Diophantine triples: Meandering towards a solution to the N-matchstick challenge

Neville de Mestre*

Mathematicians spend many hours seeking a solution to a particular problem, attempting many unsuccessful approaches, building up a deeper understanding of the nuances of the problem, and hopefully evolving towards a solution. Yet, when it is finally published, only a direct track to the solution is given. Research papers, textbooks, lectures or classroom teaching all follow this line in general, but our students need to be exposed to the unsuccessful trials and broader methods involved in solving a problem, just as we need to be exposed to the thinking processes that the students are following. In this article I propose to give an account of such meanderings towards a possible solution, and hope that it may be of benefit to budding lecturers and their students.

In *The Australian Mathematics Teacher*, volume 75, number 1, 2015, authors Pat Graham and Helen Chick present an extremely useful detailed account of how Year 7 students attacked the problem referred to as the 20-matchstick challenge.

The problem is to find as many different triangles as possible which have a perimeter of 20 units, where each side has to be an integer number of units. The use of 20 matchsticks of equal length adds a possible hands-on approach to the problem, which could help some students in communicating answers. The problem introduces the need for a systematic approach to be sure that all possibilities have been explored, the consideration of whether or not to include degenerate triangles (a, b, c) with $a + b = c$, and the triangle inequality for sides. I recommend that mathematics teachers and lecturers should read the account.

As an obvious extension, the authors suggested that the problem be investigated for other integer perimeters. Therefore, I wondered if there would be a pattern to such answers with an associated mathematical formula. So the following covers my trials, rejections, deductions and current conclusions, as I explored the general N-matchstick problem over a number of weeks.

To begin with, I explored the problem for $N = 3$ up to $N = 10$ and obtained the following results with T as the number of possible triangles.

N	3	4	5	6	7	8	9	10
T	1	0	1	1	2	1	3	2

*Bond University, Gold Coast, Queensland. Email: margnev2@bigpond.com

There seems to be no obvious pattern in this so I extended to $N = 20$.

N	11	12	13	14	15	16	17	18	19	20
T	4	3	5	4	7	5	8	7	10	8

The last entry is the already known solution to the 20-matchstick challenge as investigated in detail by Graham and Chick. The possible triangles in this case are $(2, 9, 9)$, $(3, 8, 9)$, $(4, 8, 8)$, $(4, 7, 9)$, $(5, 7, 8)$, $(5, 6, 9)$, $(6, 7, 7)$, $(6, 6, 8)$. Although again no pattern jumped out immediately, I noted that the number of triangles for an odd N was larger than the number for the next even N, except for $N = 5, 6$. So, I wrote down the patterns for the odd and even Ns separately, yielding

$$\text{(odd)}\quad 1, 1, 2, 3, 4, 5, 7, 8, 10$$
$$\text{(even)}\quad 0, 1, 1, 2, 3, 4, 5, 7, 8$$

Except for the zero at the start of the even pattern these appear to be the same, but I could not see a pattern yet. To check that there might be a clue here to a pattern, I explored further for $N = 21$ to 24.

N	21	22	23	24
T	12	10	14	12

The odd/even similarities continued, and the even pattern is three steps behind the odd pattern. I wondered why? (As they say in basketball, 'TIME OUT!'. This is the time to stop and reflect on what direction to take next.)

Since one of the properties of a triangle is that it has three sides, I decided to look at the pattern formed by the sums of each group of three consecutive T values. This yielded

$$2, 4, 9, 12, 20, 25, 36,$$

which didn't seem to help much, but a few square numbers were appearing. But the similar patterns for the odd and even T values suggested that I should do the same for successive pairs of T values. This yielded

$$1, 2, 3, 5, 7, 9, 12, 15, 18, 22, 26$$

and I had generated the beginnings of a possible breakthrough! The first three differ by 1, the next three by 2, the next three by 3, and the last two by 4. After evaluating $T(25) = 16$ and $T(26) = 14$, it was clear that this pattern was continuing. (TIME OUT again and more reflection!)

