The first (and only?) step of Gaussian elimination

Let $A^{(1)} = M_1 A$ be the matrix obtained by using the first row of A to eliminate (send to zero) the entries below a_{11}. Recall that M_1 is of the form $M_1 = I + m_1 e_1^t$, and that we want

$$M_1 A e_1 = a_{11} e_1.$$

Then $a_{11}e_1 = (I + m_1 e_1^t)Ae_1 = Ae_1 + m_1 e_1^t A e_1 = Ae_1 + a_{11} m_1$, and if $a_{11} \neq 0$, we can solve this for m_1:

$$m_1 = (\tfrac{1}{a_{11}})(a_{11}e_1 - Ae_1).$$

This is simply a vector notation for $m_{11} = 0$ and $m_{i,1} = -a_{i,1}/a_{11}, \quad i = 2, 3, \ldots, n.$

Once we know m_1, we can complete step 1 of GE, by computing $A^{(1)} = M_1 A$:

$$A^{(1)} = A + m_1(e_1^t A).$$

We do *not* explicitly form M_1 (why?), we *do* respect those parentheses (why?), and we do *not* explictly multiply e_1^t and A (why?). If one pays attention to these admonitions, then m_1 and $A^{(1)}$ can be computed in 1 division, $n-1$ multiplications, $(n-1)^2$ multiplications, and $(n-1)^2$ additions, for a total of $2n^2 + \mathrm{O}(n)$ flops.

We are finished with step 1 of GE, and partition $A^{(1)}$ as

$$A^{(1)} = \begin{bmatrix} a_{11} & \tilde{a}_1^t \\ 0 & \tilde{A}^{(1)} \end{bmatrix}.$$

Please stop reading for a bit and think about step 2... Isn't step 2 going to be just like step 1, but with the names changed? Won't we use the first row of $\tilde{A}^{(1)}$ to eliminate the entries below $\tilde{a}_{11}^{(1)}$? Then you know how it goes. In fact, you know how it goes for step $k+1$: Suppose we've completed k steps ($\tilde{A}^{(k)}$ is $(n-k) \times (n-k)$):

$$A^{(k)} = \begin{bmatrix} U^{(k)} & X^{(k)} \\ 0 & \tilde{A}^{(k)} \end{bmatrix}.$$

You know the formulas,

$$\tilde{m}_{k+1} = (\tfrac{1}{\tilde{a}_{11}})(\tilde{a}_{11}e_1 - \tilde{A}^{(k)}e_1) \quad \text{and} \quad \hat{A}^{(k+1)} = \tilde{A}^{(k)} + \tilde{m}_{k+1}(e_1^t \tilde{A}^{(k)}).$$

We get the $\tilde{A}^{(k+1)}$ simply by partioning $\hat{A}^{(k+1)}$ as we partitioned $A^{(1)}$ above.

Computing $\tilde{A}^{(k+1)}$ requires $2(n-k)^2 + \mathrm{O}(n-k)$ flops, so the flop count for all $n-1$ steps is

$$\sum_{k=1}^{n-1}(2(n-k)^2 + \mathrm{O}(n-k)) = \frac{2}{3}n^3 + \mathrm{O}(n^2).$$

The relationship of M_k to $\tilde{M}_k$ is $m_k^t = [0 \;\; \tilde{m}_k^t]$ and

$$I + m_k e_k^t \quad = \quad M_k \quad = \quad \begin{bmatrix} I & 0 \\ 0 & \tilde{M}_k \end{bmatrix} \quad = \quad \begin{bmatrix} I & 0 \\ 0 & I + \tilde{m}_k e_1^t \end{bmatrix}.$$