Remainder of the Sum of Remainders

Note that directly iterating over all pairs (i, j) in $O(n^2)$ is impossible with such constraints on n.

The goal of the tutorial is to derive a formula that allows us to compute the sum in $O(\sqrt{n})$.

Part 1. Fixing j

Let $T(j) = \sum_{i=1}^{n}(i \bmod j)$. Then $S(n) = \sum_{j=1}^{n} T(j)$.

Let's compute $T(j)$.

Consider the numbers $1, 2, \ldots, n$ and their remainders when divided by j. They are divided into blocks of length j:

$$[1, \ldots, j], \quad [j+1, \ldots, 2j], \quad \ldots$$

and possibly the last incomplete block.

Let

$$q = \left\lfloor \frac{n}{j} \right\rfloor \quad \text{(number of complete blocks)}, \qquad r = n \bmod j = n - qj \quad \text{(length of the tail)}.$$

Then:

- Each complete block gives a set of remainders $0, 1, 2, \ldots, j-1$, their sum equals $0 + 1 + \cdots + (j-1) = \frac{j(j-1)}{2}$. There are q such blocks, so the contribution of complete blocks is: $q \cdot \frac{j(j-1)}{2}$.
- The last (incomplete) block gives remainders $0, 1, 2, \ldots, r$, their sum equals $0 + 1 + \cdots + r = \frac{r(r+1)}{2}$.

Thus,

$$T(j) = q \cdot \frac{j(j-1)}{2} + \frac{r(r+1)}{2}, \qquad \text{where } q = \left\lfloor \frac{n}{j} \right\rfloor, \quad r = n - qj.$$

Substituting this into the sum:

$$S(n) = \sum_{j=1}^{n} \left(q \cdot \frac{j(j-1)}{2} + \frac{(n-qj)(n-qj+1)}{2} \right).$$

For convenience, let's denote the two parts:

$$A = \sum_{j=1}^{n} q \cdot \frac{j(j-1)}{2}, \qquad B = \sum_{j=1}^{n} \frac{(n-qj)(n-qj+1)}{2}.$$

Then $S(n) = A + B$.

Part 2. Grouping by values of $q = \left\lfloor \frac{n}{j} \right\rfloor$

Key observation: the value $q = \left\lfloor \frac{n}{j} \right\rfloor$ does not change for every j, but is piecewise constant over entire intervals.

For a fixed q, the set of j that gives this value q forms an interval:

$$j \in [L, R], \text{where } L = \left\lfloor \frac{n}{q+1} \right\rfloor + 1, \qquad R = \left\lfloor \frac{n}{q} \right\rfloor.$$

An equivalent way to construct intervals (convenient for implementation): we will iterate j from left to right. Let's say we are currently at some $j = i$. Then

$$q = \left\lfloor \frac{n}{i} \right\rfloor, \quad R = \left\lfloor \frac{n}{q} \right\rfloor.$$

Clearly, for all j in the interval $[i, R]$, the value $\left\lfloor \frac{n}{j} \right\rfloor$ is the same and equals q. After processing this interval, we move to $i = R + 1$.

It is important that the number of such intervals is on the order of $O(\sqrt{n})$, since after the point $j > \sqrt{n}$, the values of q are small and take only about $\sqrt{n}$ different values, and for j up to $\sqrt{n}$, there are also about $\sqrt{n}$ such j.

Next, we need to be able to quickly compute sums over the interval $[L, R]$:

$$\sum_{j=L}^{R} 1, \qquad \sum_{j=L}^{R} j, \qquad \sum_{j=L}^{R} j^2.$$

Let's denote:

$$\text{cnt} = R - L + 1,$$
$$S_1(L, R) = \sum_{j=L}^{R} j,$$
$$S_2(L, R) = \sum_{j=L}^{R} j^2.$$

These sums can be conveniently expressed using prefix formulas:

$$\sum_{j=1}^{x} j = \frac{x(x+1)}{2}, \qquad \sum_{j=1}^{x} j^2 = \frac{x(x+1)(2x+1)}{6}.$$

Then

$$S_1(L, R) = \frac{R(R+1)}{2} - \frac{(L-1)L}{2},$$
$$S_2(L, R) = \frac{R(R+1)(2R+1)}{6} - \frac{(L-1)L(2L-1)}{6}.$$

All these formulas will be computed modulo M, using the inverse elements of 2 and 6 (since M is prime).

