Explorations into Monomer-Dimer Tilings of Planar Regions

Zachary Hall
University of Wyoming, zhall@uwyo.edu

Follow this and additional works at: http://repository.uwyo.edu/honors_theses_16-17
Part of the Discrete Mathematics and Combinatorics Commons

Recommended Citation
http://repository.uwyo.edu/honors_theses_16-17/62

This Honors Thesis is brought to you for free and open access by the Undergraduate Honors Theses at Wyoming Scholars Repository. It has been accepted for inclusion in Honors Theses AY 16/17 by an authorized administrator of Wyoming Scholars Repository. For more information, please contact scholcom@uwyo.edu.
Explorations into Monomer-Dimer Tilings of Planar Regions

Zachary Hall

Dr. Bryan Shader

University of Wyoming Department of Mathematics
Overview

It was a blessing to be able to take part in academic research these past two semesters with the help of Dr. Shader, the University of Wyoming Honors Program, and EPSCoR. From the start I knew that this experience was going to be a challenging yet fulfilling one, and I’m glad that I embraced the opportunity to further my academic endeavors and problem solving abilities through this program. I would surely do it again many times if I could.

The fellowship itself was incredibly helpful in aiding me throughout the semester. Although the topic itself was enough to keep my attention, being employed by the University gave me an extra incentive to work. Moreover, it gave me more of a sense of purpose than if I were just doing the work on my own. I felt as if I were doing the research on behalf of the institution, and I feel that I was provided with the resources necessary by them. The EPSCoR staff was always incredibly helpful, understanding, and timely in everything that they did. Even though most of my research was done on a computer or with a paper and pencil, I was provided with necessary resources when appropriate. I don’t have any suggestions for the program, and I am hoping that I will be able to take part in another fellowship again next year.

In relation to the research itself, I attained a massive amount of knowledge this year that I wouldn’t have had previously. Although my research seems very specific, I was able to test and enhance my problem solving abilities and work ethic. I was also able to more precisely investigate and broaden my skills with computer programming, probability distributions, expected values, and matrix theory. I also was able to strengthen my interpersonal communication abilities.

I feel incredibly thankful to Wyoming EPSCoR and Dr. Bryan Shader for providing me with this opportunity. I will continue to use what I have learned for years to come.
Research

Throughout each semester, my main work was in dealing with two dimensional rectangular regions and tiling them using monomers and dimers. A monomer is a 1 by 1 tile (small square) and a dimer is a 2 by 1 tile (like a domino). These regions, usually N by N, 2 by N, or 1 by N, were each investigated in different ways and for different purposes, which will be explained below. Figure 1 to the right shows a 6 by 6 region, with 15 dimers and 6 monomers. For small regions, I was able to do a lot of work by hand using probabilistic models. However, when the models became larger, computer programming was imperative.

Most of the analysis performed during this year was done using computer programs that I developed with the help of Dr. Shader. These programs were developed in Sage Math, a Python-based language. In order to develop code, a lot of theory was explained to me by Dr. Shader, or simply written down, and then developed into code. This program was used for two purposes: calculations and simulations. Many calculations that had to be found were of permanents and determinants of specific matrices, which would be incredibly difficult and lengthy to do by hand. Our simulations were ones that could be run thousands of times within a few minutes, and would provide an overall outlook of some potential areas to look at. Some specifics on both of these uses of the computer programming will be explained in this report. Supporting material for this research, though not explicitly referenced in this paper, can be found in the works cited page.
Year One

I began my research by figuring out how to enumerate the number of possible tilings of a 4 by 4 region that has two monomers and thus seven dimers. I was then able to extend this to a 6 by 6 and 8 by 8 region. Part of this was done by hand, figuring out different possible placements of the two monomers, both of which could not be one the same color if our region was colored like a checkerboard. It was discovered that there were 552 possible monomer-dimer tilings for a 4 by 4 region with the prescribed two monomers and seven dimers. I also discovered the probabilities of a random tiling in regards to where the monomers were most likely to be placed: for this region, monomers were most likely to be in the corners, then on the sides. The least likely place for the monomers were on the interior of the 4 by 4 region. We then investigated a 6 by 6 region and an 8 by 8 region with two monomers and the rest dimers. It took more time, but we found that there were 363,536 and 1,760,337,760 possible monomer dimer tilings for the regions, respectively.

