

October 9, 2026

Admin

- HW 2 due next Tuesday
- HW 1 solutions on Canvas

Today: border rank & Schönhage's asymptotic sum inequality

So far we've seen that

- 1) We can bound ω using bounds on the rank of the matrix mult tensor $\langle n, n, n \rangle$
- 2) Bounds on the rank of rectangular matrix mult tensors imply bounds on the rank of square matrix multiplication tensors.

Today we'll see border rank, another tool we can use to bound ω .

As an example, consider polynomial multiplication modulo x^2 .

$$(a_1x + a_0)(b_1x + b_0) = c_1x + c_0 \pmod{x^2}$$

$$c_0 = a_0b_0 \quad c_1 = a_0b_1 + a_1b_0.$$

Corresponds to the tensor

$$T(a, b, c) = a_0b_0c_0 + a_0b_1c_1 + a_1b_0c_1.$$

Obvious: $R(T) \leq 3$. (In fact, $R(T) = 3$.)

But we can approximate T with a rank-2 tensor.

$$T = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left(a_0b_0(\varepsilon c_0 - c_1) + (a_0 + \varepsilon a_1)(b_0 + \varepsilon b_1)c_1 \right)$$

This corresponds to the bilinear algorithm

$$m_1 = a_0 b_0$$

$$m_2 = (a_0 + \varepsilon a_1)(b_0 + \varepsilon b_1)$$

$$c_0 = m_1, \quad c_1 = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (m_2 - m_1)$$

Definition Let $T \in \mathbb{C}^{m \times n \times p}$, $q \in \mathbb{N}$.

- Order- q border rank

$$\underline{R}_q(T) = \min \left\{ r : \begin{array}{l} \exists \text{ linear forms } a_1, \dots, a_r \\ b_1, \dots, b_r \\ c_1, \dots, c_r \\ \text{w/ coefficients in } \mathbb{C}[\varepsilon] \\ \text{st } \sum_{i=1}^r a_i(x) b_i(y) c_i(z) = \varepsilon^q T + O(\varepsilon^{q+1}) \end{array} \right\}$$

- Border rank

$$\underline{R}(T) = \min_q \underline{R}_q(T).$$

Some properties of border rank.

Lemma let T, S be tensors.

$$1) R(T) \leq (3q+1) \underline{R}_q(T).$$

$$2) \underline{R}_{\max(q_1, q_2)}(T \oplus S) \\ \leq \underline{R}_{q_1}(T) + \underline{R}_{q_2}(S).$$

$$3) \underline{R}_{q_1+q_2}(T \otimes S) \leq \underline{R}_{q_1}(T) \underline{R}_{q_2}(S)$$

Proof. (1): polynomial interpolation.

Suppose

$$\sum_{i=1}^r a_i(x) b_i(y) c_i(z) = \varepsilon^q T + O(\varepsilon^{q+1}).$$

Truncate the coefficients of a_i, b_i, c_i modulo ε^{q+1} . Only the $O(\varepsilon^{q+1})$ error term on the RHS changes.

Now $\sum_i a_i b_i c_i = \epsilon^q T + O(\epsilon^{q+1})$ is a polynomial identity in ϵ of degree $\leq 3q$. This means T is in the span of

$$\left\{ \left(\sum_{i=1}^r a_i b_i c_i \right) \Big|_{\epsilon=j} : 0 \leq j \leq 3q \right\}$$

(2): Straightforward. If

$$\sum a_i b_i c_i = \epsilon^{q_1} T + O(\epsilon^{q_1+1})$$

$$\sum d_i e_i f_i = \epsilon^{q_2} S + O(\epsilon^{q_2+1}),$$

then if $q_1 = \max(q_1, q_2)$,

$$\sum_i a_i b_i c_i + \epsilon^{q_1 - q_2} \sum_i d_i e_i f_i$$

$$= \epsilon^{q_1} (T + S) + O(\epsilon^{q_1+1}).$$

Similar if $q_2 = \max(q_1, q_2)$.

(3): Suppose

$$\sum_i a_i b_i c_i = \varepsilon^{q_1} T + O(\varepsilon^{q_1+1})$$

$$\sum_i d_i e_i f_i = \varepsilon^{q_2} S + O(\varepsilon^{q_2+1})$$

Then

$$\sum_{i,j} (a_i d_j) (b_i e_j) (c_i f_j) \quad (*)$$

$$= \left(\sum_i a_i b_i c_i \right) \left(\sum_i d_i e_i f_i \right)$$

$$= \varepsilon^{q_1+q_2} TS + O(\varepsilon^{q_1+q_2+1}).$$

Observe that $a_i(x) d_j(u)$ is linear in the monomials $x_k u_\ell$, so $(*)$ can be turned into a rank decomposition of size

$$\underline{R}_{q_1}(T) \cdot \underline{R}_{q_2}(S)$$

□

Let's see border rank in the context of matrix multiplication.

