Welcome from the Department of Mathematics and Physics at the University of New Haven

Matrix Braids

Joseph Kolibal

The University of New Haven

October, 19 2017

Abstract

Braiding matrices arise as a subtopic of the the Yang-Baxter equation, which has been studied extensively due to application in numerous fields of mathematics and physics. We connect these to a simplified matrix representation and focus on obtaining solutions to matrix braids by considering special matrices where solutions are more easily found. Finally, we suggest a fixed point iteration algorithm to determine the braid complement of a given matrix, if it exists.

This lecture is a lecture within a lecture. It is constructed to also show how to write a themed lecture using Beamer, utilizing hyperlinks to provide background information to the viewer.

- One of the earliest introductions of the Yang-Baxter equation in statistical mechanics was in 1944, when Lars Onsanger[1] briefly mentioned the star-triangle relation in the introduction of his solution to the Ising model (named after Ernst Ising).
- It wasn't until November and December of 1967, when C.N. Yang published two papers on a simple one-dimensional quantum many-body problem, and for the first time introduce the equation.

$$A(u)B(u+v)A(v) = B(u)A(u+v)B(v), \quad (1)$$

  where $A(u)$ and $B(v)$ are rational functions of u and v.
- The same equation was also used by R.J. Baxter in 1972, when he was studying some classical statistical mechanics problems in two-dimensions, and discovered his solution of the eight-vertex model.
- The term Yang-Baxter equation was introduced in late 1970s by Fadeev.

[1] He and other notables are buried in Yale's Grove Street Cemetery located along Grove St. in New Haven.

- One of the earliest introductions of the Yang-Baxter equation in statistical mechanics was in 1944, when Lars Onsanger[1] briefly mentioned the star-triangle relation in the introduction of his solution to the Ising model (named after Ernst Ising).
- It wasn't until November and December of 1967, when C.N. Yang published two papers on a simple one-dimensional quantum many-body problem, and for the first time introduce the equation.

$$A(u)B(u+v)A(v) = B(u)A(u+v)B(v), \tag{1}$$

where $A(u)$ and $B(v)$ are rational functions of u and v.

- The same equation was also used by R.J. Baxter in 1972, when he was studying some classical statistical mechanics problems in two-dimensions, and discovered his solution of the eight-vertex model.
- The term Yang-Baxter equation was introduced in late 1970s by Fadeev.

[1] He and other notables are buried in Yale's Grove Street Cemetery located along Grove St. in New Haven.

- One of the earliest introductions of the Yang-Baxter equation in statistical mechanics was in 1944, when Lars Onsanger[1] briefly mentioned the star-triangle relation in the introduction of his solution to the Ising model (named after Ernst Ising).
- It wasn't until November and December of 1967, when C.N. Yang published two papers on a simple one-dimensional quantum many-body problem, and for the first time introduce the equation.

$$A(u)B(u+v)A(v) = B(u)A(u+v)B(v), \tag{1}$$

where $A(u)$ and $B(v)$ are rational functions of u and v.
- The same equation was also used by R.J. Baxter in 1972, when he was studying some classical statistical mechanics problems in two-dimensions, and discovered his solution of the eight-vertex model.
- The term Yang-Baxter equation was introduced in late 1970s by Fadeev.

[1] He and other notables are buried in Yale's Grove Street Cemetery located along Grove St. in New Haven.

- One of the earliest introductions of the Yang-Baxter equation in statistical mechanics was in 1944, when Lars Onsanger[1] briefly mentioned the star-triangle relation in the introduction of his solution to the Ising model (named after Ernst Ising).
- It wasn't until November and December of 1967, when C.N. Yang published two papers on a simple one-dimensional quantum many-body problem, and for the first time introduce the equation.

$$A(u)B(u+v)A(v) = B(u)A(u+v)B(v), \quad (1)$$

where $A(u)$ and $B(v)$ are rational functions of u and v.

- The same equation was also used by R.J. Baxter in 1972, when he was studying some classical statistical mechanics problems in two-dimensions, and discovered his solution of the eight-vertex model.
- The term Yang-Baxter equation was introduced in late 1970s by Fadeev.

[1] He and other notables are buried in Yale's Grove Street Cemetery located along Grove St. in New Haven.

- In the past two decades, the Yang-Baxter equation has been studied extensively, and its implementations can be found in quantum mechanics, classical statistical mechanics, knot theory, braid theory, quantum groups[2], and other fields.
- There are two forms of the YBE. The parameter-dependent form is given by

$$R_{12}(u)R_{13}(u+v)R_{23}(v) = R_{23}(v)R_{13}(u+v)R_{12}(u), \tag{2}$$

where $R(u)$ is a parameter-dependent invertible element of the tensor product $A \otimes A$, for some unital associative algebra, A.
- If we drop u and v, then we have the parameter-independent YBE

$$R_{12}R_{13}R_{23} = R_{23}R_{13}R_{23}. \tag{3}$$

[2] A group G is a finite or infinite set of elements together with a binary operation that satisfy 1) Closure: if $a, b \in G$, then the product $ab \in G$; 2) Associativity: $\forall a, b, c \in G, (ab)c = a(bc)$; 3) Identity: $\exists e$ such that $ea = ae \; \forall A \in G$; and, 4) Inverse: if $a \in G$, then $\exists b \equiv a^{-1} \in G$ such that $aa^{-1} = a^{-1}a = e$.

- We will briefly introduce the Artin's braid group and the representation of the Yang-Baxter equation. An n-strand *Artin braid group*, denoted by B_n, is generated by $\{\sigma_i | 1 \leq i \leq n-1\}$, with the following relationships

$$\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1},$$

for all $1 \leq i \leq n-1$, and

$$\sigma_i \sigma_j = \sigma_j \sigma_i,$$

when $|i-j| \geq 1$. If we let each n-strand braid represent a linear mapping from $V^{\otimes n}$ to $V^{\otimes n}$, with generator σ_i associated with the map $R : V \otimes V \longrightarrow V \otimes V$ in the following manner:

$$\sigma_i \longmapsto I \otimes I \ldots \otimes R \otimes \ldots I \otimes I, \tag{4}$$

with R in the i th position and I representing the identity map, then the three strand case yields the Yang-Baxter equation.

