Остаток от суммы остатков

Заметим, что прямой перебор всех пар (i, j) за $O(n^2)$ невозможен при таких ограничениях на n. Цель разбора — вывести формулу, позволяющую посчитать сумму за $O(\sqrt{n})$.

Часть 1. Фиксируем j

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

Посчитаем $T(j)$.

Рассмотрим числа $1, 2, \ldots, n$ и их остатки при делении на j. Они разбиваются на блоки длины j:

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

и, возможно, последний неполный блок.

Пусть

$$q = \left\lfloor \frac{n}{j} \right\rfloor \quad \text{(сколько полных блоков)}, \qquad r = n \bmod j = n - qj \quad \text{(длина хвоста)}.$$

Тогда:

- Каждый полный блок даёт набор остатков $0, 1, 2, \ldots, j-1$, их сумма равна $0 + 1 + \cdots + (j-1) = \frac{j(j-1)}{2}$. Таких блоков q, значит вклад полных блоков: $q \cdot \frac{j(j-1)}{2}$.
- Последний (неполный) блок даёт остатки $0, 1, 2, \ldots, r$, их сумма равна $0 + 1 + \cdots + r = \frac{r(r+1)}{2}$.

Следовательно,

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

Подставим это в сумму:

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

Для удобства обозначим две части:

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

Тогда $S(n) = A + B$.

Часть 2. Группируем по значениям $q = \left\lfloor \frac{n}{j} \right\rfloor$

Ключевое наблюдение: значение $q = \left\lfloor \frac{n}{j} \right\rfloor$ меняется не при каждом j, а кусочно-постоянно на целых отрезках.

Для фиксированного q множество j, дающих это значение q, образует интервал:

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

Эквивалентный способ построения интервалов (удобный в реализации): будем перебирать j слева направо. Пусть сейчас мы стоим в некотором $j = i$. Тогда

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

Ясно, что для всех j из отрезка $[i, R]$ значение $\left\lfloor \frac{n}{j} \right\rfloor$ совпадает и равно q. После обработки этого отрезка переходим к $i = R + 1$.

Важно, что количество таких отрезков порядка $O(\sqrt{n})$, поскольку после точки $j > \sqrt{n}$ значения q малы и принимают всего примерно $\sqrt{n}$ разных значений, а до $\sqrt{n}$ самих j тоже около $\sqrt{n}$.

Дальше нам нужно уметь быстро посчитать суммы по отрезку $[L, R]$:

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

Обозначим:

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

$$S_1(L, R) = \sum_{j=L}^{R} j,$$

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

Эти суммы удобно выражаются через префиксные формулы:

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

Тогда

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

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

Все эти формулы будем считать по модулю M, используя обратные элементы к 2 и 6 (так как M — простое).

Часть 3. Вклад отрезка $[L, R]$ в A

Напомним:

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

На всём рассматриваемом отрезке $[L, R]$ значение q постоянно, то есть

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

Заметим, что

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

Тогда

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

Отсюда

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

Часть 4. Вклад отрезка $[L, R]$ в B

Теперь рассмотрим

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

Обозначим $r = n - qj$. Тогда

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

Раскроем скобки:

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

Перемножим:

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

Следовательно,

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

Тогда вклад отрезка $[L, R]$:

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

где

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

Часть 5. Итоговая формула

Сложив вклады $A_{[L,R]}$ и $B_{[L,R]}$, получим вклад интервала $[L, R]$ в общий ответ:

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

Остаётся просуммировать это по всем отрезкам, соответствующим различным значениям $q = \left\lfloor \frac{n}{j} \right\rfloor$.

Так как количество таких отрезков $O(\sqrt{n})$, итоговая асимптотика алгоритма будет $O(\sqrt{n})$, что укладывается для $n \leqslant 10^{12}$.

Часть 6. Реализация по модулю

Вычисления производятся по модулю $M = 998244353$.

Нужно уметь:

- быстро умножать большие числа с взятием по модулю. В C++ для промежуточного произведения можно использовать тип `__int128`;
- считать обратные элементы $\frac{1}{2}$, $\frac{1}{3}$ (а значит и $\frac{1}{6}$) по модулю M через возведение в степень $M-2$ по малой теореме Ферма;
- аккуратно приводить все формулы S_1, S_2, cnt, $A_{[L,R]}$, $B_{[L,R]}$ к арифметике по модулю.

Структура кода:

1. Считать n.

2. Предварительно посчитать $\mathrm{inv2} = 2^{M-2} \bmod M$, $\mathrm{inv3} = 3^{M-2} \bmod M$.

3. Определить функции для суммы первых x чисел и суммы первых x квадратов по модулю:

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

4. Запустить цикл по отрезкам:

```
i = 1
while i <= n:
    q = n // i
    R = n // q
    L = i
    посчитать вклад от [L, R]
    i = R + 1
```

5. Вывести накопленную сумму по модулю M.

Часть 7. Проверка на примере

Пусть $n = 5$. Тогда вручную:

$i \bmod j$	$j = 1$	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

Сумма всех элементов таблицы равна 26. Алгоритм, описанный выше, тоже выдаёт 26.

Итог

Главные идеи решения:

1. Перенумеровать сумму по j и выразить $\sum(i \bmod j)$ для фиксированного j через число полных блоков $q = \lfloor n/j \rfloor$ и хвост.

2. Заметить, что $q = \lfloor n/j \rfloor$ принимает мало разных значений, и обрабатывать целые отрезки $j \in [L, R]$ с одинаковым q.

3. На каждом отрезке выражать вклад через элементарные суммы 1, j, j^2, которые считаются по формулам.

4. Всё считать по модулю 998244353 и использовать обратные по модулю числа.

Асимптотика решения — $O(\sqrt{n})$, что достаточно для $n \leqslant 10^{12}$.