We also found that the most likely location of a monomer was in the corners once again for both of these regions. Figure 2 to the right shows the ranking of likely locations for a 6 by 6 region. As we can see the corners are the most probable locations, while the spots adjacent to the corners are the least probable locations. It was then a goal of ours to mathematically prove for all regions that the corner was the most likely location for the dimer.
for all uniform tilings of any $N \times N$ region. Attempts to prove this by known combinatorial methods were unsuccessful.

After this task was completed, Dr. Shader and I began looking into developing a system to create a random, uniform tiling of dimers and monomers that were created by starting with a tiling of $nxn$ monomers, and tiling them with one another based on whether they both chose one another. If so, they would in turn become monomers. Their probability of choosing one another was one divided by the number of adjacent monomers to them. Thus, if two isolated monomers existed, they would surely pair with one another. We then did extensive simulations with these on 4x4, 6x6, and 8x8 regions and looked to make observations.

We would run these simulations about 1000 times. We found that between 4x4, 6x6, and 8x8, the mean number of times that it would take to converge into a “frozen” configuration was about 3.5, 4.3 and 4.8 respectively. We don’t have enough information to make a conjecture regarding whether this converges to a certain number for larger and larger configurations, but we would assume that they would continue to increase as $n$ increases. Furthermore, the standard deviations for these number of steps to convergence for these tilings, regardless of $n$, was around 0.85. This is an interesting observation, and suggests that the amount of variance in the time to converge for each of these is very similar if not identical. Also, there was no definitive pattern of an increase in the standard deviation for the number of times it takes until the configurations froze.

An additional observation that we made was in regards to the average (expected) number of dimers for a configuration. This number obviously increased as $N$ increased. When comparing these averages to $N^2$, the ratio was nearly identical for all three regions, approximately .4578. It then became a goal of ours to find out what this number meant. Unfortunately, we also found
that these distributions that were found were not uniform, and thus did not generate “random”
tilings in the same way that our other method would. Still, bonding tiles has applications in
chemistry and genetics.

In further investigating this ratio, we began looking at a 1 X N region. First, we were able
to use a computer program and the intermediate possible tilings to create these and find the
average number of dimers and average number of iterations. However, for larger N, this became
difficult. Because of this, we began looking at a way to generalize and justify an expected
number of dimers by hand, and find a formula, either recursive or explicit.

This was done using Markov chains and one-step probabilities. It was also the most
intricate and intensive part of the entire research project. However, I was able to come up with
the following recursive formulas:

\[ F_N = 1 + \frac{2^{N-2}}{2^{N-1}} E_N + \frac{1}{8} F_{N-3} + \frac{N-2}{2^{N-1}} (F_0 + E_{N-2}) + \sum_{k=1}^{N-4} \frac{N - k - 2}{2^{N-k}} (F_k + E_{N-k-2}) \]

\[ E_N = \sum_{k=0}^{N-3} \frac{1}{2^k+1} (F_{N-k-2} + E_k) + \frac{1}{2^{N-2}} (F_0 + E_{N-2}) \]

\( E_N \) indicates the expected number of dimers for a 1 X N region. \( F_N \) is an intermediate value used
to help calculate this, and its explanation is not within the scope of this report. As we can see
above, both of these are necessary in calculating the expected number of dimers, as previous
calculations for each are used within each calculation. After running this in SageMath, we found
that \( E_{20000} / 20000 = 0.433446533 \). This ratio is a bit different from the ratio for N by N region.
However, it is consistent for these 1 by N regions. For example, this value for the 1 by 10,000
region is the same to 5 decimal places. Because the most this could possibly be is 0.5, this value
is surely bounded as N approaches infinity. However, the exact bound on this was not discovered.

Closing out this first year, my place in this research was in attempting to find a closed, non-recursive formula for the expected number of dimers for a 1 by N region using the aforementioned method. Various questions were left open. I had made some progress with this, but a couple of things were a bit difficult, and I was forced to abandon my efforts early in the next school year. However, patterns seemed to be developing, and I think that with a little work I will be able to figure this out eventually. I also had hoped to derive this mystery ratio and either find it to be something in existence or a completely new ratio that was not well-known. Another hope was to investigate the sequence of numbers that are my $E_N$ and look to see if this is a sequence that already exists. I was unsuccessful in my attempts. Still, I was very pleased to have found a formula for these. Extending this into something like a 2 by N or N by N region would be ideal in the future, but likely would be extremely difficult.

**Year Two**

Following my attempts over the summer, Dr. Shader and I decided to take things in a new direction for the following year. Our initial inclinations included using further matrix theory in order to continue to enumerate tilings. However, instead of doing this for a prescribed number of monomers and dimers, we aimed to discover the total number of tilings possible regardless of configuration.