- Specifically, in this case we have

$$(R\otimes I)(I\otimes R)(R\otimes I)=(I\otimes R)(R\otimes I)(I\otimes R). \tag{5}$$

- Since R is a linear map, it may be expressed as a matrix. An example of a braid matrix is

$$R=\begin{pmatrix} 1-t & t & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It can be easily verified that matrix R indeed satisfies the Yang-Baxter equation, i.e., that these solve the YBE.

- We analyze the Yang-Baxter equation specialized to matrices $A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, $B : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, having the following form

$$ABA = BAB. \tag{6}$$

- We seek to characterize solutions of (6), including finding the necessary and if possible sufficient conditions under which distinct matrices A and B satisfy (6).
- In that regard, the approach is not too dissimilar to analyzing the structure of $AB = BA$, i.e., determining when two distinct matrices commute. [3]

In that sense it seems appropriate to coin the usage that two distinct matrices form a braid, or more simply braid if they satisfy (6).

[3] The physics preamble suggests the form $ABA = BAB$, but why not consider when $AAB = BAB$, or $BAA = BAB$, or . . . , i.e., consider permutations of two matrices taken three at a time. As an academic pursuit, we can consider the permutation products of n matrices taken p at a time, but we can also consider matrix multiplication schemes for public key encryption, e.g., the Simple Matrix Scheme. The simpler Hill cipher was the first attempt to do cryptography with matrices.

- We analyze the Yang-Baxter equation specialized to matrices $A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, $B : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, having the following form

$$ABA = BAB. \tag{6}$$

- We seek to characterize solutions of (6), including finding the necessary and if possible sufficient conditions under which distinct matrices A and B satisfy (6).
- In that regard, the approach is not too dissimilar to analyzing the structure of $AB = BA$, i.e., determining when two distinct matrices commute. [3]

In that sense it seems appropriate to coin the usage that two distinct matrices form a braid, or more simply braid if they satisfy (6).

[3] The physics preamble suggests the form $ABA = BAB$, but why not consider when $AAB = BAB$, or $BAA = BAB$, or . . . , i.e., consider permutations of two matrices taken three at a time. As an academic pursuit, we can consider the permutation products of n matrices taken p at a time, but we can also consider matrix multiplication schemes for public key encryption, e.g., the Simple Matrix Scheme. The simpler Hill cipher was the first attempt to do cryptography with matrices.

- We analyze the Yang-Baxter equation specialized to matrices $A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, $B : \mathbb{C}^n \longrightarrow \mathbb{C}^n$, having the following form

$$ABA = BAB. \tag{6}$$

- We seek to characterize solutions of (6), including finding the necessary and if possible sufficient conditions under which distinct matrices A and B satisfy (6).
- In that regard, the approach is not too dissimilar to analyzing the structure of $AB = BA$, i.e., determining when two distinct matrices commute. [3]

In that sense it seems appropriate to coin the usage that two distinct matrices form a braid, or more simply braid if they satisfy (6).

[3] The physics preamble suggests the form $ABA = BAB$, but why not consider when $AAB = BAB$, or $BAA = BAB$, or . . . , i.e., consider permutations of two matrices taken three at a time. As an academic pursuit, we can consider the permutation products of n matrices taken p at a time, but we can also consider matrix multiplication schemes for public key encryption, e.g., the Simple Matrix Scheme. The simpler Hill cipher was the first attempt to do cryptography with matrices.

- An obvious necessary condition to have braiding matrices, i.e., to satisfy (6), is that

$$\det(A)\det(B)\det(A) = \det(B)\det(A)\det(B)$$

$$\Longrightarrow \det(A) = \det(B) \text{ if } A, B \text{ nonsingular.}$$

- Since the determinant of a matrix is equal to the product of the eigenvalues of the matrix, we have

$$\prod_i \lambda_i(A) = \prod_j \lambda_j(B),$$

where λ_i and λ_j are all the eigenvalues of A and B, including multiplicities.

Not much information if A or B are singular.

How was this slide written?

- An obvious necessary condition to have braiding matrices, i.e., to satisfy (6), is that

$$\det(A)\det(B)\det(A) = \det(B)\det(A)\det(B)$$
$$\implies \det(A) = \det(B) \text{ if } A, B \text{ nonsingular.}$$

- Since the determinant of a matrix is equal to the product of the eigenvalues of the matrix, we have

$$\prod_i \lambda_i(A) = \prod_j \lambda_j(B),$$

where λ_i and λ_j are all the eigenvalues of A and B, including multiplicities.

Not much information if A or B are singular.

How was this slide written?

- An obvious necessary condition to have braiding matrices, i.e., to satisfy (6), is that

$$\det(A)\det(B)\det(A) = \det(B)\det(A)\det(B)$$
$$\implies \det(A) = \det(B) \text{ if } A, B \text{ nonsingular.}$$

- Since the determinant of a matrix is equal to the product of the eigenvalues of the matrix, we have

$$\prod_i \lambda_i(A) = \prod_j \lambda_j(B),$$

where λ_i and λ_j are all the eigenvalues of A and B, including multiplicities.

Not much information if A or B are singular.

How was this slide written?

- An obvious necessary condition to have braiding matrices, i.e., to satisfy (6), is that

$$\det(A)\det(B)\det(A) = \det(B)\det(A)\det(B)$$
$$\implies \det(A) = \det(B) \text{ if } A, B \text{ nonsingular.}$$

- Since the determinant of a matrix is equal to the product of the eigenvalues of the matrix, we have

$$\prod_i \lambda_i(A) = \prod_j \lambda_j(B),$$

where λ_i and λ_j are all the eigenvalues of A and B, including multiplicities.

Not much information if A or B are singular.

How was this slide written?

