Lecture 7

[1] In the previous lectures we saw that we could approach the problem of multicast network code design by defining the transfer functions in terms of indeterminates. If we succeed in assigning values to the indeterminates so that all the transfer matrices are non-singular, then we can satisfy the demands of all the terminals. We saw some examples of transfer matrices as well. To develop a general theory we need a means of specifying the transfer matrices in a more formal manner. Towards this end we introduced the matrix (i) $\mathbf{A}$ (of dimension $\mu \times |E|$ that specifies the transformation from the sources to the edges of the network, (ii) the matrix $\mathbf{F}$ (of dimension $|E| \times |E|$ that specifies the adjacency matrix of the line graph of the network and (iii) the matrix B (of dimension $\nu \times |E|$) that specifies the transformation from the edges of the network to the outputs. We can now express the transfer function from the sources to a given terminal as explained below.

Theorem 1. *Consider a directed acyclic graph G with designated sources and terminals. Suppose the sources are represented by a vector $\vec{x}$, and the outputs are represented by a vector $\vec{z}$. Assume also that we know the matrices $\mathbf{A}$, $\mathbf{F}$ and $\mathbf{B}$. Then:*

$$\vec{z} = \vec{x} \cdot \mathbf{M}$$

where

$$\mathbf{M} = \mathbf{A}(\mathbf{I} - \mathbf{F})^{-1}\mathbf{B}^T$$

Proof. Matrix $\mathbf{A}$ reflects the way inputs are injected into the network, i.e. the one-hop contribution from the input edges to the network edges that are adjacent to all source nodes. Similarly, Matrix $\mathbf{B}$ reflects the one-hop contribution from the network edges that are adjacent to terminals to the outputs edges on those terminals. Thus, $\mathbf{A}$ and $\mathbf{B}$ clarifies how the data enters and leaves the network, but they do not give us any hints about how it is propagated through the network, or equivalently how each edge (other than the input or output edges) contributes to the other edges in the network. In general, for any pair of edges (e_i,e_j), the contribution of e_i to the linear combination carried by e_j is captured through all the possible paths that start from e_i and ends at e_j, it turns out that all these contributions are captured in the sum

$$\mathbf{I} + \mathbf{F} + \mathbf{F}^2 + \dots$$

To clarify this, recall the example we discussed in previous lecture. The original graph is shown in Figure1, and its corresponding line graph is shown in Figure2.

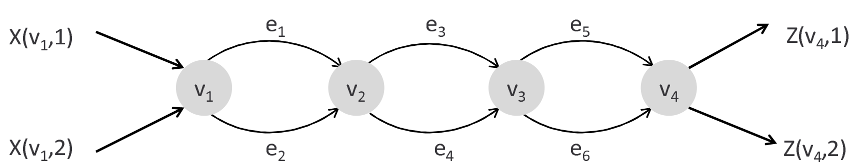

Figure 1: Original graph

[1]Based on scribed notes by Osameh Al-Kofahi from Fall 2007.

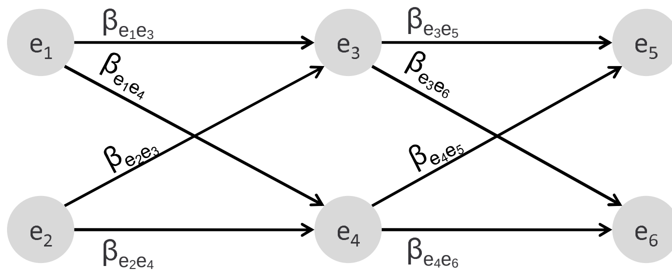

Figure 2: Line graph

The adjacency matrix for this example is:

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & \beta_{e_1e_3} & \beta_{e_1e_4} & 0 & 0 \\ 0 & 0 & \beta_{e_2e_3} & \beta_{e_2e_4} & 0 & 0 \\ 0 & 0 & 0 & 0 & \beta_{e_3e_5} & \beta_{e_3e_6} \\ 0 & 0 & 0 & 0 & \beta_{e_4e_5} & \beta_{e_4e_6} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

