Yuval's proof of Strassen's result

ho boon suan, july 2021

Instead of writing

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} Q & R \\ S & T \end{pmatrix},$$

we may write

$$\begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{pmatrix} \begin{pmatrix} e \\ f \\ g \\ h \end{pmatrix} = \begin{pmatrix} Q \\ R \\ S \\ T \end{pmatrix}.$$

Let's use a dot in place of 0 to simplify things visually. Since

$$\begin{pmatrix} a & \cdot & b & \cdot \\ \cdot & a & \cdot & b \\ c & \cdot & d & \cdot \\ \cdot & c & \cdot & d \end{pmatrix}$$

$$= \begin{pmatrix} a & \cdot & a & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a & \cdot & a & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} + \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & d & \cdot & d \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & d & \cdot & d \end{pmatrix}$$

$$+ \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & a-d & a-d & \cdot \\ \cdot & d-a & d-a & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

$$+ \begin{pmatrix} \cdot & \cdot & b-a & \cdot \\ \cdot & \cdot & b-a & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} + \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & d-b & b-d \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix} + \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & c-d & \cdot & \cdot \\ \cdot & c-d & \cdot & \cdot \end{pmatrix} + \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ c-a & a-c & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix},$$

we may set

$$\begin{cases} x_1 = a(e+g), \\ x_2 = d(f+h), \\ x_3 = (a-d)(f+g), \\ x_4 = (b-a)g, \\ x_5 = (d-b)(g-h), \\ x_6 = (c-d)f, \\ x_7 = (c-a)(e-f) \end{cases}$$

to obtain

$$\begin{pmatrix} ae+bg \\ af+bh \\ ce+dg \\ cf+dh \end{pmatrix} = \begin{pmatrix} x_1 \\ \cdot \\ x_1 \\ \cdot \end{pmatrix} + \begin{pmatrix} \cdot \\ x_2 \\ \cdot \\ x_2 \end{pmatrix} + \begin{pmatrix} \cdot \\ x_3 \\ -x_3 \\ \cdot \end{pmatrix} + \begin{pmatrix} x_4 \\ x_4 \\ \cdot \\ \cdot \end{pmatrix} + \begin{pmatrix} \cdot \\ x_5 \\ \cdot \\ \cdot \end{pmatrix} + \begin{pmatrix} \cdot \\ \cdot \\ x_6 \\ x_6 \end{pmatrix} + \begin{pmatrix} \cdot \\ \cdot \\ x_7 \\ \cdot \end{pmatrix}.$$

Thus we may multiply 2×2 matrices with only seven multiplications, instead of the usual eight, leading us to a recursive algorithm for matrix multiplication with time complexity $\Theta(n^{\log_2 7}) \approx \Theta(n^{2.8074})$.