We were able to do this and more with a bit of work. In order to do so, characteristic matrices were developed for various planes. These characteristic matrices, denoted as $A_M$,
describe monomer dimer tilings of M by 1 regions. Each entry in this $2^M$ by $2^M$ matrix represents the number of ways to extend a tiling of an M by N region with a given right hand boundary to an M by (N-1) region with another given right hand border. These configurations are represented by the row and column, respectively. The first entry, then, represents the total number of ways to tile this region. Even more interestingly, if we take $A^N_M$, we will receive a characteristic matrix for tilings of an M by N region. The first entry, again, describes the total number of tilings possible. These matrices were created using numerical methods through coding in SageMath.

Beyond simply enumerating the number of configurations that are able to be constructed by monomers and dimers in a particular region, these characteristic equations have an additional property. Before describing this, I will explain how this was discovered. The Cayley-Hamilton Theorem in matrix theory states that every square matrix satisfies its own characteristic equation. In other words, the $k^{th}$ power of a k by k matrix is a linear combination of the $(k-1)^{th}$ power of that matrix down to the $0^{th}$ power of that matrix, also known as the identity matrix, $I$. Because our characteristic matrices’ first entry represents the number of possible tilings of each type of region M by N region for $A^N_M$, we can conclude that the number of tilings of an M by N region satisfies a linear recurrence, and can thus be constructed from matrices of size M by (N-1) and smaller. For this to work, N must be greater than or equal to $2^M$. However, it may also be possible to factor this characteristic equation such that a linear recurrence of a lower order is allowed, lowering how great N needs to be to satisfy this recurrence.

Specifically, in order to employ this technique, we wanted to look at an 8 by N region and its tilings. This required creation of a 256 by 256 matrix, $A_8$. Once this matrix is created, is it possible to raise it to a power N so that the first entry of the $A^N_8$ matrix represents the number of possible tiling of an 8 by N region. However, when we utilize the Cayley-Hamilton theorem, we
are able to create a recursion. Furthermore, this can be factored into a lower order polynomial, of degree 136 (instead of 256), where the number of tilings of an 8 by N region satisfy the recursion. This recursion is seen in the Appendix, where $T_N$ represents the number of monomer-dimer tilings of an 8 by N region. It should also be noted that a ratio appears to exist for a enumeration of an 8 by N region in comparison to the enumeration of an 8 by (N-1) region:

$$\lim_{n \to \infty} \frac{T_N}{T_{N-1}} \approx 170.8771576.$$ 

This ratio seems to converge relatively quickly, and is likely the largest real root of the characteristic equation for our matrix $A_8$.

Although it may seem futile to a non-mathematical individual, this result is rather profound. The fact that we are able to create an equation in the form of a linear recursion for the number of monomer-dimer tilings of a region is peculiar. Furthermore, it creates an opportunity to save a lot computing power, as taking a 256 by 256 matrix to extremely high powers can be computationally expensive. Meanwhile, storing these numbers of tilings and entering them into a recursion is very inexpensive, and is able to be done very quickly. This is why this clever technique is so important.

This result is not only profound, but it may be original. Although recursive formulas for M by N regions have been established (online and otherwise) in various encyclopedias, we were unable to find any for M greater than or equal to 8. With this technique, we were able to find a potentially unique result not found by any others. Moreover, we created a way to find recursions for any M greater than 8, as our code is very flexible in its structure, able to adjust for various sizes of M (originally this was not the case, but we made more discoveries -- see below).

Moving forward, it was in our interest to not only characterize these regions’ number of tilings, but the compositions of their tilings as well. In order to do this, we were able to use these
same characteristic matrices in order to discern between vertical and horizontal dimers, and thus enumerate how many different tilings were possible using various combinations of vertical and horizontal dimers. This allows us to characterize our tilings with an increased degree of detail.

