Sum-Product Graphs: Diamonds and other interesting structures

Nila Cibu (UC Berkeley) - *Pure Mathematics & Statistics major*
Prof. Eric Babson (UC Davis)- *Supervisor*

Abstract

We delve into how sum-product games motivate the definition of a sum-product graph. We see that these graphs are largely dominated by cycles and which leads to questions about their graphical structure, such as finding their genus. We also look at generating $K_{2,n}$ subgraphs found within sum-product graphs and what they have to say about the overall structure of sum-product graphs.

Introduction

In our summer research, our aim was to go find solutions for the sum-product game and in the process of doing so, we were led to a beautiful, graphical representation of the game. In this paper, we leave behind the game itself to instead look at the properties of the sum-product graphs generated by the game.

This paper aims to parse the interconnected web of sum nodes and product nodes found within the central region of these graphs. We look at why this is the case, as well as identify graph properties that may help pin down the states of sum-product graphs at various sizes, like genus or important subgraphs.

1 Context

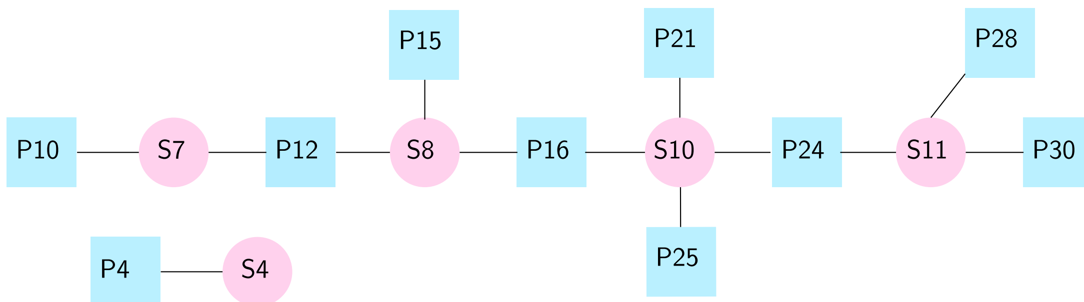
Sum-product graphs are a result of a game outlined in a 1979 column by Martin Gardner in Scientific American[1].

For the game, a pair of positive integers is chosen within the range of 2 to n, not necessarily distinct. Then one person is given the product while the other is given the sum. Both players then proceed to try and guess the pair of numbers they were given. At every turn, they reveal whether or not they know the two integers. The game ends only when one of the players reveals they know the two integers. It turns out this game can be represented as a graph.

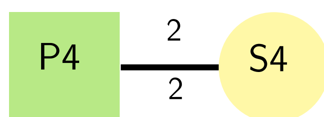
[1] Here is a link to a discussion on this game, courtesy of Torsten Sillke.

1.1 The games

Before looking at the graphs and their properties, we should first look at a few example games where S is the player who knows the sum and person P is given the product. These examples will be referring to the following subset of the graph where $n = 8$. [2]



Example. Suppose the integer pair given is **(2,2)**.



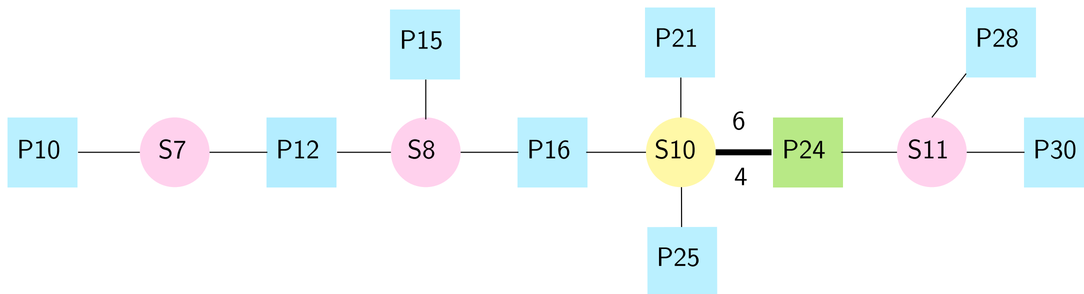
Here S and P immediately know what each other's numbers are since the only way to get either the product or sum of 4 is via the integer pair (2,2). As a result, the dialogue is composed of a single line:

S: *Yes, I know what your number is.*

S wins here since the person with the sum always starts first. This was a rather short example, but our next example will be a bit longer.

