

CIRCULANTS (Extract)

Alun Wyn-jones

Last revised in January 2008.

Please copy this book for your own reading only. Refers others to this website. Thank You.

CHAPTER 6.
Tensor Products.

6.1 Tensor Circulant Matrices

Tensor products of circulant matrices cannot in general be constructed by substituting circulant matrices for entries in circulant matrices. As Proposition 5.2.3 showed, the Kronecker product of circulant matrices is circulant iff the Kronecker product is subrepeating. In the first section of this chapter, another matrix algebra is constructed which includes the subrepeating circulant matrices as a subalgebra, and so is a generalization of the subrepeating circulants. Although matrices in the new algebra are not necessarily circulant, the algebra nevertheless enjoys some of the key properties of the circulant matrix algebra. In particular, it is simultaneously diagonalizable.

Although the generalization of subrepeating circulants supplies a tensor-like algebra, it suffers from a serious defect. One cannot take two arbitrary circulant matrices of given dimensions, and expect their tensor product to behave like a circulant matrix. Indeed, there is no solution to this requirement, but there is a solution if we restrict the dimensions of the circulants we take into a tensor product. In the second and subsequent sections of this chapter the possibility of a general tensor product is investigated, and a satisfactory tensor product is shown to exist whenever the dimensions of the factors in the tensor product are coprime.

6.1.1 The Kronecker Product. In this section, a matrix algebra is defined as the intersection of two Kronecker product algebras. This approach is taken from Davies's book (see [Dav4]) and is briefly summarised here.

The **Kronecker product** of a matrix $A = (a_{i,j})_{i,j}$ in $M_m(R)$ with a matrix $B \in M_n(R)$ is the matrix $A \times B \in M_{mn}(R)$ defined by

$$A \times B := \begin{pmatrix} a_{1,1}B & a_{1,2}B & \dots & a_{1,m}B \\ a_{2,1}B & a_{2,2}B & \dots & a_{2,m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}B & a_{m,2}B & \dots & a_{m,m}B \end{pmatrix}$$

The above definition shows a Kronecker product in block form; each entry is an $n \times n$ matrix.

We proceed by introducing two matrix algebras namely, $\text{CIRC}_m(M_n(R))$ and $M_m(\text{CIRC}_n(R))$ which are called respectively the **block circulant** ("circulant in blocks") and the **circulant block** ("blocks of circulants") matrices.

A nice example of a block circulant is a partitioned $mn \times mn$ circulant matrix. For instance, let $m = 3$ and $n = 2$, then the general 6×6 circulant matrix can be partitioned as follows.

$$\begin{pmatrix} a & b & c & d & e & f \\ f & a & b & c & d & e \\ e & f & a & b & c & d \\ d & e & f & a & b & c \\ c & d & e & f & a & b \\ b & c & d & e & f & a \end{pmatrix}$$

This matrix is circulant over $M_2(R)$! Unfortunately, block circulants are not generally circulant. For instance, the matrix below is a block circulant matrix but is not circulant.

$$\begin{pmatrix} a & b & e & f \\ c & d & g & h \\ e & f & a & b \\ g & h & c & d \end{pmatrix}$$

The block circulant $\text{CIRC}_m(0, I_n, 0, \dots, 0) = U_m \otimes I_n$ is analogous to the U matrix in the circulants. It can be shown that a matrix is a block circulant iff it commutes with $U_m \otimes I_n$. This is analogous to Corollary 2.1.1 for ordinary circulants. Also analogous to circulants, is that block circulants are diagonalized by the Kronecker product $F_m \otimes F_n$ where F_n is the $n \times n$ Fourier matrix. However, the diagonal entries are not scalar, rather they are general $n \times n$ matrices.

For the circulant block matrices, on the other hand, the centralizing matrix is $I_m \otimes U_n$. It is the matrix with U_n blocks down the main diagonal and zeroes elsewhere. It can be shown that a matrix M is a circulant block matrix iff it commutes with $I_m \otimes U_n$ iff $(F_m \otimes F_n)^\dagger M (F_m \otimes F_n)$ is in $M_m(\text{Diag}_n(R))$ where $\text{Diag}_n(R)$ is the set of all $n \times n$ diagonal matrices over R .

