Solution

To solve this problem, we need to understand the coefficients of the expanded form of $p(x)$. We will be able to do this by showing that the coefficients of $p(x)$ are “counting” something. This will allow us to change the problem into a counting problem. Using the coefficients of a polynomial to count things is a very common and useful technique in an area of mathematics called enumerative combinatorics.

This translation into a counting problem is based on the following observation: After $p(x)$ is expanded and like terms are collected, the coefficient of $x^k$, namely $c_k$, is equal to the number of ways $k$ can be expressed as a sum of distinct numbers from the list $1, 3, 3^2, 3^3, \ldots, 3^{20}$.

To get an idea of what happens when $p(x)$ is expanded, first consider the simpler product $(a + b)(c + d)(e + f)$. Expanding in two steps, we get

\[
(a + b)(c + d)(e + f) = (ac + ad + bc + bd)(e + f) = ace + acf + ade + adf + bce + bcf + bde + bdf.
\]

Notice that there are 8 terms in this product of three binomials and that $2^3 = 8$.

Another way to think of the expansion of $(a + b)(c + d)(e + f)$ is as follows: Each term in the resulting sum is formed by choosing 3 elements, one from each of the 3 sets $\{a, b\}, \{c, d\},$ and $\{e, f\}$, and multiplying them together. For example, if we choose $a$, $c$, and $f$, then we get the product $acf$ which is indeed one of the terms in the expanded sum. Since there are 2 choices for each of the 3 elements, we have $2^3 = 8$ products in the expanded sum.

This generalizes to products of more than 3 binomials. In this question, we are interested in the following product of 21 binomials:

\[
(1 + x)(1 + x^3)(1 + x^9)(1 + x^{27})(1 + x^{81}) \cdots (1 + x^{3^{20}}).
\]

Similar to the earlier product, the terms in the expanded sum can each be formed by choosing 21 elements, one from each of the 21 sets

\[
\{1, x\}, \{1, x^3\}, \{1, x^9\}, \ldots, \{1, x^{3^{20}}\},
\]
and multiplying them together. In this case there will be $2^{21}$ such terms. For example, one of the terms will be $1 \times 1 \times x^{3^2} \times 1 \times 1 \times \cdots \times 1 \times x^{3^{19}} \times x^{3^{20}}$. Using exponent rules, we can rewrite this term as $x^{3^{2} + 3^{19} + 3^{20}}$.

Each term in the expanded sum is a product of some number of 1s and distinct terms of the form $x^{3^n}$ with $0 \leq n \leq 20$. Thus, each term can be simplified into the form $x^k$ where $k$ is a sum of distinct numbers from the list $1, 3, 3^2, \ldots, 3^{20}$.

Therefore, when $p(x)$ is expanded, there is a term $x^k$ for every way there is to write $k$ as a sum of distinct numbers in the list $1, 3, 3^2, \ldots, 3^{20}$. After any like terms in the expanded sum are collected, the coefficient of $x^k$, namely $c_k$, will be equal to the number of ways that $k$ can be expressed as a sum of distinct numbers from the list $1, 3, 3^2, \ldots, 3^{20}$.

Note: It turns out that every coefficient of $p(x)$ is either 0 or 1. This corresponds to the fact that every positive integer either cannot be expressed as a sum of distinct powers of 3, or it can only be expressed as a sum of distinct powers of 3 in one way. With respect to the argument in the previous paragraph, this means that there are actually no like terms to be collected and so we have $c_k = 0$ or $c_k = 1$. This fact is not needed for the solution, but it is still interesting and will be discussed further later.

Since $c_k$ is equal to the number of ways that $k$ can be expressed as a sum of distinct numbers from the list $1, 3, 3^2, \ldots, 3^{20}$, the sum

$$c_{1000000} + c_{1000001} + c_{1000002} + \cdots + c_{1999999} + c_{2000000}$$

is equal to the number of sums of distinct numbers from the list $1, 3, 3^2, \ldots, 3^{20}$ which are between $10^6$ and $2 \times 10^6$ (inclusive). All that remains is to count these sums.

First, note that

$$3^{13} = 1\,594\,323 < 2 \times 10^6 < 4\,782\,969 = 3^{14}.$$

Since $3^{14}$ exceeds $2 \times 10^6$ on its own, and hence so do all higher powers of 3, the largest power of 3 that can occur in a sum less than $2 \times 10^6$ is $3^{13}$.

Also, note that

$$3^0 + 3^1 + 3^2 + \cdots + 3^{12} = 797\,161$$

which is less than $10^6$. This means that any sum of distinct powers of 3 that is at least $10^6$ must include a power at least as large as $3^{13}$. Since this is the largest allowable power, any such sum must include $3^{13}$.

Furthermore, $3^{12} + 3^{13} = 2\,125\,764$, which is larger than $2 \times 10^6$, so the sums we wish to count cannot include $3^{12}$.
It can be verified using a calculator (or a pencil, some paper, and spare time) that

\[ 10^6 < 3^{13} < 3^0 + 3^1 + 3^2 + \cdots + 3^{11} + 3^{13} = 1 860 \, 043 < 2 \times 10^6, \]

and so there are no other restrictions on which powers can be included in such a sum. In summary, we have that a sum of distinct powers of 3 which is between 10^6 and 2 \times 10^6

1. includes 3^{13},

2. does not include 3^{12}, and

3. includes some (possibly none) of the powers of 3 from 3^0 to 3^{11} inclusive.

