

Distance Multiplicities and Triangle Bases of the Cycle Space

Tao Gaede

February 23, 2024

1 Summary

This document summarizes a connection between graph distance multiplicities and cycle spaces. We present some basic lower and upper bounds that are immediately implied from this connection, and also explore the possibility of finding a formula for distance multiplicities. Using an example of a tree, we present and explore some structural connections to distance multiplicities, specifically, algebraic and combinatorial (special sets of triples). We conclude with some general observations and an initial set of questions.

2 Preliminary Notions

Throughout this discussion, we assume that G is a connected graph with no induced cycles of length at least 5. We denote G^k as the k -th power of G , which has vertex set $V(G)$ and $u, v \in V(G)$ are adjacent if and only if $d(u, v) \leq k$. Our restriction that G has no induced cycle of length at least 5 means that all of its powers are chordal, which is necessary for our discussion insofar as it relates to triangles. For each $k \in \{1, 2, \dots, \text{diam}(G)\}$, we use $m(k)$ to denote the multiplicity of distance k in G . Notice that $m(k) = |E(G^k)| - |E(G^{k-1})|$. So, to find $m(k)$ it is sufficient to bound the size of powers of G .

Recall that the cycle space of a graph G , denoted $\mathcal{C}(G)$, is the subspace of the edge space $\mathbb{F}_2^{|E(G)|} \subseteq \mathbb{F}_2^{\binom{n}{2}}$ that contains all eulerian subgraphs of G . Let c be the number of components of G . The dimension of the cycle space has a formula:

$$\dim \mathcal{C}(G) = |E(G)| - n + c.$$

3 Enumeration Discussion: Bounding and Approximating Multiplicities

The cycle space has a basis of special cycles called *fundamental cycles*, which are characterized as follows: let T be a spanning tree of G ; then for each edge $uv \in E(G) \setminus E(T)$, the unique cycle of $T + uv$ is a fundamental cycle. Since the base field is \mathbb{F}_2 , addition of two subgraphs is the symmetric difference of their edges. It follows that all fundamental cycles can be generated by the induced cycles of G . Note that if G is chordal, then all induced cycles are triangles. Since G is connected with no induced cycles of length at least 5, its powers are chordal, so their cycle spaces are spanned by triangles. However, the set of all triangles do not necessarily form a basis.

Let $A(G)$ be the adjacency matrix of G . Then based on the above, we immediately have an upper bound on $\dim \mathcal{C}(G^k)$ as the number of triangles in G^k :

$$\dim \mathcal{C}(G^k) \leq t_k,$$

where $t_k := \frac{\text{Tr}(A^3(G^k))}{6}$. For a lower bound, we use a result of Hong Yuan¹ on an upper bound for the spectral radius ρ_k of $A(G^k)$ in terms of size and order:

$$\rho \leq \sqrt{2|E| - n + 1}.$$

Rearranging, we can lower bound $\dim \mathcal{C}(G^k)$ as follows:

$$\frac{\rho_k^2 - (n - 1)}{2} \leq \dim \mathcal{C}(G^k).$$

Notice that $m(k) = \dim \mathcal{C}(G^k) - \dim \mathcal{C}(G^{k-1})$, because the $(n - 1)$ from each cancel. So, putting the above two bounds together gives us bounds on $m(k)$:

$$\frac{1}{2}(\rho_k^2 - n + 1) - t_{k-1} \leq m(k) \leq t_k - \frac{1}{2}(\rho_{k-1}^2 - n + 1).$$

We call a basis for $\mathcal{C}(G^k)$ involving only triangles a *k-triangle basis* for G . Since *k-triangle bases* are obviously a subset of all triangles, it is probably reasonable to expect that t_k can be scaled using an algebraic parameter related to $A(G^k)$ so that the result counts the number of independent triangles in a basis for $\mathcal{C}(G^k)$.