There is an apparent problem with this development. Block circulants are circulant over a non-commutative ring and hence form a non-commutative algebra. Consequently much of the theory of circulants is inapplicable to block circulants. On the other hand, circulant block matrices do have commutative entries, but they are non-commutative because matrix multiplication is generally non-commutative. However, the intersection of the two algebras, $\text{CIRC}_m(M_n(R)) \cap M_m(\text{CIRC}_n(R))$ is commutative and members of this joint algebra are simultaneously diagonalizable by the unitary matrix $F_m \otimes F_n$.

The subrepeating circulants have a special rôle here. From the definitions (see 5.2.2) we have

$$\text{QR}_{mn}^n \subset \text{CIRC}_m(M_n(R)) \cap M_m(\text{CIRC}_n(R)) = \text{CIRC}_m(\text{CIRC}_n(R))$$

and by Proposition 5.2.3, we have

$$\begin{aligned} \text{QR}_{mn}^n &= \text{CIRC}_{mn}(R) \cap M_m(\text{CIRC}_n(R)) \\ \therefore \text{QR}_{mn}^n &= \text{CIRC}_{mn}(R) \cap \text{CIRC}_m(M_n(R)) \cap M_m(\text{CIRC}_n(R)) \\ &= \text{CIRC}_{mn}(R) \cap \text{CIRC}_m(\text{CIRC}_n(R)) \end{aligned}$$

In summary, the above approach bestows upon suitable Kronecker products all the main properties of circulant matrices. However, the approach still suffers from the problem that the Kronecker product of general circulants is not circulant. The second approach will remedy this at the cost that the tensor product is not defined for all dimensions m and n .

6.2 General Tensor Products of Circulant Matrices.

It is sometimes erroneously assumed that a tensor product of matrices must be a Kronecker product. This is not so, the Kronecker product is merely the simplest tensor product. To find a suitable tensor product for circulants, we shall start with a set of desiderata for such a general tensor product, and then construct a tensor product which satisfy these desiderata.

A tensor product of matrices A and B is a matrix T which is constructed by a rule which assigns to $T_{r,s}$ the product of pairs of entries from A and B , $A_{i,j}B_{k,l}$, say. The rule is a map from pairs of matricial indices $((i,j), (k,l))$ to indices of the tensor product matrix, (r,s) . To qualify as a tensor product on A and B , the domain of the map should be the set of all pairs of indices for A and B . For the tensor matrix to be fully defined, the range of the map must be the set of all indices of the tensor product matrix. For the tensor matrix to be well-defined, the map must be one-to-one. There is one other criterion, the so-called product rule: Given matrices A_1, A_2 and B_1, B_2 then $(A_1 A_2) \otimes (B_1 B_2)$ must equal $(A_1 \otimes B_1)(A_2 \otimes B_2)$.

Now we apply these criteria to circulant matrices. We need to find a tensor product of $\text{CIRC}_m(R)$ and $\text{CIRC}_n(R)$ so that $\text{CIRC}_m(R) \otimes \text{CIRC}_n(R) \subset \text{CIRC}_N(R)$ for some N , and there must be a one-to-one map $\phi: \mathbb{Z}_N^2 \rightarrow \mathbb{Z}_m^2 \times \mathbb{Z}_n^2$ such that $T_{x,y} = A_{\phi_1(x,y)} B_{\phi_2(x,y)}$ where $\phi_1 \times \phi_2 = \phi$.

6.2.1 Theorem A tensor product $\text{CIRC}_m(R) \otimes \text{CIRC}_n(R)$ can be defined such that $T = A \otimes B$ is circulant in $\text{CIRC}_N(R)$ for some N and for every $A \in \text{CIRC}_m(R)$, $B \in \text{CIRC}_n(R)$ if and only if m and n are coprime. Let $T = \text{CIRC}(t)$, $A = \text{CIRC}(a)$, and $B = \text{CIRC}(b)$. Then, $t_x = a_{\phi_1(x)} b_{\phi_2(x)}$ with $\phi = \phi_1 \oplus \phi_2$ an additive isomorphism, $\phi: \mathbb{Z}_N \rightarrow \mathbb{Z}_m \oplus \mathbb{Z}_n$. In particular, $N = nm$.