Theorem

If A and B are two simultaneously diagonalizable matrices that satisfy that braid, then $A = B$.

Proof: [4] Suppose that matrices A and B are simultaneously diagonalizable, .e., there exists a matrix P such that $\Lambda_A = P^{-1}AP$ and $\Lambda_B = P^{-1}BP$, where Λ_A and Λ_B are diagonal matrices.

Since $ABA = BAB$ we have

$$(P\Lambda_A P^{-1})(P\Lambda_B P^{-1})(P\Lambda_A P^{-1}) = (P\Lambda_B P^{-1})(P\Lambda_A P^{-1})(P\Lambda_B P^{-1}). \quad (7)$$

Thus, $\Lambda_A\Lambda_B\Lambda_A = \Lambda_B\Lambda_A\Lambda_B$, hence

$$a_i^2 b_i = a_i b_i^2 \implies a_i = b_i \implies \Lambda_A = \Lambda_B.$$

This means that $A = B$. ∎

Trivial. Indeed it is not difficult to come up with trivial solutions.

[4]– Some connections to the essence of mathematical thought, and our own MATH 1121 Foundations of Math.

Theorem

If A and B are two simultaneously diagonalizable matrices that satisfy that braid, then $A = B$.

Proof:[4] Suppose that matrices A and B are simultaneously diagonalizable, .e., there exists a matrix P such that $\Lambda_A = P^{-1}AP$ and $\Lambda_B = P^{-1}BP$, where Λ_A and Λ_B are diagonal matrices.

Since $ABA = BAB$ we have

$$(P\Lambda_A P^{-1})(P\Lambda_B P^{-1})(P\Lambda_A P^{-1}) = (P\Lambda_B P^{-1})(P\Lambda_A P^{-1})(P\Lambda_B P^{-1}). \quad (7)$$

Thus, $\Lambda_A\Lambda_B\Lambda_A = \Lambda_B\Lambda_A\Lambda_B$, hence

$$a_i^2 b_i = a_i b_i^2 \implies a_i = b_i \implies \Lambda_A = \Lambda_B.$$

This means that $A = B$. □

Trivial. Indeed it is not difficult to come up with trivial solutions.

[4]– Some connections to the essence of mathematical thought, and our own MATH 1121 Foundations of Math.

Theorem

If A and B are two simultaneously diagonalizable matrices that satisfy that braid, then $A = B$.

Proof:[4] Suppose that matrices A and B are simultaneously diagonalizable, .e., there exists a matrix P such that $\Lambda_A = P^{-1}AP$ and $\Lambda_B = P^{-1}BP$, where Λ_A and Λ_B are diagonal matrices.

Since $ABA = BAB$ we have

$$(P\Lambda_A P^{-1})(P\Lambda_B P^{-1})(P\Lambda_A P^{-1}) = (P\Lambda_B P^{-1})(P\Lambda_A P^{-1})(P\Lambda_B P^{-1}). \quad (7)$$

Thus, $\Lambda_A\Lambda_B\Lambda_A = \Lambda_B\Lambda_A\Lambda_B$, hence

$$a_i^2 b_i = a_i b_i^2 \implies a_i = b_i \implies \Lambda_A = \Lambda_B.$$

This means that $A = B$. □

Trivial. Indeed it is not difficult to come up with trivial solutions.

[4]– Some connections to the essence of mathematical thought, and our own MATH 1121 Foundations of Math.

Theorem

If A and B are two simultaneously diagonalizable matrices that satisfy that braid, then $A = B$.

Proof: [4] Suppose that matrices A and B are simultaneously diagonalizable, .e., there exists a matrix P such that $\Lambda_A = P^{-1}AP$ and $\Lambda_B = P^{-1}BP$, where Λ_A and Λ_B are diagonal matrices.

Since $ABA = BAB$ we have

$$(P\Lambda_A P^{-1})(P\Lambda_B P^{-1})(P\Lambda_A P^{-1}) = (P\Lambda_B P^{-1})(P\Lambda_A P^{-1})(P\Lambda_B P^{-1}). \quad (7)$$

Thus, $\Lambda_A\Lambda_B\Lambda_A = \Lambda_B\Lambda_A\Lambda_B$, hence

$$a_i^2 b_i = a_i b_i^2 \implies a_i = b_i \implies \Lambda_A = \Lambda_B.$$

This means that $A = B$. □

Trivial. Indeed it is not difficult to come up with trivial solutions.

[4]– Some connections to the essence of mathematical thought, and our own MATH 1121 Foundations of Math.

- Consider solving $ABA = BAB$ by letting

$$W = BA, \tag{8}$$

yielding the revised form of the problem

$$AW = WB. \tag{9}$$

- Observe that by regrouping (8), an alternative factoring[5] can be achieved by letting $V = AB$, yielding

$$VA = BV. \tag{10}$$

- Solutions to Sylvester's equation, i.e., $AW + WB = C$ when $C = 0$ also provide solutions to braiding matrices.

Theorem

(Sylvester – 1884) The equation $AW = WB$ has a non-zero solution if and only if A and B have a characteristic root in common.

[5] The area of matrix factorization and decomposition has been extensively studied.

- Consider solving $ABA = BAB$ by letting

$$W = BA, \tag{8}$$

yielding the revised form of the problem

$$AW = WB. \tag{9}$$

- Observe that by regrouping (8), an alternative factoring[5] can be achieved by letting $V = AB$, yielding

$$VA = BV. \tag{10}$$

- Solutions to Sylvester's equation, i.e., $AW + WB = C$ when $C = 0$ also provide solutions to braiding matrices.

Theorem

(Sylvester – 1884) The equation $AW = WB$ has a non-zero solution if and only if A and B have a characteristic root in common.

[5] The area of matrix factorization and decomposition has been extensively studied.

- Consider solving $ABA = BAB$ by letting

$$W = BA, \tag{8}$$

yielding the revised form of the problem

$$AW = WB. \tag{9}$$

- Observe that by regrouping (8), an alternative factoring[5] can be achieved by letting $V = AB$, yielding

$$VA = BV. \tag{10}$$

- Solutions to Sylvester's equation, i.e., $AW + WB = C$ when $C = 0$ also provide solutions to braiding matrices.