Example. Now suppose our integer pair is **(4,6)**. S has the number 10 marked in yellow and P has the number 24 marked in green.

[2] Note that in general, sum nodes will be circular and/or pink whereas product nodes will be either light blue and/or square. Sum node labels will always begin with an S and product node labels will always begin with a P. In the following examples, green and yellow have been used to highlight the nodes of interest.



Unlike the previous case where the sum was only connected to one product node, here S does not know which of the 4 possible product nodes (P15, P21, P24, P25) P could have.

S: *No, I don't know what your number is.*

Similarly, P also does not know which of the two sums, 10 or 11, S could have. For both sums, S10 and S11, the product node P24 is connected to, the sums would be unable to tell which of the many products nearby is P's number.

P: *No, I don't know what your number is.*

Since P also doesn't know what number person S has, S knows that P's number cannot be either 21 or 25. If P were to have either of these numbers they would immediately know what the number of the sum is and would have said as such. However, there are two more product nodes connected to the S's sum node these being S16 and S24. If person P had either of these numbers they would give the same response.

S: *No, I don't know what your number is.*

At this point, the tides of the game turn in P's favour. P notices that if S were to have the number 11, they would have said "Yes" in the previous round.

The sum node S11 is only connected to one product node that is connected to more than one sum node, that product node being P24. For all the other product nodes S11 is connected to, the product person would immediately determine that their partner's sum node is S11. Since P knows that person S was unable to deduce this, it means S cannot be 11 and must have the number 10.

P: *Yes, I know what your number is.*

In this game, P has won.

1.2 Observer

There is a third role in this game, the observer. The observer only gets to see the conversation between S and P and nothing else.

In the first example, we gave with the pair (2,2) the observer would see just the singular "Yes". For the next example, the observer would see "No", "No", "No", and "Yes".

Notation. We can write these answers instead as just Y (0 N before Y) and NNNY (3 N before Y).

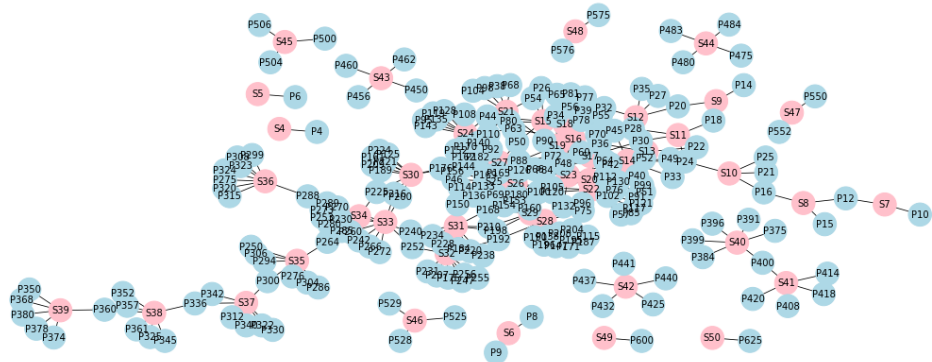
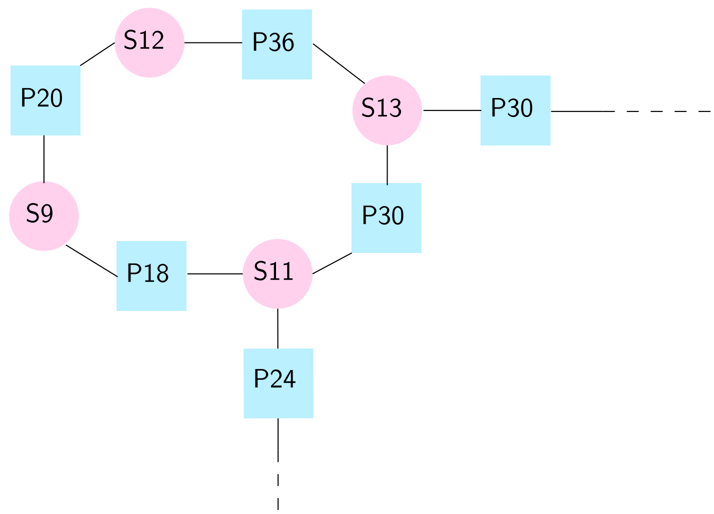