Proof. Take the most general possible definition of $T = A \otimes B$ and assume that all three matrices are circulant.

$$T_{x,y} := A_{\nu_1(x,y),\rho_1(x,y)} B_{\nu_2(x,y),\rho_2(x,y)} \quad \text{where } \nu_1(x,y), \rho_1(x,y) \in \mathbb{Z}_m, \text{ and } \nu_2(x,y), \rho_2(x,y) \in \mathbb{Z}_n.$$

$$\therefore t_{y-x} = a_{\rho_1(x,y)-\nu_1(x,y)} b_{\rho_2(x,y)-\nu_2(x,y)} \quad (1)$$

$$\therefore t_y = a_{\rho_1(0,y)-\nu_1(0,y)} b_{\rho_2(0,y)-\nu_2(0,y)}$$

$$\therefore t_{y-x} = a_{\rho_1(0,y-x)-\nu_1(0,y-x)} b_{\rho_2(0,y-x)-\nu_2(0,y-x)} \quad (2)$$

Identifying terms in equations (1) and (2),

$$\therefore \rho_1(x,y) - \nu_1(x,y) = \rho_1(0,y-x) - \nu_1(0,y-x) \quad \text{with a similar equation for } \rho_2 \text{ and } \nu_2.$$

This shows that $\rho_1(x,y) - \nu_1(x,y)$ is a function of $y-x$. Let $\rho_1(x,y) - \nu_1(x,y) = \phi_1(y-x)$ and let $\rho_2(x,y) - \nu_2(x,y) = \phi_2(y-x)$ then

$$t_x = a_{\phi_1(x)} b_{\phi_2(x)}$$

We now apply the product rule. Let $t, t' \in \mathbb{Z}_N$ and let $a, a' \in \mathbb{Z}_m, b, b' \in \mathbb{Z}_n$.

$$\begin{aligned} (t * t')_x &= \sum_{y \in \mathbb{Z}_N} t_y t'_{x-y} \\ &= \sum_{y \in \mathbb{Z}_N} a_{\phi_1(y)} b_{\phi_2(y)} a'_{\phi_1(x-y)} b'_{\phi_2(x-y)} \end{aligned} \quad (3)$$

$$= (a * a')_{\phi_1(x)} (b * b')_{\phi_2(x)} \quad \text{by the product rule}$$

$$\begin{aligned} &= \left(\sum_{j \in \mathbb{Z}_m} a_j a'_{\phi_1(x)-j} \right) \left(\sum_{k \in \mathbb{Z}_n} b_k b'_{\phi_2(x)-k} \right) \\ &= \sum_{j \in \mathbb{Z}_m} \sum_{k \in \mathbb{Z}_n} a_j b_k a'_{\phi_1(x)-j} b'_{\phi_2(x)-k} \end{aligned} \quad (4)$$

Comparing (3) and (4), we see that the (j, k) subscripts in expression (4) must range over the same set as $(\phi_1(y), \phi_2(y))$ respectively in expression (3).

$$\therefore \phi = \phi_1 \times \phi_2 : \mathbb{Z}_N \rightarrow \mathbb{Z}_m \times \mathbb{Z}_n \quad \text{onto}$$

Since summation order is immaterial, we can substitute $\phi_1(y)$ for j and $\phi_2(y)$ for k in expression (4) and then identify terms with expression (3) showing that,

$$\phi_1(x) - \phi_1(y) = \phi_1(x-y)$$

$$\phi_2(x) - \phi_2(y) = \phi_2(x-y)$$

That is, $\phi = \phi_1 \times \phi_2$ is a linear map and so is an additive group homomorphism. $\therefore \phi = \phi_1 \oplus \phi_2$. It was shown that ϕ must be onto, and by the tensor requirements, ϕ must be one-to-one, therefore, ϕ is an additive isomorphism. But, \mathbb{Z}_N is cyclic, whereas $\mathbb{Z}_m \oplus \mathbb{Z}_n$ cyclic if and only if m, n are coprime. Therefore, m, n are coprime and $N = mn$. \square