To demonstrate this change, observe the two matrices below. The first represents our original characteristic matrix, $A_3$, for a 3 by 1 region. The second is updated to reflect vertical and horizontal dimers:

\[
\begin{bmatrix}
3 & 2 & 1 & 1 & 2 & 1 & 1 & 1 \\
2 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

\[
\begin{bmatrix}
2v + 1 & hv + h & h & h^2 & hv + h & h^2 & h^2 & h^3 \\
\nu + 1 & 0 & h & 0 & h & 0 & h^2 & 0 \\
1 & h & 0 & 0 & h & h^2 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
\nu + 1 & h & h & h^2 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & h & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]
These two matrices contain the same information in different forms. As one can observe, the coefficients of each entry in the updated matrix add up to the entry in the original matrix. This is because each combination of v’s and h’s represents one tiling possible with that combination vertical and horizontal dimers. For example, the first entry of 2v+1 indicates that a 3 by 1 region can be tiled in two different ways with one vertical dimer and one way with no dimers (monomers only). This is a total of three times. Other examples would include hv, which describes a possibility of one vertical and one horizontal, and v²h³, which describes a possibility for two vertical dimers and three horizontal dimers. Then, when one takes the updated matrix to a power, it can see not only the number of ways to tile the region, but also the number of ways to tile region with a specific set of monomers and dimers.

This code is also flexible, and can be used for any region of any size (subject to computing power restrictions). This is due to a discovery we made while creating these characteristic matrices. At a low k, finding the updated Aₖ is somewhat easy, and can be done by hand. Fortunately, at low k, we were able to observe a relationship between Aₖ and the characteristic of the two previous regions, where M = k-1 and k-2. The relationship is described below:

$$A_k = \begin{bmatrix} A_{k-1} & h * A_{k-1} \\ A_{k-1} & 0 \end{bmatrix} + \begin{bmatrix} v * A_{k-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In the above equation, for the first matrix, 0 represents a matrix of the same number of rows and columns as Aₖ₋₁ ( # of rows = # of columns = 2ᵏ⁻¹). In the second matrix, 0 represents a matrix of the same number of rows and columns as Aₖ₋₂ ( # of rows = # of columns = 2ᵏ⁻²). These block matrices described above show that we can construct Aₖ using just the previous two characteristic matrices. Furthermore, this technique can be extended to our original characteristic
matrices, which don’t include v’s and h’s. Here, we can replace v and h with 1, and this recursion of characteristic matrices will remain true. Because of this discovery, we are able to create a linear recurrence for M of size greater than 8, and also characterize any M by N region within our computing abilities.

The final work that Dr. Shader and I completed related to the work done at the end of Year 1. Previously, we had looked at ways to dynamically pair monomers with one another to create dimers in order to eventually get a “frozen” formation of monomers and dimers where monomers were isolated. This was first done using a simulation that paired monomers together and eventually froze, stored this information, and started over, storing the information and then presenting it (e.g. mean number of dimers, mean number of iterations, standard deviation, etc.). This was then taken into a more theoretical realm, where a two-part recursion was found for 1 X N regions and the number of dimers possible.

Still, we hoped to extend our techniques here. The main hope was to once again find out the expected number of dimers for any region. In order to do this, we planned to create probability distributions for any M by N region. These would be in matrix form, and would each entry would be the probability of transitioning from one configuration (represented by the row) to another configuration (represented by the column). Each terminal, or “frozen”, configuration would be the last rows and columns in the matrix.

For small M and N, this is not particularly difficult by hand. We must create every possible tiling of the region and then use these tilings to create probability distributions. For instance, for a 2 by 2 region, the corresponding probability matrix is as follows:
The first entry, $1/8$, represents the probability of a matrix going from all monomers to all monomers again in one step. The two terminal configurations are the rightmost ones, indicating two horizontal or two vertical dimers with equal probability of $1/16$ in one step. To find the probability distributions, we want to take this matrix and raise it to a power that is essentially infinite, exhausting the possibility of it not freezing. This can be done through known combinatorial methods.

In order to do this for greater regions, the plan was to first generate every possible tiling of the given region. These would be stored as matrices, where a 0 represented a monomer, consecutive 1’s represented a horizontal dimer, and consecutive 2’s represented a vertical dimer. We would then find probabilities of the transitions between each. These steps would both be done using extensive coding, and were likely to be computationally expensive.

This first step of creating every possible tiling was somewhat troublesome but was accomplished. Two different techniques were implemented. The first was very computationally inefficient, and included running through every possible tiling generated from every starting position on the region, then comparing it to previous tilings to see if any replications were made. The second technique, which was slightly more efficient, included creating dimers by moving from left to right on any region and creating every possibility.
Even with this slight improvement, some computational issues continued to persist. Just for a $6 \times 6$ region, which was previously confirmed to have $2,989,126,727$ different possible tilings, it took over two hours to compute and store all of the possibilities. Furthermore, even if we had found an even more efficient way to compute and store all of these configurations, the probabilities were an even greater problem. Computationally difficult “permanent” calculations were required, and the best known algorithm for calculating the permanent grows exponentially in the size of the matrix being used. Although techniques to calculate the probability distributions do exist, the computational difficulties are too much to overcome at this moment.

