 2$.

Proof: Define $G_i : \mathcal{R}_k \to R_{>0}$ by

$$G_i(t_1, \ldots, t_k) := \frac{k-1}{\log k} \cdot \frac{1}{1 - t_1 - \cdots - t_k + kt_i}$$

Then $\int_0^1 G_i(t_1, \ldots, t_k) dt_i \leq 1$, and $\sum_i \frac{1}{G_i(t_1,\ldots,t_k)} = \frac{k}{k-1} \log k \geq M_k$. $\square$

One can extend this argument to show $M_{k,\epsilon} \leq \frac{k}{k-1} \log(2k-1)$.

This implies $M_4 < 2$, so $M_5 \geq 2$ (proved by Maynard) is best possible.
And $M_{50} < 4$, which means the ϵ-trick was necessary to get $H_1 \leq 246$;
for $k > 50$ every admissible k-tuple has diameter at least $H(51) = 252$.

A lower bound on M_k

Maynard proves $M_k \geq \log k - 2\log\log k - 2$ for $k \gg 1$ using $F \in L^2(\mathcal{R}_k)$,

$$F(t_1, \ldots, t_k) := g(t_1)\cdots g(t_k),$$

where $g\colon [0,T] \to \mathbb{R}$ has the form $g(t) = \frac{1}{c+dt}$, for some $c, d, T > 0$.

We refine this approach by introducing an additional parameter $\tau > 0$ that allows us to replace the $\log\log k$ term with a small constant. Explicitly, let

$$g(t) := \frac{1}{c + (k-1)t},$$

and define

$$m_2 := \int_0^T g(t)^2\,dt, \quad \mu := \frac{1}{m_2}\int_0^T tg(t)^2\,dt, \quad \sigma^2 := \frac{1}{m_2}\int_0^T t^2 g(t)^2\,dt \,-\, \mu^2.$$

We require τ and T to satisfy

$$k\mu \leq 1 - \tau, \qquad k\mu < 1 - T, \qquad k\sigma^2 < (1 + \tau - k\mu)^2.$$

A lower bound on M_k

Theorem (D.H.J. Polymath 2014)

For $k \geq 2$ and $c, T, \tau > 0$ satisfying the inequalities above, we have

$$M_k \geq \frac{k}{k-1} \log k - E(k, c, \tau, T),$$

where $E(k, c, \tau, T)$ is an explicit function that is bounded as $k \to \infty$ for suitably chosen c, T, τ. Suitable choices include

$$c := \frac{1}{\log k} - \frac{1}{\log^2 k}, \qquad T := \frac{1}{\log k}, \qquad \tau := \frac{1}{\log k}.$$

For any $\alpha \geq T$ this bound also applies to $M_k^{[\alpha]}$.

For the k of interest we can generally keep $E(k, c, \tau, T) < 3$ by choosing

$$c := \frac{a}{\log k}, \qquad T := \frac{b}{\log k}, \qquad \tau := 1 - k\mu.$$

with $a \approx 1$ and b slightly less than 1.

Explicit lower bounds on M_k for large k

Lower bounds on M_k and $M_k^{[T]}$ given by the theorem with $E := E(k, c, t, T)$ determined by k and the parameters a, b as above.

k	a	b	E	$\frac{k}{k-1}\log k - E$	result
5511	0.965 000	0.973 000	2.616	6.000 048 609	DHL$[k, 4]^*$
35 410	0.994 790	0.852 130	2.645	7.829 849 259	DHL$[k, 3]$
41 588	0.978 780	0.943 190	2.636	8.000 001 401	DHL$[k, 5]^*$
309 661	0.986 270	0.920 910	2.643	10.000 000 320	DHL$[k, 6]^*$
1 649 821	1.004 220	0.801 480	2.659	11.657 525 560	DHL$[k, 4]$
75 845 707	1.007 120	0.770 030	2.663	15.481 250 900	DHL$[k, 5]$
3 473 955 908	1.007 932	0.749 093	2.665	19.303 748 720	DHL$[k, 6]$

The starred DHL$[k, m+1]^*$ use $M_k \geq 2m$ and are conditional on EH.
The unstarred DHL$[k, m+1]$ are unconditional via MPZ$[\varpi, \delta]$ using

$$M_k^{[T]} = M_k^{\left[\frac{\delta}{1/4+\varpi}\right]} > \frac{m}{1/4+\varpi},$$

with ϖ maximized subject to $600\varpi + 180\delta < 7$ with $\delta = T(\frac{1}{4} + \varpi)$.