Theorem

(Sylvester – 1884) The equation $AW = WB$ has a non-zero solution if and only if A and B have a characteristic root in common.

[5] The area of matrix factorization and decomposition has been extensively studied.

Theorem

(Cecioni-Frobenius) The number of linearly independent solutions of the equation $AW = WB$ is $\sum_{i=1}^{n}\sum_{j=1}^{n} e_{ij}$, where e_{ij} is the degree of the greatest common divisor of the invariant factor a_i of $\lambda I - A$, and the invariant factor b_j of $\lambda I - B$, i.e., $e_{ij} = deg(gcd(a_i, b_j))$. [6]

- Clearly, this applies to solutions to braiding matrices, however the deeper problem of demonstrating the existence of a factorisable $W = BA$ remains.
- Note that if W is invertible, then A is W similar to B and B is W similar to A, i.e.,

$$A = WBW^{-1} \text{ and } B = W^{-1}AW. \tag{11}$$

[6] Where computing SNF $(\lambda I - A) =$RCF(A).

Theorem

(Cecioni-Frobenius) The number of linearly independent solutions of the equation $AW = WB$ *is* $\sum_{i=1}^{n}\sum_{j=1}^{n} e_{ij}$*, where* e_{ij} *is the degree of the greatest common divisor of the invariant factor* a_i *of* $\lambda I - A$*, and the invariant factor* b_j *of* $\lambda I - B$*, i.e.,* $e_{ij} = deg(gcd(a_i, b_j))$*.* [6]

- Clearly, this applies to solutions to braiding matrices, however the deeper problem of demonstrating the existence of a factorisable $W = BA$ remains.
- Note that if W is invertible, then A is W similar to B and B is W similar to A, i.e.,

$$A = WBW^{-1} \text{ and } B = W^{-1}AW. \tag{11}$$

[6] Where computing SNF $(\lambda I - A) =$RCF(A).

Theorem

(Cecioni-Frobenius) The number of linearly independent solutions of the equation $AW = WB$ is $\sum_{i=1}^{n}\sum_{j=1}^{n} e_{ij}$, where e_{ij} is the degree of the greatest common divisor of the invariant factor a_i of $\lambda I - A$, and the invariant factor b_j of $\lambda I - B$, i.e., $e_{ij} = deg(gcd(a_i, b_j))$. [6]

- Clearly, this applies to solutions to braiding matrices, however the deeper problem of demonstrating the existence of a factorisable $W = BA$ remains.
- Note that if W is invertible, then A is W similar to B and B is W similar to A, i.e.,

$$A = WBW^{-1} \text{ and } B = W^{-1}AW. \tag{11}$$

[6] Where computing SNF $(\lambda I - A) =$RCF(A).

- Solving problems of the form $AW = WB$ will be denoted as the Cecioni-Frobenius equation (CFE).
- This form still poses difficulties in solving for braiding matrices due to the requirements that a factorization of W must yield the matrices A and B when a solution W of the CFE is found, however for a class of problems this factorization can be trivially achieved.
- Suppose $A = (a_{ij})$ and $B = (b_{ij})$ are doubly-stochastic matrices. To construct doubly stochastic solutions we need to find factorization such that $W = BA$.
- One such matrix is $W = U/n$, where U is the $n \times n$ unit matrix, i.e., the matrix in which all coefficients are ones.

We know we have one solution since the eigenvalue 1 is common to all stochastic matrices.

- Solving problems of the form $AW = WB$ will be denoted as the Cecioni-Frobenius equation (CFE).
- This form still poses difficulties in solving for braiding matrices due to the requirements that a factorization of W must yield the matrices A and B when a solution W of the CFE is found, however for a class of problems this factorization can be trivially achieved.
- Suppose $A = (a_{ij})$ and $B = (b_{ij})$ are doubly-stochastic matrices. To construct doubly stochastic solutions we need to find factorization such that $W = BA$.
- One such matrix is $W = U/n$, where U is the $n \times n$ unit matrix, i.e., the matrix in which all coefficients are ones.

We know we have one solution since the eigenvalue 1 is common to all stochastic matrices.

Theorem

Let A and B be $n \times n$ doubly stochastic matrices. If $B = W = U/n$, where U is the unit matrix, then A and B are solutions of the CFE.

Proof:

- Since A and B are doubly-stochastic, the column vector $(1 \dots 1)^{\mathrm{T}}$ is a right eigenvector for A or B with eigenvalue 1, and the row vector $(1 \dots 1)$ is a left eigenvector for A or B with eigenvalue 1.
- Thus, the matrix U, in which every entry is one, is a left eigenvector matrix for A or B with eigenvalue 1, and is also a right eigenvector matrix for A or B with eigenvalue 1.
- This means that $W = U/n$, is also a right and left eigenvector matrix for A or B, i.e., $AW = BW = WA = WB = W$.
- Let $B = W$. Since $WA = W$, then for this choice of B, $BA = W$, and so $AW = WB$. □

Theorem

Let A and B be $n \times n$ doubly stochastic matrices. If $B = W = U/n$, where U is the unit matrix, then A and B are solutions of the CFE.

Proof:

- Since A and B are doubly-stochastic, the column vector $(1 \ldots 1)^{\mathrm{T}}$ is a right eigenvector for A or B with eigenvalue 1, and the row vector $(1 \ldots 1)$ is a left eigenvector for A or B with eigenvalue 1.
- Thus, the matrix U, in which every entry is one, is a left eigenvector matrix for A or B with eigenvalue 1, and is also a right eigenvector matrix for A or B with eigenvalue 1.
- This means that $W = U/n$, is also a right and left eigenvector matrix for A or B, i.e., $AW = BW = WA = WB = W$.
- Let $B = W$. Since $WA = W$, then for this choice of B, $BA = W$, and so $AW = WB$. □

Theorem

Let A and B be $n \times n$ doubly stochastic matrices. If $B = W = U/n$, where U is the unit matrix, then A and B are solutions of the CFE.