6.2.2 Single Residue Definition of ϕ and Tensor Eigenvalues The ϕ map can be fully defined by a single residue, g , in \mathbb{Z}_{mn}^* .

$$\phi_g(x) = (gx \bmod m) \oplus (gx \bmod n)$$

With this definition of the map, the relationship of the eigenvalues of $A \otimes B$ to the eigenvalues of A and B can be found quite easily. Consider a product of arbitrary eigenvalues from A and B .

$$\begin{aligned} \lambda_i(A)\lambda_j(B) &= \sum_{k \in \mathbb{Z}_m} a_k \zeta_m^{ik} \sum_{l \in \mathbb{Z}_n} b_l \zeta_n^{jl} \\ &= \sum_{k \in \mathbb{Z}_m} \sum_{l \in \mathbb{Z}_n} a_k b_l \zeta_{mn}^{ikn+jlm} \\ &= \sum_{x \in \mathbb{Z}_{mn}} a_{(gx \bmod m)} b_{(gx \bmod n)} \zeta_{mn}^{gxin+gxm} \\ &= \sum_{x \in \mathbb{Z}_{mn}} t_x \zeta_{mn}^{x(gin+gjm)} \\ &= \lambda(T)_{g(in+jm)} \end{aligned}$$

The result is a general eigenvalue of $T = A \otimes B$. So (as one would expect), each eigenvalue of the tensor product is the product of eigenvalues of the tensor factors. The map relating the eigenvalues is most simply given by

$$\tilde{\phi}(in + jm) = (i, j)$$

This map can also be defined in terms of a single residue, $h \in \mathbb{Z}_{mn}$.

$$\tilde{\phi}(x) = (hx \bmod m, hx \bmod n)$$

where $h = g^{-1}(\bar{n}^2 n + \bar{m}^2 m)$, \bar{n} is the inverse residue of $n \pmod{m}$, and \bar{m} is the inverse residue of $m \pmod{n}$.

6.2.3 Example Let $N = 12$, $m = 3$, $n = 4$, and take $g = 1$. We have,

$$\begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix} \otimes \begin{pmatrix} d & e & f & g \\ g & d & e & f \\ f & g & d & e \\ e & f & g & d \end{pmatrix} = \begin{pmatrix} ad & be & cf & ag & bd & ce & af & bg & cd & ae & bf & cg \\ cg & ad & be & cf & ag & bd & ce & af & bg & cd & ae & bf \\ bf & cg & ad & be & cf & ag & bd & ce & af & bg & cd & ae \\ ae & bf & cg & ad & be & cf & ag & bd & ce & af & bg & cd \\ cd & ae & bf & cg & ad & be & cf & ag & bd & ce & af & bg \\ bg & cd & ae & bf & cg & ad & be & cf & ag & bd & ce & af \\ af & bg & cd & ae & bf & cg & ad & be & cf & ag & bd & ce \\ ce & af & bg & cd & ae & bf & cg & ad & be & cf & ag & bd \\ bd & ce & af & bg & cd & ae & bf & cg & ad & be & cf & ag \\ ag & bd & ce & af & bg & cd & ae & bf & cg & ad & be & cf \\ cf & ag & bd & ce & af & bg & cd & ae & bf & cg & ad & be \\ be & cf & ag & bd & ce & af & bg & cd & ae & bf & cg & ad \end{pmatrix}$$

or, more succinctly,

$$\text{CIRC}_3(a, b, c) \otimes \text{CIRC}_4(d, e, f, g) = \text{CIRC}_{12}(ad, be, cf, ag, bd, ce, af, bg, cd, ae, bf, cg)$$