Some days later I returned to my list of triangles (a, b, c) for each N from 3 to 24 and placed them in groups starting with $a = 1$ up to the maximum value allowable to satisfy the triangle inequality. For $N = 24$ this was clearly $a = 8$ for the triangle $(8, 8, 8)$. By insisting on $a \leq b \leq c$, I averted duplication of triangles and knew that the triangle inequality was covered for $a + b \leq c$. For example, the list for $N = 13$ to 16 is

N	$a = 1$	$a = 2$	$a = 3$	$a = 4$	$a = 5$
13	$(1, 6, 6)$	$(2, 5, 6)$	$(3, 5, 5)$ $(3, 4, 6)$	$(4, 4, 5)$	
14		$(2, 6, 6)$	$(3, 5, 6)$	$(4, 5, 5)$ $(4, 4, 6)$	
15	$(1, 7, 7)$	$(2, 6, 7)$	$(3, 6, 6)$ $(3, 5, 7)$	$(4, 5, 6)$ $(4, 4, 7)$	$(5, 5, 5)$
16		$(2, 7, 7)$	$(3, 6, 7)$	$(4, 6, 6)$ $(4, 5, 7)$	$(5, 5, 6)$

The detailed list to $N = 24$ certainly indicated the presence of a pattern, but could I find a formula? I decided to create a new list showing the number of triangles in each group as follows:

N	$a = 1$	2	3	4	5	6	7	8	Total
3	1								1
4									0
5	1								1
6		1							1
7	1	1							2
8		1							1
9	1	1	1						3
10		1	1						2
11	1	1	2						4
12		1	1	1					3
13	1	1	2	1					5
14		1	1	2					4
15	1	1	2	2	1				7
16		1	1	2	1				5
17	1	1	2	2	2				8
18		1	1	2	2	1			7
19	1	1	2	2	3	1			10
20		1	1	2	2	2			8
21	1	1	2	2	3	2	1		12
22		1	1	2	2	3	1		10
23	1	1	2	2	3	3	2		14
24		1	1	2	2	3	2	1	12

There is certainly a predictability about $a = 1$ and $a = 2$, and after the first two non-zero entries also for $a = 3$ and $a = 4$. I was not quite sure yet about the patterns for $a = 5$ and $a = 6$, and certainly not enough entries were available for $a = 7$ and $a = 8$. Now it has already been noted that the number of triangles for N even is exactly the number for N odd, but three steps behind. But the more detailed list shows that their structures are exactly the same with the a-values out of step by one. As well, the final a-values seem to be in groups of three from $N = 6$ onwards with successive values $1, 1, 2$. Was this progress? (TIME OUT again!)

After a reflective time interval, I decided to go back to the start, but include $N = 1$ and $N = 2$ in the analysis. Obviously $T(1) = T(2) = 0$, and so the sequence for the first 24 $T(N)$ becomes

$$0, 0, 1, 0, 1, 1, 2, 1, 3, 2, 4, 3, 5, 4, 7, 5, 8, 7, 10, 8, 12, 10, 14, 12.$$

The sum of successive pairs yields

$$0, 1, 2, 3, 5, 7, 9, 12, 15, 18, 22, 26,$$

and then the sum of successive groups of three in this sequence is

$$3, 15, 36, 66.$$

The first differences of this sequence are $12, 21, 30$, while the second differences are $9, 9$. I wondered if this could be indicating a pattern for groups of six (2×3)?

In order to investigate this, I needed to introduce a notation for these groups of six. I started with $\{6k, 6k+1, 6k+2, 6k+3, 6k+4, 6k+5\}$ with $k = 0, 1, 2, \ldots$, but this introduced $N = 0$. Therefore, I settled on $\{6k-5, 6k-4, 6k-3, 6k-2, 6k-1, 6k\}$ with $k = 1, 2, 3, \ldots$.