¹The paper is *A Bound on the Spectral Radius of Graphs* (1988). There are better more recent bounds if needed. The most recent one I saw was from 2019, but Hong's proof and bound are relatively simple

Computer experimentation on small order trees suggests that one can get quite close to $\dim \mathcal{C}(G^k)$, for $k \in [3, \text{diam}(G)]$, by simply dividing by $\frac{\rho_k}{2}$:

$$\dim \mathcal{C}(G^k) \sim \frac{2t_k}{\rho_k} + \Theta(n).$$

So, it appears that a formula for $m(k)$ in terms of $A(G^k)$ could be within reach. I don't know why this approximation is so close – somehow dividing by ρ removes a lot of (oriented) dependent triangles. This estimate is really good for $n \leq 13$ or so, in that it approximates $m(k)$ within an error of about 2 or 3, however the error increases with the order, and the estimate seems to become a lower bound for $k \in [3, \text{diam}(G)]$, so the estimate misses some independent triangles.

Again, this experimentation was only done on trees, so maybe tree structure is coming into play here to make the estimate closer than in general. However, I would expect this estimate to work equally well for any G with G^k chordal. Not sure though.

4 Structural Discussion: Related Matrices and Triple Systems

Example 1. Below is an example of a tree of order 7 and diameter 4 and its associated k powers and triangle bases. Note that these bases form a nested sequence. The triangle bases for these graphs are as follows:

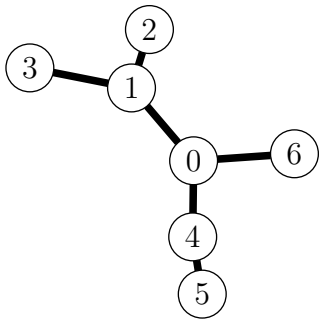
$$\mathcal{T}(T) = \emptyset$$

$$\mathcal{T}(T^2) = \{160, 460, 450, 140, 130, 230, 120\}$$

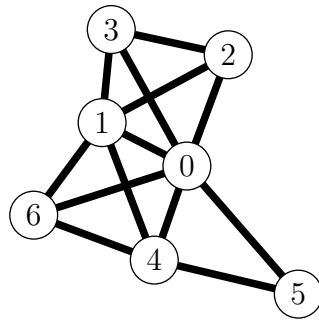
$$\mathcal{T}(T^3) = \{160, 460, 450, 140, 130, 230, 120, 260, 360, 560, 150, 240, 340\}$$

$$\mathcal{T}(T^4) = \{160, 460, 450, 140, 130, 230, 120, 260, 360, 560, 150, 240, 340, 350, 250\}$$

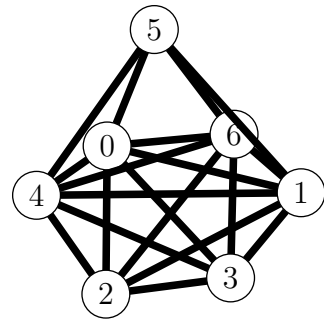
Observe that the differences in dimensions of consecutive cycle spaces are the multiplicities of the distances $k \in \{2, 3, 4\}$ of T . That is, we have $m(2) = 7$, $m(3) = 6$, $m(4) = 2$. Since $m(1) = |E(T)| = n - 1 = 6$, and there are $\binom{n}{2} = 21$ distances, all are accounted for. Below are the graph powers of T .



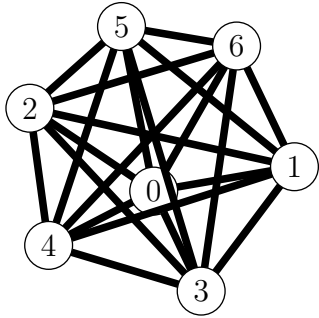
(a) A tree T .



(b) The graph T^2 .



(c) The graph T^3 .



(d) The graph T^4 .

When the triangle bases are nested (when can this be the case?), it is easy to partition the cycle space of G^{diam} into orthogonal subspaces, each with dimension equal the distance multiplicities. In general, it is always possible to do this, however, if the triangle bases aren't nested, it seems less obvious how to partition into orthogonal subspaces corresponding to the distances. We denote such a subspace for distance k by \mathcal{M}_k . In the above example, these subspaces are:

$$\mathcal{M}_2 = \text{span}(\{160, 460, 450, 140, 130, 230, 120\})$$

$$\mathcal{M}_3 = \text{span}(\{260, 360, 560, 150, 240, 340\})$$

$$\mathcal{M}_4 = \text{span}(\{350, 250\}).$$

I believe it's a coincidence that all triangles in this example contain 0. There are k -triangle bases that must contain disjoint triangles, specifically when k is not near the diameter.