Error term in lower bound on M_k

The error term $E(k,c,\tau,T)$ is the explicitly computable function

$$E(k,c,\tau,T) := \frac{k}{k-1}\frac{Z+Z_3+WX+VU}{(1+\tau/2)(1-\frac{k\sigma^2}{(1+\tau-k\mu)^2})},$$

$$Z := \frac{1}{\tau}\int_1^{1+\tau}\left(r\left(\log\frac{r-k\mu}{T}+\frac{k\sigma^2}{4(r-k\mu)^2\log\frac{r-k\mu}{T}}\right)+\frac{r^2}{4kT}\right)\,dr,$$

$$Z_3 := \frac{1}{m_2}\int_0^T kt\log\left(1+\frac{t}{T}\right)g(t)^2\,dt,$$

$$W := \frac{1}{m_2}\int_0^T \log\left(1+\frac{\tau}{kt}\right)g(t)^2\,dt,$$

$$X := \frac{\log k}{\tau}c^2,$$

$$V := \frac{c}{m_2}\int_0^T \frac{1}{2c+(k-1)t}g(t)^2\,dt,$$

$$U := \frac{\log k}{c}\int_0^1\left((1+u\tau-(k-1)\mu-c)^2+(k-1)\sigma^2\right)\,du.$$

Comparison with upper bounds

Lower and upper bounds on k needed to obtain DHL$[k, m+1]$ (or DHL$[k, m+1]^*$ under EH) implied by upper and lower bounds on M_k.

claim	$M_k^{[T]}$	$\min k$	$\max k$
DHL$[k, 2]$	4.000	51	$54^\dagger$
DHL$[k, 4]^*$	6.000	398	5511
DHL$[k, 3]$	7.830	2508	35 410
DHL$[k, 5]^*$	8.000	2973	41 588
DHL$[k, 6]^*$	10.000	22 017	309 661
DHL$[k, 4]$	11.658	115 601	1 649 821
DHL$[k, 5]$	15.481	5 288 246	75 845 707
DHL$[k, 6]$	19.304	241 891 521	3 473 955 908

† Obtained using explicitly constructed $F(t_1, \ldots, t_k) \neq g(t_1) \cdots g(t_k)$.

Lower bounds on M_k for small k

Lemma

$M_k := \sup \rho(F)$ *is unchanged by restricting to* ***symmetric*** $F \in L^2(\mathcal{R}_k)$.

We thus restrict our attention to functions

$$F = \sum_{i=1}^{n} a_i b_i$$

that are linear combinations of a fixed set of $\mathbb{R}$-linearly independent symmetric s $b_i \in L^2(\mathcal{R}_k)$. We wish to choose

$$\mathbf{a} := \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n$$

to maximize $\rho(F)$ over the real vector space spanned by $b_1, \ldots, b_n$.

Reduction to linear algebra

We thus fix $\mathbf{b} := (b_1, \ldots, b_n)$ with $b_i \in L^2(\mathcal{R}_k)$ linearly independent, and consider the real, symmetric, positive definite matrices

$$\mathbf{I} := \left[\int_{[0,1]^k} b_i(t_1, \ldots, t_k) b_j(t_1, \ldots, t_k) dt_1 \ldots dt_k \right]_{ij},$$

$$\mathbf{J} := \left[k \int_{[0,1]^{k+1}} b_i(t_1, \ldots, t_k) b_j(t_1, \ldots, t_{k-1}, t) dt_1 \ldots dt_k \, dt \right]_{ij},$$

For $F := \mathbf{a} \cdot \mathbf{b}$ we may compute

$$I(F) = \mathbf{a}^\intercal \mathbf{I} \mathbf{a}, \qquad J(F) = \mathbf{a}^\intercal \mathbf{J} \mathbf{a}, \qquad \rho(F) = \frac{\mathbf{a}^\intercal \mathbf{J} \mathbf{a}}{\mathbf{a}^\intercal \mathbf{I} \mathbf{a}}.$$

We may rescale $\mathbf{a}$ so that $I(F) = 1$ without changing $\rho(F)$.
We thus wish to maximize $\mathbf{a}^\intercal \mathbf{J} \mathbf{a}$ subject to $\mathbf{a}^\intercal \mathbf{I} \mathbf{a} = 1$.