Figure 1: Graph for when $n = 25$, i.e. $G(25)$. Notice how cluttered the diagram is despite n being relatively small.

A question we can ask is if the observer can identify the original integer pair given to S and P with just their responses. However, this is not a question that will be answered in this write-up. The observer and their role will be covered in more detail by Mariam Abu-Adas and by Yuanyuan Shen in their write-ups.

2 Cycles

It turns out, although for small n like $n = 8$ the graphs are quite simple; every edge (which corresponds to an integer pair) starts a game that has a definite end. However, once n grows large enough, every game generated by an edge might no longer end. This is because of the cycles in these graphs. The first cycle appears when $n = 10$ and it has been displayed below.



When the edge starting off the game is caught in a cycle, the game has no end since the continuous "no" answers are possible with any connected node within the cycle. this is not necessarily true for edges like the one between S11 and P24, only edges like those between S11 and P18 would do this.

As n grows larger, these graphs become dominated by cycles, with most sum nodes eventually connecting into a cycle. Proof of this will be shown in Theorem 2.1.

The generated graphs for even $n = 25$ are cluttered; it is difficult to discern the centre of the graph, $G(25)$ as seen in Figure 1.

Notation. $G(n)$ is the graph generated when the max integer value is n.

As mentioned before, it can be shown that almost every sum node will eventually connect into a cycle, more explicitly:

Theorem 2.1. *In a graph $G(n)$, every sum node with value greater than or equal to 14 will eventually*[3] *connect into a cycle for a sufficiently large n.*

Proof. There are two possible cases that will need to be looked at separately. We will use induction in both to show how sum nodes link [4] to smaller valued sum nodes.

Base case: In $G(12)$ (refer to Figure 2) we will see that every sum node, $s_0 \geq 14$ will be 2 steps away from two different sum nodes with lower values. Additionally, since the path P24 to P10 is stable and cannot change (alongside other stable, floating paths like S2-P2, P8-S6-P9, S5-P6)[5], the only sum nodes

[3] in the sense of increasing the value of n

[4] Using the term connect or link is not technically correct since no two sum nodes can be neighbours, but for the sake of brevity in this proof for two sum nodes to be connected/linked will mean that the two sum nodes share a product node neighbour.

[5] This is because the numbers involved are so small that you can manually check all of their possible connections and see that they remain unchanged