5 Observations

1. The number of non-zero entries of $A(G)^k$ is equal to $|E(G^k)|$. This seems like a good way to relate $A(G^k)$ with $A(G)$. Note that this means that $m(k)$ is the difference of number of vertex pairs with walks of length k and number of vertex pairs with walks of length $k - 1$.
2. The cycle space of G is the null space of the $n \times \binom{n}{2}$ vertex-edge incidence matrix of G . Maybe we can use this to study \mathcal{M}_k or k -triangle bases somehow.
3. Since triangles are either disjoint, share 1 edge, or are equal, it seems like it ought to be a good idea to consider an triangle-edge incidence matrix for G^k here too.
4. There is a relationship between $\mathcal{C}(G^{\text{diam}})$ and the distance matrix for G . Let T be a spanning tree of G . Then colour all edges of G with 1, and then for each fundamental cycle of the form $T + uv$, where $uv \in E(G^{\text{diam}}) \setminus E(G)$, colour uv with $d(u, v)$. The result is a distance edge weighted complete graph, whose adjacency matrix is the distance matrix for G .

6 Questions

1. It could be interesting to explore questions that relate these cycle subspaces $\mathcal{T}(G^k)$ and \mathcal{M}_k with graph structure. For instance, suppose T_1 and T_2 have

the same multiplicity of distance k . When is it possible for them to have intersecting, or equal, triangle bases?

2. Inversely, suppose T_1 and T_2 have the same k -triangle basis. Then how do they relate in structure as trees? Merely having the same multiplicity for distance k doesn't necessarily imply (I don't think) that T_1 and T_2 can share a k -triangle basis.
3. How exactly does distance k show up in all vectors in \mathcal{M}_k ? I think each subgraph in \mathcal{M}_k has a pair of vertices at distance k . Also, I suppose any graph with all pairs at distance k in G , must be in \mathcal{M}_k . I think it might be the case that a graph R is in \mathcal{M}_k if and only if its max (non-infinity if disconnected) distance is equal to k in G .
4. Can we characterize the graphs corresponding to $\sum_{x \in \mathcal{M}_k} x$?
5. Suppose we find a formula for $m(k)$ in terms of $A(G^k)$. Can we find a way to express it in terms of $A(G)$?
6. Anyway, I think there could be a variety of interesting questions to ask here!

7 Cycle Space Partition

Let \mathcal{M}_k be the subspace of $\mathcal{C}(G^k)$ generated by all triangles containing a distance k edge. Equivalently, \mathcal{M}_k is the set of characteristic vectors in $\mathbb{F}_2^{\binom{n}{2}}$ of all eulerian subgraphs of G^k that are not in G^{k-1} . Observe that $\dim \mathcal{M}_k = m(k)$. Below we prove a lemma that shows the existence of a basis for \mathcal{M}_k involving triangles with exactly one distance k edge.

Lemma 1. *Let T be a tree. Then the set of triangles in T^k involving exactly one distance k edge are in bijection with the distance k edges.*

Proof. Contradict acyclicity. □

Lemma 2. *There exists a basis for \mathcal{M}_k in which all triangles have exactly one distance k edge.*

Proof. Show that the triangles must either be all dist. k edges, or have a unique dist. k edge. Show that we may swap the “all dist. k ” triangles with unique dist. k triangles outside the span of the basis. □

Proposition 7.1. *It holds that $m(2) = t_2 - \#K_{1,3}$.*

Theorem 7.2. *It holds that*

$$m(k) = \begin{cases} t_k - t_{k-1}, & \text{if } k \text{ is odd;} \\ (t_k - \#K_{1,k/2}) - t_{k-1}, & \text{if } k \text{ is even.} \end{cases}$$

Observe that we may order the edges of G^k based on their distances, from low to high. This means by Lemma [], the $m(k)$ triangles all have exactly one edge in the distance k region of the characteristic vectors.