Reduction to a generalized eigenvalue problem

To maximize $\mathbf{a}^\intercal\mathbf{J}\mathbf{a}$ subject to $\mathbf{a}^\intercal\mathbf{I}\mathbf{a} = 1$ we introduce a Lagrange multiplier λ. Let $f(\mathbf{a}) := \mathbf{a}^\intercal\mathbf{J}\mathbf{a}$ and $g(\mathbf{a}) := \mathbf{a}^\intercal\mathbf{I}\mathbf{a} - 1$. We require

$$\nabla f - \lambda \nabla g = 0.$$

Since $\mathbf{I}$ and $\mathbf{J}$ are symmetric, $\nabla f = 2\mathbf{J}\mathbf{a}$ and $\nabla g = 2\mathbf{I}\mathbf{a}$, we thus have

$$2(\mathbf{J} - \lambda\mathbf{I})\mathbf{a} = 0.$$

Equivalently (since $\mathbf{I}$ is invertible), $\mathbf{I}^{-1}\mathbf{J}\mathbf{a} = \lambda\mathbf{a}$. Thus λ is an eigenvalue of $\mathbf{I}^{-1}\mathbf{J}$ and $\mathbf{a}$ is a corresponding eigenvector (scaled to make $\mathbf{a}^\intercal\mathbf{I}\mathbf{a} = 1$).

Note that $\mathbf{J}\mathbf{a} = \lambda\mathbf{I}\mathbf{a}$ implies $\mathbf{a}^\intercal\mathbf{J}\mathbf{a} = \lambda\mathbf{a}^\intercal\mathbf{I}\mathbf{a} = \lambda$, so we want to maximize λ. We thus seek a maximal solution to the generalized eigenvalue problem

$$\mathbf{J}\mathbf{a} = \lambda\mathbf{I}\mathbf{a}.$$

Fast methods to approximate $\mathbf{a}$ and λ are well known.

Symmetric polynomials

The standard monomial basis of symmetric polynomials $P_\alpha(t_1,\ldots,t_k)$ is indexed by partitions $\alpha \coloneqq (\alpha_1,\ldots,\alpha_r)$ of weight $r \le k$.

For example, with $k = 3$ we have,

$$P_{(1,1,1)} = t_1t_2t_3, \quad P_{(2,1,1)} = t_1^2t_2t_3 + t_2^2t_1t_3 + t_3^2t_1t_2, \quad P_{(3)} = t_1^3 + t_2^3 + t_3^3$$

The set $\{P_{(1)}^a P_\alpha : a \ge 0, 1 \notin \alpha\}$ is also a basis, as is the set

$$\{(1 - P_{(1)})^a P_\alpha : a \ge 0, 1 \notin \alpha\}.$$

It turns out to be computationally more convenient to work with the subset

$$\mathcal{B} \coloneqq \{(1 - P_{(1)})^a P_\alpha : a \ge 0, \alpha \subseteq 2\mathbb{N}\},$$

which empirically works nearly as well and is a basis for the subalgebra it generates (its span is closed under multiplication).

Computing the matrices $\mathbf{I}$ and $\mathbf{J}$

To compute $\mathbf{I}$ and $\mathbf{J}$ we use the finite subset $\mathcal{B}_d := \{b \in \mathcal{B} : \deg b \leq d\}$ for some fixed degree d (ideally $d \geq k/2$, but this is only feasible for small k).

We view each $b \in \mathcal{B}_d$ as a function $\mathcal{R}_k \to \mathbb{R}$ by restriction.

We first compute a lookup table of coefficients $c_{\alpha,\beta,\gamma} \in \mathbb{Z}$ defined by

$$P_\alpha P_\beta = \sum_\gamma c_{\alpha,\beta,\gamma} P_\gamma$$

indexed by pairs (α, β) with $\alpha, \beta \subseteq 2\mathbb{N}$ and $\deg(P_\alpha) + \deg(P_\beta) \leq d$.

To compute the entries of $\mathbf{I}$ we use

$$\int_{\mathcal{R}_k} (1 - P_{(1)})^a P_\alpha = \frac{k!}{r_1! \cdots r_s!(k-r)!} \cdot \frac{a!\alpha_1! \cdots \alpha_r!}{(a + \alpha_1 + \cdots + \alpha_r + k)!},$$

where $r_1, \ldots, r_s$ are the multiplicities of the blocks of α.
Computing $\mathbf{J}$ is more work, but it can be reduced to integrals of this form.