With this notation I can denote each successive group of six in the sequence of $T(N)$ by $H(k)$, where H stands for 'hexa'. The sum of each $H(k)$ is denoted by $SH(k)$, and hence

$$SH(1) = 3, \quad SH(2) = 15, \quad SH(3) = 36, \quad SH(4) = 66.$$

Since the second differences may be the same (even though there is only a sample of two values at present), I tried a quadratic form for $SH(k)$. The first three terms in the sequence yielded

$$SH(k) = 9k^2/2 - 3k/2 = 3k(3k-1)/2$$

and this formula gives $SH(4) = 66$. The formula predicted $SH(5) = 105$. I evaluated $T(N)$ for $N = 25$ to 30 by first principles, and obtained $H(5) = \{16, 14, 19, 16, 21, 19\}$ confirming that $SH(5)$ is indeed 105. (TIME OUT! Is there any way to prove this result for all k by mathematical induction? Perhaps an exercise for later.)

At this stage the earlier discovery concerning pairs within a group of six proved to be useful. When the three successive pairs in $H(k)$ are denoted by $P1(k), P3(k), P5(k)$, I noted that the gap between $SP1(k), SP3(k), SP5(k)$ is k. Therefore

$$SH(k) = SP1(k) + SP3(k) + SP5(k) = 3SP3(k),$$

that is,

$$SP_3(k) = k(3k-1)/2 = T(6k-3) + T(6k-2)$$
$$SP_1(k) = k(3k-1)/2 - k = T(6k-5) + T(6k-4)$$
$$SP_5(k) = k(3k-1)/2 + k = T(6k-1) + T(6k)$$

This is one simple equation for each pair in our strange sequence. The second equation could come from the pattern of differences within each pair, if there is one. Denoting these differences by $DP1(k), DP3(k), DP5(k)$, I observed that

$$DP_1(1) = 0, \quad DP_3(1) = 1, \quad DP_5(1) = 0$$
$$DP_1(2) = 1, \quad DP_3(2) = 1, \quad DP_5(2) = 1$$
$$DP_1(3) = 1, \quad DP_3(3) = 2, \quad DP_5(3) = 1$$
$$DP_1(4) = 2, \quad DP_3(4) = 2, \quad DP_5(4) = 2$$
$$DP_1(5) = 2, \quad DP_3(5) = 3, \quad DP_5(5) = 2$$

Therefore it appears that $DP_3(k) = k/2$ (when k is even) and $(k+1)/2$ (when k is odd), while $DP_1(k) = DP_5(k) = DP_3(k)$ (when k is even) and $DP_3(k) - 1$ (when k is odd).

Summarising the discoveries after solving the two simple systems of equations for each case yielded:

(k even)	(k odd)
$T(6k-5) = 3k^2/4 - k/2$	$T(6k-5) = 3k^2/4 - k/2 - 1/4$
$T(6k-4) = 3k^2/4 - k$	$T(6k-4) = 3k^2/4 - k + 1/4$
$T(6k-3) = 3k^2/4$	$T(6k-3) = 3k^2/4 + 1/4$
$T(6k-2) = 3k^2/4 - k/2$	$T(6k-2) = 3k^2/4 - k/2 - 1/4$
$T(6k-1) = 3k^2/4 + k/2$	$T(6k-1) = 3k^2/4 + k/2 - 1/4$
$T(6k) = 3k^2/4$	$T(6k) = 3k^2/4{+}1/4$

These formulae have not been proved for general k, but they are correct for any N that I have tried. They give $H(5) = \{16, 14, 19, 16, 21, 19\}$ as obtained earlier, and $H(6) = \{24, 21, 27, 24, 30, 27\}$ which can be verified by first-principle evaluations, confirming the gap of three between the sums of successive pairs. Therefore, given any N, it is now appears possible to obtain T without going through the long method of evaluation by first principles. Not that it is really important to know how many Diophantine triangles can be constructed with, say 99 matchsticks, but the procedure in arriving at the final formulae should emphasise to students the importance of persevering with various trials. In the current climate of short-term concentration via technological entertainment and communication, academics and teachers need to provide examples of problems that are neither impossible to solve nor able to be solved in a short time. The above example is just one of those.