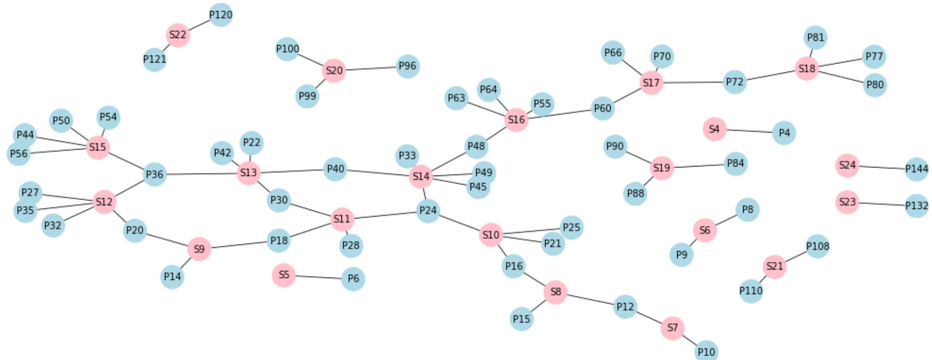


Figure 2: $G(12)$, base case for Theorem 2.1

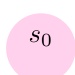
with a value smaller than the sum node, s_0, can connect to are those that are part of a cycle.

Inductive hypothesis: Every sum-node $s* < s_0$, given sufficiently large n is connected to a cycle.

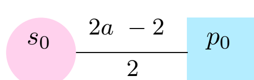
All we need to do is ensure that every sum node will eventually have two other sum nodes with smaller values two edges away.

Case 1: Sum node's value is even. Let the sum node $s_0 = 2a \geq 14$ where $a \in \mathbb{Z}$ and let $n = n_0$ for this graph, $G(n)$.

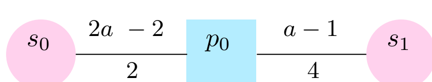
$n = n_0$



$n \geq 2a - 2$



$n \geq 2a - 2$



If we increase n such that $n \geq 2a - 2$ then we can guarantee that n is connected to the product node, $p_0 = 2(2a - 2)$ via the edge $(2, 2a - 2)$. This product is then connected to a sum node with a lower value, s_1, via the edge $(a-1, 4)$. We need to check that $s_0 > s_1$ is indeed the case. From the assumption

$s_0 \geq 14$:

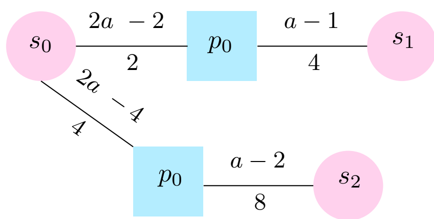
$$a > 3$$

$$2a > a + 3$$

$$2a > a - 1 + 4$$

$$s_0 > s_1$$

$n \geq 2a - 2$



s_0 is also connected to a product node, p_1 via the edge $(2a - 4, 4)$. Using the same idea from before we see that p_1 is connected to another sum node s_2. Once again let us check whether $s_0 > s_2$:

$$a > 6$$

$$2a > a + 6$$

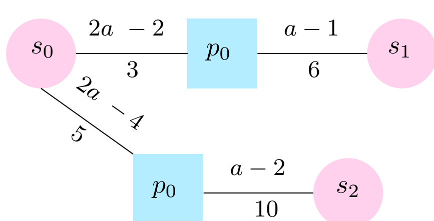
$$s_0 > s_2$$

Now we need to check the odd case.

Case 2: Sum node's value is odd. Let the sum node $s_0 = 2a + 1 \geq 14$ where $a \in \mathbb{Z}$ and let $n = n_0$ for this graph, $G(n)$.

Since the main proof idea is identical, we will only be showing the diagrammatic proof of smaller nodes that s_0 will connect to given n large enough as well as checks to make sure the bound $s_0 > 14$ is enough to ensure extra two sum nodes.

$n \geq 2a - 2$



Checking $s_0 > s_1$:

$$a > 4$$

$$2a + 1 > a + 5$$

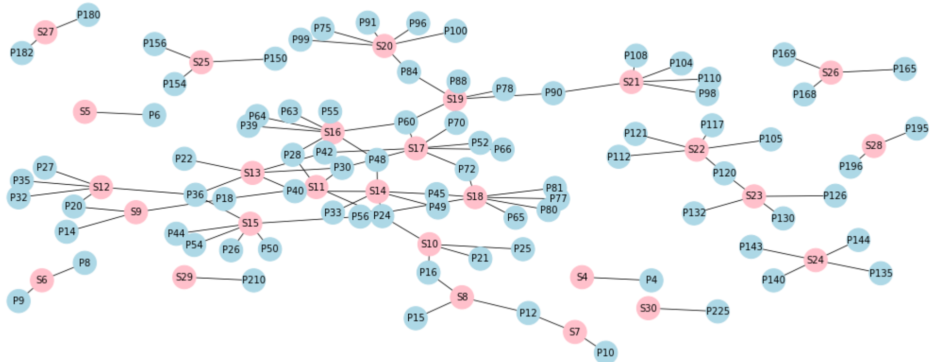


Figure 3: $G(15)$, the first non-planar sum-product graph

$$s_0 > s_1$$

Checking $s_0 > s_2$:

$$a > 7$$

$$2a + 1 > a + 8$$

$$s_0 > s_2$$

This means every sum node is "connected" to (or two edges away from) two other smaller valued sum nodes. $\square$

3 Graph genus

There are various ways to understand what exactly is going on within these graphs and one such way is to see whether or not you can draw the graph on a sheet of paper i.e. check whether it is a planar graph. We can then broaden the scope of our investigation and look for the genus of these graphs in general.

3.1 Planar graphs

The more technical definition of what it means to be a planar graph is:

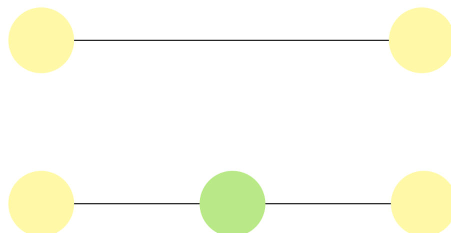
Definition 3.1. Planar graph is a graph that can be embedded in the plane.

Although the first few graphs like $G(10)$ are all planar, however for $G(15)$, as seen in Figure 3, this is no longer true. Note that because of the rules of the original game, every graph $G(n) \subset G(n+1)$, which means that if $G(15)$, then every following sum-product graph will also be non-planar.

To prove that $G(15)$ is the first non-planar graph we need Kuratowski's Theorem.

Theorem 3.2. ***Kuratowski's Theorem*** *A finite graph is planar if and only if it does not contain a subgraph that is a subdivision of the complete graph* K_5 *or the complete bipartite graph* $K_{3,3}$ *(utility graph)[1].*

Here they mention the term **subdivsion**. For an edge that simply means taking the edge and then dividing it into two new edges with a vertex connecting them in middle, as shown below, the green circle representing the new vertex.



Definition 3.3. **Subdivision** of a graph G is a graph resulting from subdividing the edges of G.

Definition 3.4. The reverse of subdividing an edge is called **smoothing**.

Note that the addition of leaves or removal of leaves (vertices with degree 1) will not affect a graph's genus. With these tools in hand we can finally begin proving this statement.

Proposition 3.5. $G(15)$ *is the first non-planar sum-product graph.*

Proof. There are two things we need to show here:

- $G(14)$ is planar
- $G(15)$ is non-planar

Showing that $G(14)$, Figure 4, is planar is relatively straightforward. All we need to show is that $G(14)$ can be drawn without any edge crossings. That being said these diagrams are still difficult to parse due to the number of vertices present. Upon repeatedly smoothing out $G(14)$ and removing its leaves we get the graph:

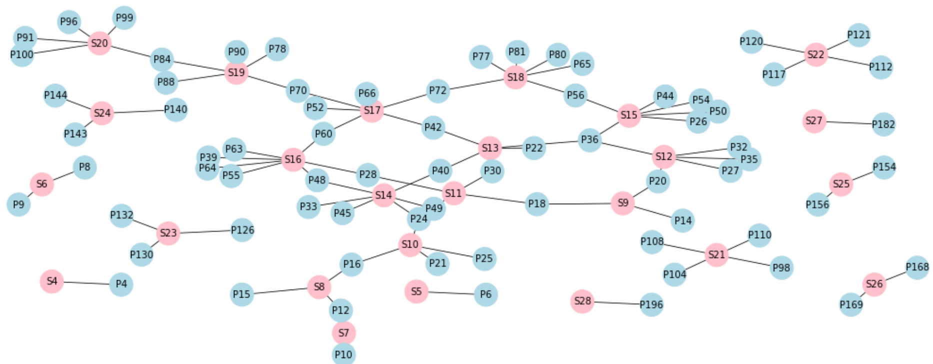
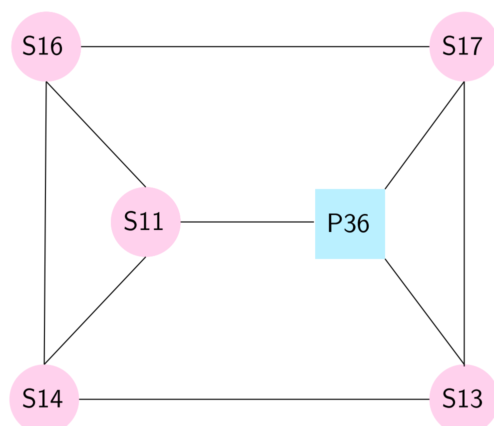


Figure 4: G(14), the last planar graph



This clearly has no edge crossings and is therefore planar. We can do a similar thing for G(15), the smoothed out graph seen in Figure 5.

Using the smoothed G(15) we find the subgraph shown below. This subgraph is a subdivision of the utility graph since we can smoothen out the node S16 to get a straight edge from S11 to S17.

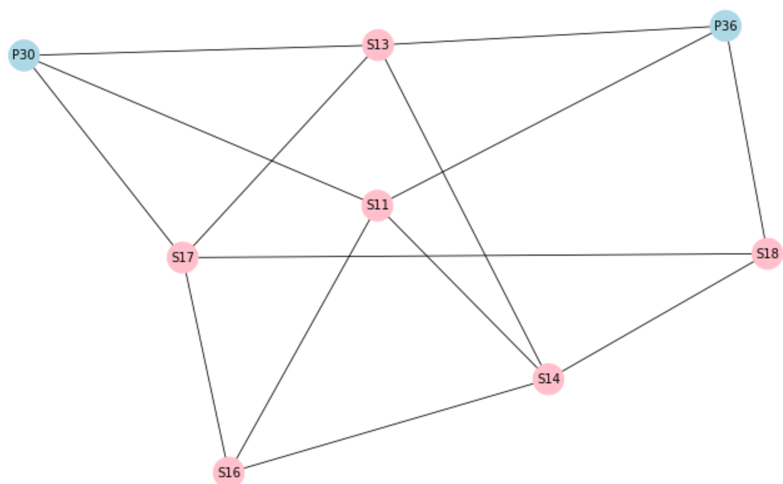
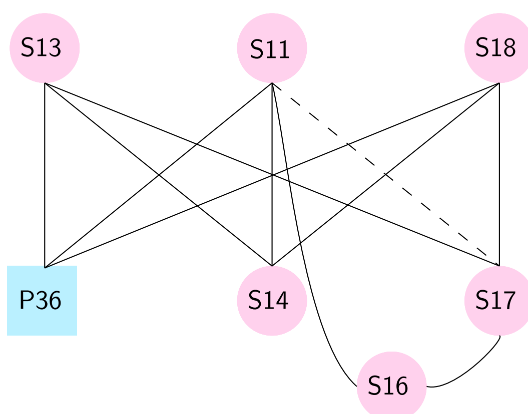


Figure 5: Smoothened G(15)



By Kuratowski's Theorem, G(15) must be non-planar. □

3.2 Genus of sum-product graphs

Checking for planarity is a special case of checking a graph's genus.

Definition 3.6. A graph G's **genus** is the minimal integer n such that it can be embedded on a sphere with n handles.