Proof:

- Since A and B are doubly-stochastic, the column vector $(1\dots1)^{\mathrm{T}}$ is a right eigenvector for A or B with eigenvalue 1, and the row vector $(1\dots1)$ is a left eigenvector for A or B with eigenvalue 1.
- Thus, the matrix U, in which every entry is one, is a left eigenvector matrix for A or B with eigenvalue 1, and is also a right eigenvector matrix for A or B with eigenvalue 1.
- This means that $W = U/n$, is also a right and left eigenvector matrix for A or B, i.e., $AW = BW = WA = WB = W$.
- Let $B = W$. Since $WA = W$, then for this choice of B, $BA = W$, and so $AW = WB$. □

Theorem

Let A and B be $n \times n$ doubly stochastic matrices. If $B = W = U/n$, where U is the unit matrix, then A and B are solutions of the CFE.

Proof:

- Since A and B are doubly-stochastic, the column vector $(1 \dots 1)^{\mathrm{T}}$ is a right eigenvector for A or B with eigenvalue 1, and the row vector $(1 \dots 1)$ is a left eigenvector for A or B with eigenvalue 1.
- Thus, the matrix U, in which every entry is one, is a left eigenvector matrix for A or B with eigenvalue 1, and is also a right eigenvector matrix for A or B with eigenvalue 1.
- This means that $W = U/n$, is also a right and left eigenvector matrix for A or B, i.e., $AW = BW = WA = WB = W$.
- Let $B = W$. Since $WA = W$, then for this choice of B, $BA = W$, and so $AW = WB$. □

Corollary

If A and B are doubly stochastic matrices and at least one of them is equal to W, then A and B are braiding matrices.

Proof:

- Assume $B = W$, then by the previous Theorem, A and B are solutions of (9).
- Since $BA = W$ and $AW = WB$, then $ABA = BAB$. □

- Some elementary cases that can be dispensed with quickly in solving $ABA = BAB$ include when the matrices A and B are invertible and commuting. In this case we have $AB = BA$, and thus

$$ABA = BAB \implies ABA = ABB \implies A = B. \quad (12)$$

Corollary

If A and B are doubly stochastic matrices and at least one of them is equal to W, then A and B are braiding matrices.

Proof:

- Assume $B = W$, then by the previous Theorem, A and B are solutions of (9).
- Since $BA = W$ and $AW = WB$, then $ABA = BAB$. □

- Some elementary cases that can be dispensed with quickly in solving $ABA = BAB$ include when the matrices A and B are invertible and commuting. In this case we have $AB = BA$, and thus

$$ABA = BAB \implies ABA = ABB \implies A = B. \tag{12}$$

Corollary

If A and B are doubly stochastic matrices and at least one of them is equal to W, then A and B are braiding matrices.

Proof:

- Assume $B = W$, then by the previous Theorem, A and B are solutions of (9).
- Since $BA = W$ and $AW = WB$, then $ABA = BAB$. □

- Some elementary cases that can be dispensed with quickly in solving $ABA = BAB$ include when the matrices A and B are invertible and commuting. In this case we have $AB = BA$, and thus

$$ABA = BAB \implies ABA = ABB \implies A = B. \tag{12}$$

Theorem

Given A and B two invertible circulant matrices, then A and B are braiding matrices only when $A = B$.

Proof:

- Assume that A and B are circulant matrices which braid.
- We recall that any circulant matrix is diagonalized by the Fourier matrix, F. Then, $A = F^{-1}D_AF$ and $B = F^{-1}D_BF$. Thus, $AB = F^{-1}D_AFF^{-1}D_BF = F^{-1}D_AD_BF$.
- Since two diagonal matrices always commute, $AB = F^{-1}D_BD_AF = F^{-1}D_BFF^{-1}D_AF = BA$.
- But since A and B are invertible, then $A = B$. □

Theorem

Given A and B two invertible circulant matrices, then A and B are braiding matrices only when $A = B$.

Proof:

- Assume that A and B are circulant matrices which braid.
- We recall that any circulant matrix is diagonalized by the Fourier matrix, F. Then, $A = F^{-1}D_AF$ and $B = F^{-1}D_BF$. Thus, $AB = F^{-1}D_AFF^{-1}D_BF = F^{-1}D_AD_BF$.
- Since two diagonal matrices always commute, $AB = F^{-1}D_BD_AF = F^{-1}D_BFF^{-1}D_AF = BA$.
- But since A and B are invertible, then $A = B$. □

Theorem

Given A and B two invertible circulant matrices, then A and B are braiding matrices only when $A = B$.

Proof:

- Assume that A and B are circulant matrices which braid.
- We recall that any circulant matrix is diagonalized by the Fourier matrix, F. Then, $A = F^{-1}D_AF$ and $B = F^{-1}D_BF$. Thus, $AB = F^{-1}D_AFF^{-1}D_BF = F^{-1}D_AD_BF$.
- Since two diagonal matrices always commute, $AB = F^{-1}D_BD_AF = F^{-1}D_BFF^{-1}D_AF = BA$.
- But since A and B are invertible, then $A = B$. □

Theorem

Given A and B two invertible circulant matrices, then A and B are braiding matrices only when $A = B$.

Proof:

- Assume that A and B are circulant matrices which braid.
- We recall that any circulant matrix is diagonalized by the Fourier matrix, F. Then, $A = F^{-1}D_A F$ and $B = F^{-1}D_B F$. Thus, $AB = F^{-1}D_A F F^{-1}D_B F = F^{-1}D_A D_B F$.
- Since two diagonal matrices always commute, $AB = F^{-1}D_B D_A F = F^{-1}D_B F F^{-1}D_A F = BA$.
- But since A and B are invertible, then $A = B$. □

- If only one of the two matrices is invertible, for example matrix A, and both A and B are commuting, then $ABA = BAB$ implies

$$ABA = ABB \implies BA = BB.$$

- A trivial solution is $B = 0$. Since there is no left cancellation since B is singular, $BA = BB$ may have nontrivial solutions.
- A class of nontrivial solutions of $BA = BB$ can be obtained by requiring B to be an idempotent singular matrix.
- Note that the minimal polynomial of an idempotent matrix, i.e., the polynomial ψ of the smallest degree such that $\psi(B) = 0$ is $\psi = B^2 - B$. With the exception of the identity matrix, every idempotent matrix is singular.