To understand the meaning of $\mathbf{F}^2$, let us do the math:

$$\mathbf{F^2} = \begin{bmatrix} 0 & 0 & 0 & 0 & \beta_{e_1e_3}\beta_{e_3e_5} + \beta_{e_1e_4}\beta_{e_4e_5} & \beta_{e_1e_3}\beta_{e_3e_6} + \beta_{e_1e_4}\beta_{e_4e_6} \\ 0 & 0 & 0 & 0 & \beta_{e_2e_3}\beta_{e_3e_5} + \beta_{e_2e_4}\beta_{e_4e_5} & \beta_{e_2e_3}\beta_{e_3e_6} + \beta_{e_2e_4}\beta_{e_4e_6} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Consider the entry $(1, 5)$ in $\mathbf{F}^2$, which is equal to $\beta_{e_1e_3}\beta_{e_3e_5} + \beta_{e_1e_4}\beta_{e_4e_5}$, it shows that e_1 can contribute to e_5 via the two paths $(e_1 \rightarrow e_3 \rightarrow e_5)$ and $(e_1 \rightarrow e_4 \rightarrow e_5)$. If N denotes the number of hops in the longest path in G the contributions between all edge pairs is captured by

$$\sum_{k=0}^{N} \mathbf{F^k}$$

. To complete our proof we first need to prove the following two claims.

Claim 2. *Suppose* $\mathbf{F}$ *is the adjacency matrix for a topologically ordered line graph, i.e.,* $\mathbf{F}$ *is strictly upper triangular. Then:*

$$\mathbf{F}^N = 0$$

Proof. Consider the i^{th} row in $\mathbf{F}$ (starting from 0), since $\mathbf{F}$ is strictly upper triangular the position of the first non-zero element in this row is $i+1$. Therefore, multiplying row i with column j will give a zero for all $j \leq i+1$ and will depend on the matrix elements for $j > i+1$. Which means that all the elements $(i, i+1)$ (i.e. the elements on the first upper diagonal above the main diagonal) in the resulting matrix $\mathbf{F}^2$ are equal to zero.

Multiplying $\mathbf{F}^2$ by $\mathbf{F}$ will make all the elements (i,j) for $j \leq i+2$ equal to zero, since the position of the first non-zero element in the i^{th} row in $\mathbf{F}^2$ is $i+2$. Thus making the elements on the second upper diagonal above the main diagonal equal to zero. Therefore, after k multiplications the elements of the k-upper diagonal above the main will be all zeros. That is, after N multiplications all the elements in the matrix $\mathbf{F}^N$ will be zeros, which concludes the proof. ■

Matrices that have the property that $\mathbf{F}^k = 0$ for some k are called *nilpotent*.

Claim 3. *Assume that matrix* $\mathbf{F}$ *is an adjacency matrix for a topologically ordered line graph, i.e.,* $\mathbf{F}$ *is strictly upper triangular. Then:*

$$\mathbf{I} + \mathbf{F} + \mathbf{F}^2 + \cdots + \mathbf{F}^N = (\mathbf{I} - \mathbf{F})^{-1}$$

Proof. Multiply both sides by $(\mathbf{I} - \mathbf{F})$, we now need to show that the LHS is equal to $\mathbf{I}$:

$$(\mathbf{I} - \mathbf{F})(\mathbf{I} + \mathbf{F} + \mathbf{F}^2 + \cdots + \mathbf{F}^N)$$

$$(\mathbf{I} + \mathbf{F} + \mathbf{F}^2 + \cdots + \mathbf{F}^N) - (\mathbf{F} + \mathbf{F}^2 + \cdots + \mathbf{F}^{N+1})$$

which evaluates to

$$\mathbf{I} - \mathbf{F}^{N+1} = \mathbf{I}$$

$\mathbf{F}^{N+1} = 0$ because $\mathbf{F}^N$ is equal to zero. This concludes the proof. ■

Claim 3 shows that $(\mathbf{I} - \mathbf{F})^{-1}$ is the network transfer matrix that captures the contributions between all the edges using all the possible number of hops. Now recall that $\mathbf{A}$ is the transfer matrix from the input links to the network, and that $\mathbf{B}$ is the transfer matrix from the network to the output edges. It is now clear that

$$\mathbf{M} = \mathbf{A}(\mathbf{I} - \mathbf{F})^{-1}\mathbf{B}^T$$

This concludes the proof of Theorem 1. ■

Theorem 4. *Let* $\mathbf{M} = \mathbf{A}(\mathbf{I} - \mathbf{F})^{-1}\mathbf{B}^T$, *and suppose that* $\mathbf{A}$ *and* $\mathbf{B}$ *have the same dimensions. Then:*

$$\det(\mathbf{M}) = \pm \det\left(\begin{bmatrix} \mathbf{A} & 0 \\ (\mathbf{I} - \mathbf{F}) & \mathbf{B}^T \end{bmatrix}\right)$$

Proof. Note that

$$\det(\mathbf{M}) = \pm \det\left(\begin{bmatrix} (\mathbf{I} - \mathbf{F}) & \mathbf{B}^T \\ \mathbf{A} & 0 \end{bmatrix}\right)$$

, since the matrix on the RHS above is only a permutation of rows of the matrix of interest. Next, we introduce the Schur factorization of block matrices.