Therefore, the number of sums of distinct powers of 3 between 10^6 and 2 \times 10^6 is 2^{12} = 4096 since for each of the 12 numbers 3^0, 3^1, \ldots, 3^{11}, such a sum can either include this number or not include it. Thus, there are 4096 such sums and so

\[ c_{1000000} + c_{1000001} + c_{1000002} + \cdots + c_{1999999} + c_{2000000} = 4096 \]

**Ternary expansions and why \( c_k \) is always 0 or 1:**

As promised, we now include some discussion about the fact that an integer either cannot be expressed as a sum of distinct powers of 3, or it can be expressed in only one way. This may remind you of binary expansions of positive integers. Essentially, a binary expansion of a positive integer is an expression of that integer as a sum of powers of 2 where each power can be used at most once. There is also such a thing as a ternary expansion which is an expression of an integer as a sum of powers of 3 where each power may be used at most two times. Both of these are similar to the more familiar base 10 or decimal expansion of a positive integer, which is really an expression of the number as a sum of powers of 10 where each power is allowed to be used at most nine times. Binary, ternary, and decimal expansions of positive integers always exist and are always unique.

Relating back to the the problem, the fact that the coefficients of \( p(x) \) are either 0 or 1 is related to ternary expansions. In this question we are able to use each power of 3 at most once which means that we will lose the ability to write some integers using powers of 3 resulting in \( c_k = 0 \) for some \( k \). For example, \( c_6 = 0 \), and this corresponds to the fact that the only way to write 6 as a sum of powers of 3 using any power of 3 at most twice is \( 6 = 3 + 3 \). Hence, 6 cannot be expressed as a sum of distinct powers of 3. However, if a positive integer can be written as a sum of distinct powers of 3 (and hence can be written as a sum of powers of 3 where each power is used at most twice), the uniqueness of ternary expansions says there is only one way to do it, which means \( c_k \leq 1 \) for all \( k \). If you are interested, I suggest you read a bit about ternary expansions of positive
integers and try to understand how they relate to this problem. Here is a stand-alone justification that if a positive integer can be expressed as a sum of distinct powers of 3, then this can be done in only one way.

Suppose \( k \) is a positive integer, \( n_1 < n_2 < \cdots < n_r \) are nonnegative integers, and \( m_1 < m_2 < \cdots < m_s \) are nonnegative integers satisfying

\[
k = 3^{n_1} + 3^{n_2} + \cdots + 3^{n_{r-1}} + 3^{n_r} = 3^{m_1} + 3^{m_2} + \cdots + 3^{m_{s-1}} + 3^{m_s}.
\]

(Note that the two given sums express \( k \) as the sum of distinct powers of 3 and that we are not assuming that \( r = s \).) We will show that \( r = s \), and that \( n_1 = m_1, n_2 = m_2, \) and so on up to \( n_r = m_s \). In other words, we will show that the two expressions of \( k \) as a sum of distinct powers of 3 must actually be the same expression.

Suppose \( n_r < m_s \). Since these are integers, it follows that \( n_r + 1 \leq m_s \). Since \( n_1, n_2, \ldots, n_r \) are all distinct and between 0 and \( n_r \) inclusive (\( n_r \) is the largest), we certainly have that

\[
k = 3^{n_1} + 3^{n_2} + \cdots + 3^{n_r} \leq 3^0 + 3^1 + 3^2 + \cdots + 3^{n_r}.
\]

It can be shown that the sum on the right equals \( \frac{3^{n_r+1} - 1}{2} \). (You can prove this formula yourself, or observe that the sum on the right is a finite geometric series.) Therefore,

\[
k = 3^{n_1} + 3^{n_2} + \cdots + 3^{n_r} \leq \frac{3^{n_r+1} - 1}{2} < \frac{3^{n_r+1}}{2} < 3^{n_r+1} \leq 3^{m_s}.
\]

Reading the chain of inequalities, this implies \( k < 3^{m_s} \).

We also have

\[
k = 3^{m_1} + 3^{m_2} + \cdots + 3^{m_s} \geq 3^{m_s},
\]

which implies \( k \geq 3^{m_s} \). We cannot have \( k \geq 3^{m_s} \) and \( k < 3^{m_s} \) simultaneously, so our assumption that \( n_r < m_s \) must have been false. This means \( m_s \leq n_r \). Similar reasoning can be used to show that \( n_r \leq m_s \). Together with \( m_s \leq n_r \), we get \( n_r = m_s \).

This means we can subtract the largest term in each sum and establish that the following two “reduced” sums are equal

\[
3^{n_1} + 3^{n_2} + \cdots + 3^{n_{r-1}} = 3^{m_1} + 3^{m_2} + \cdots + 3^{m_{s-1}}.
\]

The argument above can now be repeated to show that the largest power of 3 in each of the sums above must be the same, that is, \( n_{r-1} = m_{s-1} \). Continuing this process, we get \( n_{r-2} = m_{s-2} \), and so on.
Eventually, one of the two sums must “run out” of terms. Since the two “reduced” sums are equal at every stage, the two sums must run out of terms at the same time. (If one side ran out before the other then we would have 0 equal to a sum of powers of 3 which is impossible.) This means we must have $r = s$. At the second to last stage we will have $3^{n_1} = 3^{m_1}$ and so we will conclude $n_1 = m_1$ to complete the argument.

This argument is using what is called inductive reasoning. The principle of mathematical induction, which you may have heard of, is a technique in mathematics which solidifies parts of this kind of argument like “eventually one of the two sums must run out”. If you have seen mathematical induction before (or would like to read up on it), a nice exercise might be to try to write a more rigourous proof of the above fact.