Consider the tensor corresponding to matrix mult

$$\begin{pmatrix} * & * \\ * & 0 \end{pmatrix} \begin{pmatrix} * & * \\ * & * \end{pmatrix} = \begin{pmatrix} * & * \\ * & * \end{pmatrix}$$

$$\begin{aligned} T = & X_{11} Y_{11} Z_{11} + X_{12} Y_{21} Z_{11} \\ & + X_{11} Y_{12} Z_{12} + X_{12} Y_{22} Z_{12} \\ & + X_{21} Y_{11} Z_{21} + X_{21} Y_{12} Z_{22}. \end{aligned}$$

Textbook algorithm $\Rightarrow R(T) \leq 6$.

But T can be approximated by a rank-5 tensor.

$$T = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \left(\begin{aligned} &(X_{12} + \varepsilon X_{11})(Y_{12} + \varepsilon Y_{22})Z_{12} \\ &+ (X_{21} + \varepsilon X_{11})Y_{11}(Z_{11} + \varepsilon Z_{21}) \\ &- X_{12}Y_{12}(Z_{11} + Z_{12} + \varepsilon Z_{22}) \\ &- X_{21}(Y_{11} + Y_{12} + \varepsilon Y_{21})Z_{11} \\ &+ (X_{12} + X_{21})(Y_{12} + \varepsilon Y_{21})(Z_{11} + \varepsilon Z_{22}) \end{aligned} \right)$$

We say that T has border rank 5,
written $\underline{R}(T) \leq 5$.

By computing two copies of T ,
we get matrix mult of format

$$\begin{array}{|c|c|c|} \hline & & \\ \hline & & \\ \hline & & \\ \hline \end{array} \times \begin{array}{|c|c|} \hline & \\ \hline & \\ \hline \end{array},$$

i.e., the tensor $\{3, 2, 2\}$.

We saw $\underline{R}_1(\langle 3, 2, 2 \rangle) \leq 10$.

Similarly, $\underline{R}_1(\langle 2, 3, 2 \rangle) \leq 10$

$$\underline{R}_1(\langle 2, 2, 3 \rangle) \leq 10.$$

Taking Kronecker products,

$$\underline{R}_3(\langle 12, 12, 12 \rangle) \leq 1000$$

$$\text{so } R(\langle 12, 12, 12 \rangle) \leq 10,000$$

$$\Rightarrow w \leq \log_{12} 10,000 \approx 3.7.$$

If we instead take the q^{th} power of $\langle 12, 12, 12 \rangle$, we get

$$\underline{R}_{3q}(\langle 12^q, 12^q, 12^q \rangle) \leq 10^{3q}$$

$$\Rightarrow R(\langle 12^q, 12^q, 12^q \rangle) \leq (3q+1)10^{3q}$$

$$\text{So } w \leq \frac{3q \log 10 + \log(3q+1)}{q \log 12}$$

$$\leq 2.78 + \frac{\log(3q+1)}{q \log 12}$$

So for $q \rightarrow \infty$, get $w \leq 2.78$.

(Technically, w is an infimum over complexities, not a minimum, so really we have $w \leq \log_{12} 1000$.)

The same proof idea leads to:

Lemma IF $\underline{R}(\langle m, n, p \rangle) \leq r$,

then $w \leq 3 \log_{mnp} r$.

We'll look at one more tool: direct sums of matrix mult tensors.

Fact $R(\langle m, 1, p \rangle) = mp$

$$R(\langle 1, n, 1 \rangle) = n$$

$$R(\langle m, 1, p \rangle \oplus \langle 1, n, 1 \rangle) = mp + n.$$

But in 1981, Schönhage proved something amazing!