Part 3. Contribution of the interval $[L, R]$ to A

Recall:

$$A = \sum_{j=1}^{n} q \cdot \frac{j(j-1)}{2}.$$

Over the entire interval $[L, R]$, the value of q is constant, that is

$$A_{[L,R]} = \sum_{j=L}^{R} q \cdot \frac{j(j-1)}{2} = q \cdot \frac{1}{2} \sum_{j=L}^{R} (j(j-1)).$$

Note that

$$j(j-1) = j^2 - j.$$

Then

$$\sum_{j=L}^{R}(j(j-1)) = \sum_{j=L}^{R} j^2 - \sum_{j=L}^{R} j = S_2(L,R) - S_1(L,R).$$

From this, we have

$$A_{[L,R]} = q \cdot \frac{1}{2}\big(S_2(L,R) - S_1(L,R)\big).$$

Part 4. Contribution of the interval $[L, R]$ to B

Now consider

$$B = \sum_{j=1}^{n} \frac{(n-qj)(n-qj+1)}{2}.$$

Let $r = n - qj$. Then

$$\frac{r(r+1)}{2} = \frac{(n-qj)(n-qj+1)}{2}.$$

Expanding the brackets:

$$r(r+1) = (n-qj)(n-qj+1) = (n-qj)(n+1-qj).$$

Multiplying:

$$(n-qj)(n+1-qj) = n(n+1) - qj(2n+1) + q^2j^2.$$

Thus,

$$\frac{(n-qj)(n-qj+1)}{2} = \frac{1}{2}\left(n(n+1) - q(2n+1)\cdot j + q^2 \cdot j^2\right).$$

Then the contribution of the interval $[L, R]$:

$$B_{[L,R]} = \sum_{j=L}^{R} \frac{1}{2}\left(n(n+1) - q(2n+1)\cdot j + q^2 \cdot j^2\right) = \frac{1}{2}\Big[n(n+1)\cdot \mathrm{cnt} - q(2n+1)\cdot S_1(L,R) + q^2 \cdot S_2(L,R)\Big],$$

where

$$\mathrm{cnt} = R - L + 1.$$

Part 5. Final Formula

By summing the contributions $A_{[L,R]}$ and $B_{[L,R]}$, we obtain the contribution of the interval $[L, R]$ to the overall answer:

$$S_{[L,R]} = A_{[L,R]} + B_{[L,R]}.$$

It remains to sum this over all intervals corresponding to different values of $q = \left\lfloor \frac{n}{j} \right\rfloor$.

Since the number of such intervals is $O(\sqrt{n})$, the overall asymptotic complexity of the algorithm will be $O(\sqrt{n})$, which is feasible for $n \le 10^{12}$.

Part 6. Implementation Modulo

Calculations are performed modulo $M = 998244353$.

You need to be able to:

- quickly multiply large numbers with modulo. In C++, for intermediate products, you can use the type `__int128`;
- compute inverse elements $\frac{1}{2}$, $\frac{1}{3}$ (and thus $\frac{1}{6}$) modulo M using exponentiation to $M-2$ by Fermat's little theorem;
- carefully bring all formulas S_1, S_2, cnt, $A_{[L,R]}$, $B_{[L,R]}$ to arithmetic modulo.

Code structure:

1. Read n.
2. Precompute $\text{inv2} = 2^{M-2} \bmod M$, $\text{inv3} = 3^{M-2} \bmod M$.
3. Define functions for the sum of the first x numbers and the sum of the first x squares modulo:

$$\sum_{k=1}^{x} k = \frac{x(x+1)}{2}, \qquad \sum_{k=1}^{x} k^2 = \frac{x(x+1)(2x+1)}{6}.$$

4. Start a loop over intervals:

```
i = 1
while i <= n:
    q = n // i
    R = n // q
    L = i
    compute contribution from [L, R]
    i = R + 1
```

5. Output the accumulated sum modulo M.

Part 7. Verification with Example

Let $n = 5$. Then manually:

$i \bmod j$	$j = 1$	2	3	4	5
$i = 1$	0	1	1	1	1
2	0	0	2	2	2
3	0	1	0	3	3
4	0	0	1	0	4
5	0	1	2	1	0

The sum of all elements in the table equals 26. The algorithm described above also yields 26.

Conclusion

Main ideas of the solution:

1. Renumber the sum over j and express $\sum(i \bmod j)$ for a fixed j in terms of the number of complete blocks $q = \lfloor n/j \rfloor$ and the tail.
2. Notice that $q = \lfloor n/j \rfloor$ takes few different values, and process entire intervals $j \in [L, R]$ with the same q.
3. For each interval, express the contribution using elementary sums 1, j, j^2, which are computed using formulas.
4. Compute everything modulo 998244353 and use modular inverses.

The asymptotic complexity of the solution is $O(\sqrt{n})$, which is sufficient for $n \le 10^{12}$.