Notation. $\gamma(G)$ is the genus of graph G.

This means a planar graph is a graph with genus 0 since it can be drawn on the plane without any crossings. A natural question to then ask is what the genus of an arbitrary sum-product graph $G(n)$ is. This is not quite as simple as proving certain sum-product graphs are non-planar. Finding a graph's genus is an NP-hard problem [2]. While there are general bounds, as seen in equation 3.1 [2], they require knowing the number of vertices v and edges e. Determining these values for sum-product graphs is not so simple since counting the number of product nodes requires knowledge of primes and how they are distributed.

$$\gamma(G) \geq \left\lceil 1 - \frac{v}{2} + \frac{e}{4} \right\rceil \tag{3.1}$$

Another option would be to identify subgraphs whose genus is known such as the complete bipartite graphs, $K_{m,n}$.

$$\gamma(K_{m,n}) = \left\lceil \frac{(m-2)(n-2)}{4} \right\rceil \tag{3.2}$$

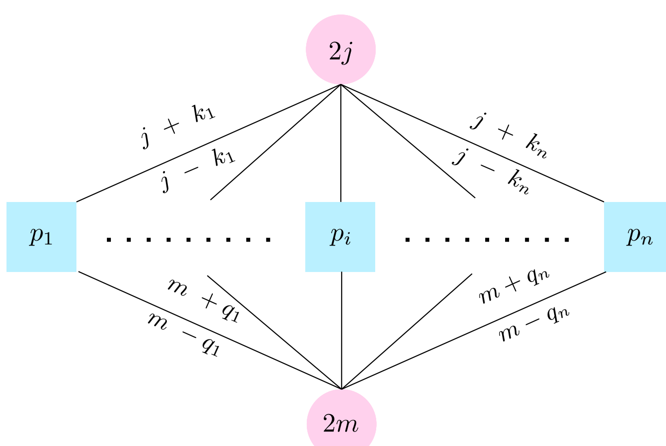
4 Diamonds

Trying to identify or construct possible complete bipartite subgraphs leads one to the world of Diophantine equations. This can get complicated when $m, n \geq 3$, but when $m = 2$ it is possible to identify such structures in these graphs; m and n here refer to the number of sum nodes and product nodes respectively.

It turns out all possible sum-product $K_{2,n}$ subgraphs can be generated by using a simple algorithm.

4.1 Generating $K_{2,n}$ subgraphs

To find this algorithm we need to first look at what numbers a possible $K_{2,n}$ subgraph would be composed of. From there we can work backwards to find what our generating numbers should look like and what properties they should have.



Here we have two sum nodes S_1, S_2 and we let them equal $2j$ and $2m$[6] respectively. Note that j and m are not necessarily integers, they could instead be positive multiples of $\frac{1}{2}$. Similarly, k_i and q_i may also be $\frac{1}{2}$. Additionally, since this is a sum-product graph $j > k_i$ and $m > q_i$ for all i.

Looking at the product node p_i,

$$\begin{aligned} p_i &= (j - k_i)(j + k_i) = (m - q_i)(m + q_i) \\ &= j^2 - k_i^2 = m^2 - q_i^2 \end{aligned}$$

Rearranging these equations we get

$$j^2 - m^2 = k_i^2 - q_i^2 \quad \text{for all } i$$

Let

$$\begin{aligned} z &= j^2 - m^2 \\ r &= j + m \\ s &= j - m = \frac{z}{r} \\ t_i &= k_i + q_i \end{aligned}$$

Notice that these new variables, z, r, s, t_i can be multiples of $\frac{1}{4}$. Finally, we define our generating variables.