Lemma If $n = (m-1)(p-1)$,
then

$$R(\langle m, 1, p \rangle \oplus \langle 1, n, 1 \rangle) \leq mp + 1.$$

Proof We want a decomposition of the tensor

$$\sum_{i=1}^m \sum_{j=1}^p a_i b_j c_{i,j} + \sum_{i=1}^{m-1} \sum_{j=1}^{p-1} u_{i,j} v_{i,j} w.$$

Write

$$u_{i,p} = 0$$

$$v_{i,p} = -\sum_{j < p} v_{i,j}$$

$$u_{m,j} = -\sum_{i < m} u_{i,j}$$

$$v_{m,j} = 0$$

Then

$$\sum_{i=1}^m \sum_{j=1}^p (a_i + \varepsilon u_{i,j})(b_j + \varepsilon v_{i,j})$$
$$(\varepsilon^2 c_{ij} + w)$$

$$= \left(\sum_{i=1}^m a_i \right) \left(\sum_{j=1}^p b_j \right) w$$

$$+ \varepsilon \sum_{i=1}^m \sum_{j=1}^p (u_{ij} b_j w + a_i v_{ij} w)$$

$$+ \varepsilon^2 \left(\sum_{i,j} a_i b_j c_{ij} + \sum_{i,j} u_{ij} v_{ij} w \right)$$
$$+ O(\varepsilon^3).$$

Rearranging gives

$$R_2(\langle m, 1, p \rangle \oplus \langle 1, n, 1 \rangle) \leq mp + 1$$

when $n = (m-1)(p-1)$. $\square$

We can use this to bound ω through Schönhage's asymptotic sum inequality.

Theorem Suppose

$$R\left(\bigoplus_{i=1}^k \langle m_i, n_i, p_i \rangle\right) \leq r$$

and $k \leq r$. Then $\omega \leq 3\tau$,
where τ defined by

$$\sum_{i=1}^k (m_i n_i p_i)^\tau = r.$$

Corollary $\omega < 2.565\dots$

Proof $R(\langle 4, 1, 3 \rangle \oplus \langle 1, 6, 1 \rangle) \leq 13$

$12^\tau + 6^\tau$ solved by $\tau \approx 0.8551$

$\Rightarrow \omega \leq 3\tau \approx 2.565\dots \quad \square$

We'll prove a weaker version of Schönhage's theorem.

Lemma $R\left(\bigoplus_{i=1}^k \langle m, n, p \rangle\right) \leq r$

$\Rightarrow \omega \leq 3 \frac{\log \lceil r/k \rceil}{\log(mnp)}$.

Proof We first show

$$R\left(\bigoplus_{i=1}^k \langle m^s, n^s, p^s \rangle\right) \leq \left\lceil \frac{r}{k} \right\rceil^s \cdot k$$

by induction on s .

Case $s = 1$: true by assumption.

Case $s > 1$:

$$\begin{aligned} & \bigoplus_{i=1}^k \langle m^s, n^s, p^s \rangle \\ &= \bigoplus_{i=1}^k \left(\langle m, n, p \rangle \otimes \langle m^{s-1}, n^{s-1}, p^{s-1} \rangle \right) \\ &= \left(\bigoplus_{i=1}^k \langle m, n, p \rangle \right) \otimes \langle m^{s-1}, n^{s-1}, p^{s-1} \rangle \end{aligned}$$

By assumption,

$$R(\langle m, n, p \rangle) \leq r$$

Let $\langle N \rangle$ be the $N \times N \times N$ diagonal tensor. Any decomposition of $\langle N \rangle$ yields a decomposition of a tensor of rank $\leq N$. So

$$R\left(\bigoplus_{i=1}^k \langle m, n, p \rangle\right) \otimes \langle m^{s-1}, n^{s-1}, p^{s-1} \rangle$$

$$\leq R\left(\langle r \rangle \otimes \langle m^{s-1}, n^{s-1}, p^{s-1} \rangle\right)$$

$$= R\left(\bigoplus_{i=1}^r \langle m^{s-1}, n^{s-1}, p^{s-1} \rangle\right)$$

$$\leq \lceil r/k \rceil \cdot \lceil r/k \rceil^{s-1} \cdot k = \lceil r/k \rceil^s \cdot k.$$

Now we bound w :

$$R(\langle m^s, n^s, p^s \rangle) \leq R\left(\bigoplus_{i=1}^k \langle m^s, n^s, p^s \rangle\right)$$

$$\leq \lceil r/k \rceil^s \cdot k$$

$$\Rightarrow w \leq \frac{3s \log \lceil r/k \rceil + 3 \log k}{3s \log(mnp)}$$

$$= \frac{3 \log \lceil r/k \rceil}{\log(mnp)} + O\left(\frac{\log s}{s}\right)$$

$$\begin{matrix} s \rightarrow \infty \\ \longrightarrow \end{matrix} \frac{3 \log \lceil r/k \rceil}{\log(mnp)}$$