Thus, we assume idempotent matrices are always singular.

- If only one of the two matrices is invertible, for example matrix A, and both A and B are commuting, then $ABA = BAB$ implies

$$ABA = ABB \implies BA = BB.$$

- A trivial solution is $B = 0$. Since there is no left cancellation since B is singular, $BA = BB$ may have nontrivial solutions.
- A class of nontrivial solutions of $BA = BB$ can be obtained by requiring B to be an idempotent singular matrix.
- Note that the minimal polynomial of an idempotent matrix, i.e., the polynomial ψ of the smallest degree such that $\psi(B) = 0$ is $\psi = B^2 - B$. With the exception of the identity matrix, every idempotent matrix is singular.

Thus, we assume idempotent matrices are always singular.

- If only one of the two matrices is invertible, for example matrix A, and both A and B are commuting, then $ABA = BAB$ implies

$$ABA = ABB \implies BA = BB.$$

- A trivial solution is $B = 0$. Since there is no left cancellation since B is singular, $BA = BB$ may have nontrivial solutions.
- A class of nontrivial solutions of $BA = BB$ can be obtained by requiring B to be an idempotent singular matrix.
- Note that the minimal polynomial of an idempotent matrix, i.e., the polynomial ψ of the smallest degree such that $\psi(B) = 0$ is $\psi = B^2 - B$. With the exception of the identity matrix, every idempotent matrix is singular.

Thus, we assume idempotent matrices are always singular.

Theorem

Given a (singular) idempotent matrix B, $\exists A$, such that A and B commute, and A and B are solutions that braid.

Proof:

- Construct $A = aI + bB$. Since A is a polynomial function of B, then A and B must commute, i.e., $AB = BA$.
  Next, we need to determine a and b such that

$$(aI + bB)B(aI + bB) = B(aI + bB)B,$$
$$\implies a^2B + 2abB^2 + b^2B^3 = aB^2 + bB^2.$$

Since $BB = B$, then (13) reduces to

$$(a+b)^2B = (a+b)B, \tag{13}$$

yielding $(a+b) = (a+b)^2 \implies a+b = 1$, or $a+b = 0$. □

Theorem

Given a (singular) idempotent matrix B, $\exists A$, such that A and B commute, and A and B are solutions that braid.

Proof:

- Construct $A = aI + bB$. Since A is a polynomial function of B, then A and B must commute, i.e., $AB = BA$.
  Next, we need to determine a and b such that

$$(aI + bB)B(aI + bB) = B(aI + bB)B,$$
$$\implies a^2B + 2abB^2 + b^2B^3 = aB^2 + bB^2.$$

Since $BB = B$, then (13) reduces to

$$(a+b)^2B = (a+b)B, \tag{13}$$

yielding $(a+b) = (a+b)^2 \implies a+b = 1$, or $a+b = 0$. □

Theorem

Given a (singular) idempotent matrix B, $\exists A$, such that A and B commute, and A and B are solutions that braid.

Proof:

- Construct $A = aI + bB$. Since A is a polynomial function of B, then A and B must commute, i.e., $AB = BA$.
  Next, we need to determine a and b such that

$$(aI + bB)B(aI + bB) = B(aI + bB)B,$$
$$\implies a^2B + 2abB^2 + b^2B^3 = aB^2 + bB^2.$$

  Since $BB = B$, then (13) reduces to

$$(a+b)^2B = (a+b)B, \tag{13}$$

  yielding $(a+b) = (a+b)^2 \implies a+b = 1$, or $a+b = 0$. □

Theorem

Let matrix A be a polynomial function of degree n of an idempotent matrix B with $p(x) = a_0 + a_1 x + \ldots + a_n x^n$, then A and B braid if we have $a_0 + a_1 + \ldots + a_n = 1$ or if we have $a_0 + a_1 + \ldots + a_n = 0$.

Proof:

- In the previous, we have constructed A such that it is a linear polynomial function of B.
- If we can show that any polynomial function of B can be reduced to a polynomial of degree 1, then we can use the previous Theorem to prove that A and B braid, provided that A satisfies the coefficient requirement of the Theorem.

Theorem

Let matrix A be a polynomial function of degree n of an idempotent matrix B with $p(x) = a_0 + a_1 x + \ldots + a_n x^n$, then A and B braid if we have $a_0 + a_1 + \ldots + a_n = 1$ or if we have $a_0 + a_1 + \ldots + a_n = 0$.

Proof:

- In the previous, we have constructed A such that it is a linear polynomial function of B.
- If we can show that any polynomial function of B can be reduced to a polynomial of degree 1, then we can use the previous Theorem to prove that A and B braid, provided that A satisfies the coefficient requirement of the Theorem.

- Let $A = a_0 I + a_1 B + a_2 B^2 + \ldots + a_n B^n$, Since B is idempotent,

$$B^n = B^{n-2}BB = B^{n-2}B = B^{n-1}, \forall n \geq 2. \tag{14}$$

- Then $B^n = B$, yielding

$$\begin{aligned} A &= a_0 I + a_1 B + a_2 B + \ldots + a_n B \\ &= a_0 I + (a_1 + a_2 + \ldots + a_n) B \\ &= a_0 I + qB, \end{aligned}$$

where $q = a_1 + a_2 + \ldots + a_n$.

- If a_0 and q satisfy the algebraic expressions $a_0 + q = 1$ or $a_0 + q = 0$, then, by the proof of the previous Theorem, A and B braid. □

The more general case when A is an arbitrary function of B, can also be solved.