Consider a block matrix $\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$. If $\mathbf{A}$ is invertible, the following factorization holds.

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 \\ \mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{bmatrix} \times \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{S}_\mathbf{A} \end{bmatrix} \times \begin{bmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ 0 & \mathbf{I} \end{bmatrix},$$

where $\mathbf{S_A} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$.

Applying this factorization to the matrix $\begin{bmatrix} (\mathbf{I}-\mathbf{F}) & \mathbf{B}^T \\ \mathbf{A} & 0 \end{bmatrix}$, we obtain

$$\begin{bmatrix} (\mathbf{I}-\mathbf{F}) & \mathbf{B}^T \\ \mathbf{A} & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & 0 \\ \mathbf{A}(\mathbf{I}-\mathbf{F})^{-1} & \mathbf{I} \end{bmatrix} \times \begin{bmatrix} (\mathbf{I}-\mathbf{F}) & 0 \\ 0 & -\mathbf{A}(\mathbf{I}-\mathbf{F})^{-1}\mathbf{B}^T \end{bmatrix} \times \begin{bmatrix} \mathbf{I} & (\mathbf{I}-\mathbf{F})^{-}\mathbf{B}^T \\ 0 & \mathbf{I} \end{bmatrix},$$

Now, the determinant of the RHS is

$$\det(\mathbf{I}-\mathbf{F}) \times \det \pm \mathbf{A}(\mathbf{I}-\mathbf{F})^{-1}\mathbf{B}^T,$$

using the fact that the determinant of a product of matrices is the product of the individual determinants and the determinant of lower triangular (or upper triangular) matrices with ones on the diagonal is 1. The result follows by noting that $(\mathbf{I}-\mathbf{F})$ is also upper triangular with ones on the diagonal and has determinant 1. ■

We now state and prove the famous multicast theorem of network coding.

Theorem 5. *Consider a directed acyclic graph G with unit capacities that has a single source node s (with h sources) and a set of terminal nodes T. The multicast property with rate h is said to be satisfied if max-flow$(s, T_i) \geq h$, for all $T_i \in T$. If G satisfies the multicast property a network code that supports the multicast rate h is guaranteed to exist as long as the field size is larger than $|T|$.*

Proof. The transfer matrix from s to a given T_i can be written as:

$$\mathbf{M}_i = \mathbf{A}(\mathbf{I}-\mathbf{F})^{-1}\mathbf{B}_i^T$$

For T_i to be able to recover the h sources, we need $\det(\mathbf{M}_i) \neq 0$. Therefore the following must hold for all the terminals to be able to recover the h sources:

$$\prod_{i=1}^{|T|} \det(\mathbf{M}_i) \neq 0$$

The product of the determinants is a polynomial in the indeterminates $\{\alpha_{l,e}\}$, $\{\beta_{e',e}\}$ and $\{\epsilon_{e,l}\}$. To ensure that it is non-zero we first need to show that it is not identically zero (i.e. it is not the zero polynomial). This is true since the max-flow condition holds for all the terminals, which implies that there is at least one assignment of the indeterminates such each $\det(M_i)$ is non-zero.

Now, note that from Theorem 4:

$$\det(\mathbf{M}_i) = \det\left(\begin{bmatrix} \mathbf{A} & 0 \\ (\mathbf{I}-\mathbf{F}) & \mathbf{B}_i^T \end{bmatrix}\right)$$

Since each variable in $\begin{bmatrix} \mathbf{A} & 0 \\ (\mathbf{I}-\mathbf{F}) & \mathbf{B}_i^T \end{bmatrix}$ appears at most once, then in the determinant also each variable appears at most once, i.e., with a degree equal to 1. Therefore in the product $\prod_{i=1}^{|T|} \det(\mathbf{M}_i)$ the max degree of any variable is at most $|T|$. Now we use the *sparse zeros lemma*, which shows that an assignment of the variables such that the polynomial evaluates to a non-zero value exists as long as the size of the field is greater than $|T|$. ■