**Conclusion**

In closing, this research experience was extremely beneficial both to my personal and academic growth. More than anything, I learned the value of consistency and hard work from this project. I learned that you get out what you put in. When I put effort forth I got results both in the mathematics and my satisfaction. The mathematical discoveries that we made were (we believe) some that have not been found before, and we were able to ask and answer questions that haven’t been asked or answered. I learned not only this, but also that setting your sight on accomplishing a goal is crucial: as you work toward solving that problem, others will present themselves to you, creating new worlds of possibility and discovery. It is pretty evident here that a lot of our “goals” were not met. Still, discoveries were made and curiosities were satisfied. Dr. Shader’s patience and guidance throughout the process has helped foster character growth and intellectual stimulation. I am extremely grateful for having been given the opportunity to delve into what I love at such a degree of depth.
Appendix

\[ T_N = - \left( -133T_{N-1} + 9461T_{N-2} + 444475T_{N-3} + 15682766T_{N-4} - 722541212T_{N-5} - 7305389710T_{N-6} + 574267000999T_{N-7} - 233653534837T_{N-8} - 200490394546923T_{N-9} + 2620225796537763T_{N-10} + 22210634958891288T_{N-11} - 594192783412443668T_{N-12} + 1356777613008202584T_{N-13} + 68291719890210235765T_{N-14} - 268253654813476148557T_{N-15} - 4897897088354628204037T_{N-16} + 32106480529052355996947T_{N-17} + 240221865853213692513646\right)T_{N-18} - 2205552492873670287663828T_{N-19} - 8445503733084012538100930T_{N-20} + 105424445526242788867448129T_{N-21} + 21765659157641099027241277T_{N-22} - 37780513566432872417227619329T_{N-23} - 4118319568948717066962940555T_{N-24} + 105530204969194957840204391736T_{N-25} + 55437099892457787805787214776T_{N-26} - 2352103450390413090091917191982T_{N-27} - 46071906884487506019308882960T_{N-28} + 42476468145554942776805877738388T_{N-29} + 233003308722109473088317443912T_{N-30} - 6282585870038387040747248893795760T_{N-31} + 661441236451947656698640079306888T_{N-32} + 7674627201184383046502224357363136T_{N-33} - 1340808142309514586202784313960628T_{N-34} - 77797614948100967363352602387980392T_{N-35} + 17710060408737102039827341689606212T_{N-36} + 66310967463316167697585120227052536920T_{N-37} - 183812384084138028231288456367143408T_{N-38} - 474567386658325108778883289502615112T_{N-39} + 1580720710972574861834599012583394680T_{N-40} + 28717872018960545645147752488137757108T_{N-41} - 114501618648560677627096261433137174096T_{N-42} - 147499611823739883597971912391251223668T_{N-43} + 70139940418462634682213929007733252344T_{N-44} + 6447781993092355352638315988830302415456T_{N-45} - 3635210527695942510685968726268892553224T_{N-46} - 2402877680437916794084244598863714178800T_{N-47} + 1594463482763649534997173293235340917860T_{N-48} + 7637918093582185817749811172003319302776T_{N-49} - 5923238530179938960001709451510430307652T_{N-50} - 206922829517443380445184695920359646266120T_{N-51} + 18658921893033755174762785602941899136T_{N-52} + 4767820015321693482105750504645302122552T_{N-53} - 49909584790733215751702936922467009425928T_{N-54} - 9304997807320486607562975093238618906036T_{N-55} + 11350244145188480237204089606554364606618T_{N-56} + 152769202965685260477252485841385426860194T_{N-57} - 219689890860297149656872045775244968806202T_{N-58} - 208538581462627634177673224011398107039090T_{N-59} + 36218640851503680223657838216937887307604T_{N-60} + 231583216340388172415734989740624526648040T_{N-61} - 508815589260808676726332156376678124527856T_{N-62} - 199511556049127432079580538361588420627554T_{N-63} + 609146202822432337135677069553266362499938T_{N-64} + 115685446686219687847079510107356658097786T_{N-65} -
547063867615*T_{N-126} - 656045967301*T_{N-127} + 12548072151*T_{N-128} + 3420483719*T_{N-129} - 100111298*T_{N-130} - 11790708*T_{N-131} + 437922*T_{N-132} + 23797*T_{N-133} - 1027*T_{N-134} - 21*T_{N-135} + 1)

Works Cited