6.3 Tensors of Supercirculants. Within the supercirculants, the tensor product of circulant subalgebras of coprime orders takes a particularly intuitive form. Since the supercirculants contain all circulant

algebras, the tensor product must also be a subalgebra. Clearly, it is one of order mn . Also, in the super-circulants (sticking with the matricial circulants), $U_{mn}^{nx} = U_m^x$. So, it is natural to try the map

$$U_m^i \otimes U_n^j \rightarrow U_{mn}^{ni} U_{mn}^{mj} = U_{mn}^{ni+mj}$$

This map works; it is equivalent to that defined by $\phi_g : \mathbb{Z}_{mn} \rightarrow \mathbb{Z}_m \oplus \mathbb{Z}_n$ where $g = \bar{n}^2 n + \bar{m}^2 m \in \mathbb{Z}_{mn}$, and \bar{n} and \bar{m} are inverse residues modulo m and n respectively. Then, $\phi_g(ni + mj) \mapsto (i, j)$. Hence,

$$\text{if } A = \sum_{i \in \mathbb{Z}_m} a_i U_m^i \in \text{CIRC}_m, \text{ and } B = \sum_{j \in \mathbb{Z}_n} b_j U_n^j \in \text{CIRC}_n,$$

$$\text{then } A \otimes B = \sum_{i \in \mathbb{Z}_m} a_i (U_{mn}^n)^i \sum_{j \in \mathbb{Z}_n} b_j (U_{mn}^m)^j = AB \in \text{CIRC}_{mn}$$

In other words, with this choice of ϕ_g , the tensor product is the circulant product. This applies equally to the eigenvalues with componentwise product, so we also get a simple formula for the tensor product in the eigenspace:

$$\lambda(A \otimes B) = \lambda(A) \lambda(B)$$

In ordinary circulants, this translates to

$$\lambda_x^{(mn)}(A \otimes B) = \lambda_{(x \bmod m)}^{(m)}(A) \lambda_{(x \bmod n)}^{(n)}(B)$$

We have demonstrated the following.

6.3.1 Theorem Let m, n be coprime.

(i) $\text{circ}_m(R) \otimes \text{circ}_n(R) \approx \text{circ}_{mn}(R)$

(ii) The isomorphism is given by $\phi : u_m^i \otimes u_n^j \mapsto u_{mn}^{in+jm}$.

(iii) In the eigenspace, the map is given by $\tilde{\phi} : \lambda_x^{(mn)}(a \otimes b) \mapsto \lambda_{(x \bmod m)}^{(m)}(a) \lambda_{(x \bmod n)}^{(n)}(b)$. \square

A simple consequence of the construction of the tensor product is

6.3.2 Corollary Let $A \in \text{CIRC}_m, B \in \text{CIRC}_n$.

If $T = A \otimes B$ is $mn \times mn$ circulant then $\det T = (\det A)^n (\det B)^m$. \square

How do we recognize a tensor product matrix? The following lemma gives a necessary condition which is useful for sparse matrices.

6.3.3 Lemma Let $T = \text{CIRC}_{mn}(t) = A \otimes B$ be circulant with $A = \text{CIRC}_m(a)$ and $B = \text{CIRC}_n(b)$ and suppose a has x non-zero components and b has y non-zero components. Then, t has xy non-zero components. \square

6.3.4 Corollary If $T = A \otimes B$ has a prime number, p , say, of non-zero components, then either a has p non-zero components and b has 1 or vice versa. \square

6.4 Tensor Products and Polynomials in Several Variables

The homomorphism $\Gamma^n : R[x] \rightarrow \text{circ}_n(R)$ of §3.3 will be generalized here to $\Gamma^{m,n} : R[x, y] \rightarrow \text{circ}_{mn}(R)$.

The ring $R[x, y]$ is the tensor product $R[x] \otimes_R R[y]$ in a natural way: $(a_i x^i) \otimes (b_j y^j)$ is identified with $a_i x^i b_j y^j$. For m and n coprime, this suggests a definition of the map $\Gamma^{m,n}$ by requiring that it maps tensor products to the tensor products of §6.4. Thus

$$\Gamma^{m,n}(x^i \otimes y^j) = u_m^i \otimes u_n^j$$

$\Gamma^{m,n}$ is then extended to all of $R[x, y]$ by linearity, so that $\Gamma^{m,n} : \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} a_{i,j} x^i y^j \mapsto \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} a_{i,j} u_{mn}^{in+jm}$

Here are another two though equivalent descriptions of $\Gamma^{m,n}$.