- Let $A = a_0 I + a_1 B + a_2 B^2 + \ldots + a_n B^n$, Since B is idempotent,

$$B^n = B^{n-2}BB = B^{n-2}B = B^{n-1}, \forall n \geq 2. \tag{14}$$

- Then $B^n = B$, yielding

$$\begin{aligned} A &= a_0 I + a_1 B + a_2 B + \ldots + a_n B \\ &= a_0 I + (a_1 + a_2 + \ldots + a_n) B \\ &= a_0 I + qB, \end{aligned}$$

where $q = a_1 + a_2 + \ldots + a_n$.

- If a_0 and q satisfy the algebraic expressions $a_0 + q = 1$ or $a_0 + q = 0$, then, by the proof of the previous Theorem, A and B braid. □

The more general case when A is an arbitrary function of B, can also be solved.

- Let $A = a_0 I + a_1 B + a_2 B^2 + \ldots + a_n B^n$, Since B is idempotent,

$$B^n = B^{n-2}BB = B^{n-2}B = B^{n-1}, \forall n \geq 2. \tag{14}$$

- Then $B^n = B$, yielding

$$\begin{aligned} A &= a_0 I + a_1 B + a_2 B + \ldots + a_n B \\ &= a_0 I + (a_1 + a_2 + \ldots + a_n) B \\ &= a_0 I + qB, \end{aligned}$$

  where $q = a_1 + a_2 + \ldots + a_n$.

- If a_0 and q satisfy the algebraic expressions $a_0 + q = 1$ or $a_0 + q = 0$, then, by the proof of the previous Theorem, A and B braid. □

The more general case when A is an arbitrary function of B, can also be solved.

- Let $A = a_0 I + a_1 B + a_2 B^2 + \ldots + a_n B^n$, Since B is idempotent,

$$B^n = B^{n-2}BB = B^{n-2}B = B^{n-1}, \forall n \geq 2. \tag{14}$$

- Then $B^n = B$, yielding

$$\begin{aligned} A &= a_0 I + a_1 B + a_2 B + \ldots + a_n B \\ &= a_0 I + (a_1 + a_2 + \ldots + a_n) B \\ &= a_0 I + qB, \end{aligned}$$

where $q = a_1 + a_2 + \ldots + a_n$.

- If a_0 and q satisfy the algebraic expressions $a_0 + q = 1$ or $a_0 + q = 0$, then, by the proof of the previous Theorem, A and B braid. □

The more general case when A is an arbitrary function of B, can also be solved.

- Consider matrix functions, i.e., functions on a matrix. This takes a scalar function f and a matrix $A \in \mathbb{C}^n \times \mathbb{C}^n$ and specifies $f(A)$ to be a matrix of the same dimensions as A.

Definition

A function f is said to be defined on the spectrum of A if the values

$$f^{(j)}(\lambda_i), \quad j = 0, \ldots, n_i - 1, \quad i = 1, \ldots, s \tag{15}$$

exist, where λ_i are eigenvalues of A with the corresponding Jordan block index n_i, and s is the number of distinct eigenvalues.

These are called the values of the function f on the spectrum of A

Theorem

Let f be defined on the spectrum of $S \in \mathbb{C}^{n\times n}$ and let ψ be the minimal polynomial of S. Then $f(S) = p(S)$, where p is the polynomial of degree less than

$$\sum_{i=1}^{s} n_i = \deg \psi,$$

that satisfies the interpolation conditions

$$p^{(j)}(\lambda_i) = f^{(j)}(\lambda_i), \quad j = 0,\ldots,n_i - 1, \quad i = 1,\ldots,s. \tag{16}$$

There is a unique such polynomial, p, and it is know as the Hermite interpolating polynomial.

- Indeed, this result can be taken as the definition of a matrix function via Hermite interpolation.

- For idempotent matrices, it is possible to construct functions of these matrices to extend the polynomial results obtained, and to obtain further solutions that are braiding matrices.

Theorem

Let matrix $A = f(B)$, where f satisfies the differentiability requirements for the interpolating Hermite polynomial, and B is an idempotent matrix. If $f(1) = 1$, or $f(1) = 0$, then A and B are braiding matrices.

- For idempotent matrices, it is possible to construct functions of these matrices to extend the polynomial results obtained, and to obtain further solutions that are braiding matrices.

Theorem

Let matrix $A = f(B)$, where f satisfies the differentiability requirements for the interpolating Hermite polynomial, and B is an idempotent matrix. If $f(1) = 1$, or $f(1) = 0$, then A and B are braiding matrices.

Proof:

- Since the idempotent matrix B is singular, the two eigenvalues of B are $\lambda_1 = 0$ and $\lambda_2 = 1$. Also, the degree of the minimal polynomial of B is 2.
- Thus, by Hermite interpolation, the polynomial p that interpolates $A = f(B)$ is a linear polynomial and is given by

$$p(B) = \frac{f(1) - f(0)}{1 - 0} B + f(0)I = (f(1) - f(0))B + f(0)I. \qquad (17)$$

  Thus, A is a linear polynomial of B, and A and B commute.
- We have shown that if A and B are braiding matrices when $(f(1) - f(0)) + f(0) = 1$, or when $(f(1) - f(0)) + f(0) = 0$, i.e., $f(1) = 1$, or $f(1) = 0$. □

Proof:

- Since the idempotent matrix B is singular, the two eigenvalues of B are $\lambda_1 = 0$ and $\lambda_2 = 1$. Also, the degree of the minimal polynomial of B is 2.
- Thus, by Hermite interpolation, the polynomial p that interpolates $A = f(B)$ is a linear polynomial and is given by

$$p(B) = \frac{f(1) - f(0)}{1 - 0} B + f(0)I = (f(1) - f(0))B + f(0)I. \tag{17}$$

Thus, A is a linear polynomial of B, and A and B commute.