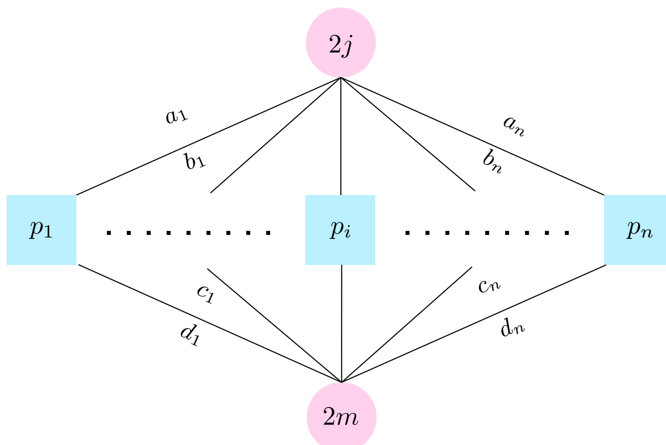
$$\begin{aligned} A &= gcd(r, s, t_1, \cdots, t_n) \\ B_i &= gcd(\frac{r}{A}, \frac{t_i}{A}) \\ C_{ij} &= gcd(\frac{t_i}{A}, \frac{t_j}{A}) \quad \text{for } i \neq j \\ D &= \frac{r}{A \prod\limits_i B_i} \\ E_i &= \frac{t_i}{B_i \prod\limits_j C_{ij}} \end{aligned}$$

These variables have all been constructed so that the previous set of variables can be gotten back. To begin generating these subgraphs we can assign values to A, B_i, C_{ij}, D, E_i. There are some restrictions on the values we are able to assign to these elements because of how we defined them.

- A can be any number
- for all i, j B_i, C_{ij} cannot have a shared common factor

[6]This m is different from the m used to denote the number of sum nodes in the bipartite subgraph.

- D, E_i need to be co-prime
- Check $D \prod_j B_j > B_i E_i \prod_j C_{ij}$ for all i, otherwise reorder them
- Ensure $B_i E_i \prod_j C_{ij} \neq B_j E_j \prod_i C_{ij}$ and $B_i D E_i \neq B_j D E_j$ for $i \neq j$
- Ensure $B_i E_i \prod_j C_{ij} \neq B_i D E_i$ for all i



Now using these A, B_i, C_{ij}, D, E_i we can get back our original edge pairs.

$$r = AD \prod_i B_i$$
$$t_i = AB_i E_i \prod_j C_{ij}$$
$$z = AD \prod_i (B_i E_i \prod_j C_{ij})$$
$$s = \frac{z}{r}$$
$$u_i = \frac{z}{t_i}$$

$$a_i = r - s + t_i - u_i$$
$$b_i = r - s - t_i + u_i$$
$$c_i = r + s + t_i + u_i$$
$$d_i = r + s - t_i - u_i$$

By doing some small-scale algebra we are back to the original diamond we started with.

4.2 Diamonds

These structures can also inform us about other aspects of a sum-product graph's structure, especially in the case where $n = 2$.

Definition 4.1. A **diamond** refers to a $K_{2,2}$ subgraph.

Identifying where these diamonds are located, particularly those with the highest valued sum nodes seems to point to the boundary of where cycles begin to end and where edges start to generate games that are no longer infinite. We have written up code to help identify all the diamonds and visualise them in these sum-product graphs and encourage you to play around with the code yourself.

5 Future

This research has led us to some unexpected discoveries, diamonds being one, however, this is far from the end. There are still a lot of unanswered questions about the genus of a sum-product graph. Although we were able to identify $K_{2,m}$ subgraphs, we can only do so in the case where $m = 2$ not n. It would be interesting to see if other types of subgraphs could be found; contained within these vast, sum-product graphs. There are other graph properties worth exploring such as their radii or their diameters.

References

[1] W. Klotz, "A constructive proof of kuratowski's theorem." [Online]. Available: https://www.researchgate.net/publication/256078009-mathematics-

[2] A. Perez, "Determining genus of a graph," 2009. [Online]. Available: https://www.yumpu.com/en/document/read/49462442/1-determining-the-genus-of-a-graph-harvard-college-mathematics-