I and J as inner products

The quadratic forms I and J both correspond to inner products on $L^2(\mathcal{R}_k)$. Indeed, $I(F) = \int_{\mathcal{R}_k} F^2 = \langle F, F \rangle$ is the standard inner product on $L^2(\mathcal{R}_k)$, and

$$\begin{aligned} J(F) &= \sum_i \int_{\mathcal{R}_{k-1}^{(i)}} \left(\int_0^{1-\sum_{j\neq i} t_j} F \, dt_i' \right)^2 d\mathcal{R}_{k-1}^{(i)} \\ &= \int_{\mathcal{R}_k} F \sum_i \left(\int_0^{1-\sum_{j\neq i} t_j} F \, dt_i' \right) d\mathcal{R}_k \\ &= \langle F, \mathcal{L}F \rangle, \end{aligned}$$

where $\mathcal{L} \colon L^2(\mathcal{R}_k) \to L^2(\mathcal{R}_k)$ is the self-adjoint linear operator

$$\mathcal{L}F := \sum_{i=1}^{k} \int_0^{1-\sum_{j\neq i} t_j} F \, dt_i' \qquad \text{(support truncated to } \mathcal{R}_k\text{)},$$

For any finite set $\{b_1, \ldots, b_n\}$ of linearly independent symmetric functions,

$$\mathbf{I} = \left[\langle b_i, b_j \rangle \right]_{ij}, \qquad \mathbf{J} = \left[\langle \mathcal{L} b_i, b_j \rangle \right]_{ij}.$$

Using a Krylov subspace

For any nonzero F and integer d we may consider the Krylov subpace

$$\operatorname{span}\{F, \mathcal{L}F, \mathcal{L}^2F, \ldots, \mathcal{L}^{d-1}F\}$$

of dimension d. With respect to this basis, $\mathbf{I}$ and $\mathbf{J}$ are Hankel matrices

$$\mathbf{I} = \left[\langle \mathcal{L}^{i+j-2}F, F\rangle\right]_{ij}, \qquad \mathbf{J} = \left[\langle \mathcal{L}^{i+j-1}F, F\rangle\right]_{ij}$$

and we only need to compute the $2d$ values $\langle \mathcal{L}^n F, F\rangle$ for $0 \le n \le 2d-1$.

It is convenient to take $F = 1$, in which case each $\mathcal{L}^n F$ is a symmetric polynomial of degree n. For example

$$\mathcal{L}1 = k + (1-k)P_{(1)},$$
$$\mathcal{L}^2 1 = \frac{k^2 + k - (2k^2 - 2k)P_{(1)} + (2k^2 - 6k + 4)P_{(1,1)} + (k^2 - 3)P_{(2)}}{2}.$$

M_k bounds with the Krylov subspace method

k	lower	upper
2	1.38593	1.38629
3	1.64644	1.64791
4	1.84540	1.84839
5	2.00714	2.01179
...	...	...
10	2.54547	2.55842
20	3.12756	3.15340
30	3.48313	3.51848
40	3.73919	3.78346
50	3.93586	3.99186
53	3.98621	4.04664
54	4.00223	4.06424
60	4.09101	4.16374
...	...	...
100	4.46424	4.65168

Using a Krylov subspace

The entries $\langle \mathcal{L}^n 1, 1\rangle$ of $\mathbf{I}$ and $\mathbf{J}$ are rational functions in k with denominator $(k+n)!$. The numerators P_n are the polynomials

n	P_n
0	1
1	$2k$
2	$5k^2 + k$
3	$14k^3 + 10k^2$
4	$42k^4 + 69k^3 + 10k^2 - k$
5	$132k^5 + 406k^4 + 196k^3 - 14k^2$
6	$429k^6 + 2186k^5 + 2310k^4 + 184k^3 - 79k^2 + 10k$
7	$1430k^7 + 11124k^6 + 21208k^5 + 8072k^4 - 1654k^3 + 124k^2 + 16k$
8	$4862k^8 + 54445k^7 + 167092k^6 + 143156k^5 - 1064k^4 - 7909k^3 + 2558k^2 - 260k$
$\vdots$	$\vdots$

The number of terms in $\mathcal{L}^n 1$ grows very rapidly with n, but W_n has only $n+1$ terms, each of which has just $O(n \log n)$ bits.

Key question: Is there a recurrence we can use to derive P_{n+1} directly from $P_0, \ldots, P_n$ without needing to compute $\mathbf{I}$ and $\mathbf{J}$?