(i) It can be described as a substitution map

$$\Gamma^{m,n} : a(x, y) \mapsto a(u_{mn}^n, u_{mn}^m)$$

(ii) $\Gamma^{m,n}$ can be defined as the composition of $\tilde{\Gamma}_m^{mn} \Gamma^m$, acting on the variable x , with $\tilde{\Gamma}_n^{mn} \Gamma^n$, acting on the variable y .

$$\tilde{\Gamma}_m^{mn} \Gamma^m|_x \left(\tilde{\Gamma}_n^{mn} \Gamma^n|_y (a(x, y)) \right) = \tilde{\Gamma}_n^{mn} \Gamma^n (a(x, u_{mn}^m)) = a(u_{mn}^n, u_{mn}^m) = \Gamma^{m,n}(a)$$

By Proposition 3.5.3, $\tilde{\Gamma}_m^{mn} \Gamma^m = \Gamma^{mn} \epsilon_x^m$ where $\epsilon_x^m : x \mapsto x^m$. Therefore, from the above,

$$\Gamma^{m,n} = \tilde{\Gamma}_m^{mn} \Gamma^m|_x \tilde{\Gamma}_n^{mn} \Gamma^n|_y = \Gamma^{mn}|_x \Gamma^{mn}|_y \epsilon_x^m \epsilon_y^n = \Gamma^{mn} S(x, y) \epsilon_x^m \epsilon_y^n$$

where $S(x, y) : f(x, y) \mapsto f(x, x)$.

6.4.1 Proposition Let $\Gamma^{m,n} : R[x, y] \rightarrow \mathbf{circ}_{mn}(R)$ be the map $\Gamma^{m,n} a(x, y) = a(u_{mn}^m, u_{mn}^n)$. Then

- (i) $\Gamma^{m,n}$ is a ring homomorphism,
- (ii) $\ker \Gamma^{m,n} = \ker \Gamma^m \vee \ker \Gamma^n = (x^m - 1) + (y^n - 1)$, and
- (iii) $\Gamma^{m,n}(R[x, y]) = \tilde{\Gamma}_{mn/d}^{mn}(\mathbf{circ}_{mn/d}(R))$ where $d = \gcd(m, n)$.

Proof.

(i) By the description in 6.4(i) above, $\Gamma^{m,n}$ is a substitution map, and so must be a ring homomorphism. QED (i)

(ii) Denote the ideal $(x^m - 1) + (y^n - 1)$ by J . It is trivial that $\Gamma^{m,n}(x^m - 1) = \Gamma^{m,n}(y^n - 1) = 0$,
 $\therefore J \subset \ker \Gamma^{m,n}$.

Now assume $a(x, y) \in \ker \Gamma^{m,n}$. Let $\bar{a}(x, y)$ be the polynomial $a(x, y)$ reduced modulo the ideal J so that all powers of x and y in \bar{a} are less than m and n respectively. Since $J \subset \ker \Gamma^{m,n}$, $\bar{a}(x, y) \in \ker \Gamma^{m,n}$. Since $\deg_x(\bar{a}) < m$ and $\deg_y(\bar{a}) < n$, $\Gamma^{m,n}$ maps \bar{a} to the vector in $\mathbf{circ}_{mn}(R)$ whose components equal the coefficients of \bar{a} which are therefore all zero since $\bar{a} \in \ker \Gamma^{m,n}$. Hence, $a \equiv 0 \pmod{J}$. Therefore, $a(x, y) \in \ker \Gamma^{m,n} \Rightarrow a(x, y) \in J$. QED (ii)

(iii) Let $\gcd(m, n) = d$, then $\Gamma^{m,n} : x^i y^j \mapsto u_{mn}^{ni+mj} = u_{mn}^{d(in'+jm')}$ where $m' = m/d$, $n' = n/d$. \square