- We have shown that if A and B are braiding matrices when $(f(1) - f(0)) + f(0) = 1$, or when $(f(1) - f(0)) + f(0) = 0$, i.e., $f(1) = 1$, or $f(1) = 0$. □

Proof:

- Since the idempotent matrix B is singular, the two eigenvalues of B are $\lambda_1 = 0$ and $\lambda_2 = 1$. Also, the degree of the minimal polynomial of B is 2.
- Thus, by Hermite interpolation, the polynomial p that interpolates $A = f(B)$ is a linear polynomial and is given by

$$p(B) = \frac{f(1) - f(0)}{1 - 0} B + f(0)I = (f(1) - f(0))B + f(0)I. \tag{17}$$

  Thus, A is a linear polynomial of B, and A and B commute.
- We have shown that if A and B are braiding matrices when $(f(1) - f(0)) + f(0) = 1$, or when $(f(1) - f(0)) + f(0) = 0$, i.e., $f(1) = 1$, or $f(1) = 0$. □

- Consider,

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix},$$

and let $A = f(B)$, where $f(x) = e^x$. Clearly, A and B are not braiding matrices, because A does not satisfy the condition of the Theorem, i.e., $f(1) \neq 1$ and $f(1) \neq 0$. On the other hand, if we choose f to be $\frac{e^x}{e}$, then

$$A = \begin{pmatrix} \frac{1}{e} & \frac{e-1}{e} \\ 0 & 1 \end{pmatrix},$$

and A and B are braiding matrices.

$$ABA = BAB = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

- We can write the braiding matrices in CFE form as

$$XA = BX, \tag{18}$$

where $X = AB$.

- Adding A to both sides of (18), we obtain $(X+I)A = BX + A$.
- Solving for A on the left side of the equation yields our fixed point iteration method

$$A = (X+I)^{-1}(BX+A). \tag{19}$$

Must ensure that $(X+I)$ remains invertible during the iteration.

- We could rewrite (19) as

$$A = (X+cI)^{-1}(BX+cA), \tag{20}$$

and for large enough values of c, the diagonal dominance of $(X+cI)$ will guarantee invertibility.

- We can write the braiding matrices in CFE form as

$$XA = BX, \qquad (18)$$

where $X = AB$.

- Adding A to both sides of (18), we obtain $(X + I)A = BX + A$.
- Solving for A on the left side of the equation yields our fixed point iteration method

$$A = (X + I)^{-1}(BX + A). \qquad (19)$$

Must ensure that $(X + I)$ remains invertible during the iteration.

- We could rewrite (19) as

$$A = (X + cI)^{-1}(BX + cA), \qquad (20)$$

and for large enough values of c, the diagonal dominance of $(X + cI)$ will guarantee invertibility.

- We can write the braiding matrices in CFE form as

$$XA = BX, \tag{18}$$

where $X = AB$.

- Adding A to both sides of (18), we obtain $(X+I)A = BX + A$.
- Solving for A on the left side of the equation yields our fixed point iteration method

$$A = (X+I)^{-1}(BX + A). \tag{19}$$

Must ensure that $(X+I)$ remains invertible during the iteration.

- We could rewrite (19) as

$$A = (X+cI)^{-1}(BX + cA), \tag{20}$$

and for large enough values of c, the diagonal dominance of $(X+cI)$ will guarantee invertibility.

- We can write the braiding matrices in CFE form as

$$XA = BX, \tag{18}$$

where $X = AB$.

- Adding A to both sides of (18), we obtain $(X+I)A = BX + A$.
- Solving for A on the left side of the equation yields our fixed point iteration method

$$A = (X+I)^{-1}(BX+A). \tag{19}$$

Must ensure that $(X+I)$ remains invertible during the iteration.

- We could rewrite (19) as

$$A = (X+cI)^{-1}(BX+cA), \tag{20}$$

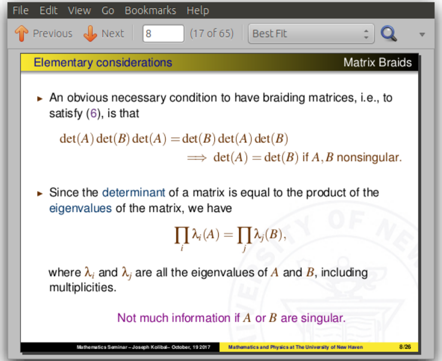
and for large enough values of c, the diagonal dominance of $(X+cI)$ will guarantee invertibility.

```
\begin{frame}
\Ft{Elementary considerations}
\begin{itemize}
\item
An obvious necessary condition to have braiding matrices, i.e., to
satisfy (\ref{eq:first}), is that
\begin{equation*}
\begin{aligned}
 \det(A)\det(B)\det(A)= &  \det(B)\det(A)\det(B) \\  \pause
 &  \implies \det(A) = \det(B)\mbox{ if } A, B \mbox{ nonsingular}.
\end{aligned}
\end{equation*}
\pause

\item
Since the \Urlx{https://wiki2.org/en/Determinant+Brights}{determinant}
of a matrix is equal to the product of the
Urlx{https://wiki2.org/en/Eigenvalues and eigenvectors+Brights}{eigenvalues}
of the matrix, we have
%
\begin{equation*}
 \prod_i{\lambda_{i}(A)} = \prod_j{\lambda_{j}(B)},
\end{equation*}
%
%
where $ \lambda_i$ and $ \lambda_j$ are all the eigenvalues of $ A$ and $ B$,
including multiplicities.
\vspace{0.8em}
\pause

{\color{\quotecolor}
\begin{center}
Not much information if $ A$ or $ B$ are singular.
\end{center}
}

\end{itemize}
\end{frame}
```



We used Beamer, a version of LaTeX that is highly optimized to produce quality presentation slides.
Interested? Consider MATH 2212 Software Tools for Math, along with some self-help research tools.

Back to **Elementary Considerations**, pg.8

The Department of Mathematics and Physics at The University of New Haven

http://math.newhaven.edu

JKolibal@newhaven.edu

http://math.newhaven.edu/kolibal

Alex Cibotarica, MS Thesis