Can. J. Math. Vol. **48** (1), 1996 pp. 196–209

POSITIVE POWERS OF POSITIVE POSITIVE DEFINITE MATRICES

LON ROSEN

ABSTRACT. Let C be an $n \times n$ positive definite matrix. If $C \geq 0$ in the sense that $C_{ij} \geq 0$ and if $p > n-2$, then $C^p \geq 0$. This implies the following "positive minorant property" for the norms $\|A\|_p = [\mathrm{tr}(A^*A)^{p/2}]^{1/p}$. Let $2 < p \neq 4, 6, \ldots$. Then $0 \leq A \leq B \Rightarrow \|A\|_p \leq \|B\|_p$ if and only if $n < p/2 + 1$.

1. **Introduction.** If C is an $n \times n$ positive definite matrix (we write $C \in \mathcal{P}_n$) and if the entries of C are nonnegative (we write $C \geq 0$), we call C a positive positive definite matrix (and we write $C \in \mathcal{P}_n^+$). Is $\mathcal{P}_n^+$ closed under taking powers? Since it's easy to see that $\mathcal{P}_n$ is, our question becomes:

$$C \in \mathcal{P}_n^+, q > 0 \overset{?}{\Rightarrow} C^q \geq 0. \tag{1.1}$$

Of course, (1.1) is completely trivial if q is an integer. However, one should not be misled by this trivial case. If $q < n-2$ and q is not an integer, counterexamples have been discovered to (1.1) [1, 4]. Thus the following Theorem, our main result, is "best possible":

THEOREM 1.1. *If $C \in \mathcal{P}_n^+$ and $q > n-2$, then $C^q \geq 0$.*

This theorem has been proved in the case $n = 3$ by Virot and by Déchamps-Gondim, Lust-Piquard and Queffelec [1], but, as far as we know, it is new for $n > 3$. While preparing this paper we learned that Weissenhofer (private communication) has proved (1.1) under the stronger assumption

$$q > (n-1)\left[\frac{n-1}{2}\right].$$

Actually, our interest in the theorem arose from the so-called "positive minorant property" which states that the p-norm of the $n \times n$ matrix A,

$$\|A\|_p = [\mathrm{tr}(A^*A)^{p/2}]^{1/p},$$

is monotonic in A in the sense that

$$0 \leq A \leq B \Rightarrow \|A\|_p \leq \|B\|_p. \tag{1.2}$$

The question of whether (1.2) is true has attracted considerable attention. Actually, people considered the stronger "minorant property" in which A is not required to be positive

Received by the editors September 15, 1994; revised September 13, 1995.
AMS subject classification: Primary 15A48, secondary 47A63, 60J15.
© Canadian Mathematical Society 1996.

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

but only to satisfy $|A_{ij}| \leq B_{ij}$. The counterexamples of Peller [2] and Simon [3] for large n apparently came as quite a surprise. But it wasn't clear exactly how large n had to be, and Simon asked: What is the critical value of n?

The connection between Theorem 1.1 and the positive minorant property was given in [1]: Letting $A(t) = A + t(B - A) \geq 0$ interpolate between A and B for $0 \leq t \leq 1$, we compute that

$$\frac{d}{dt}\|A(t)\|_p = \|A(t)\|_p^{1-p} \operatorname{tr}[C(t)^{p/2-1}A(t)^*(B-A)] \tag{1.3}$$

where $C(t) = A(t)^*A(t) \in \mathcal{P}_n^+$. Knowing that $C(t)^{p/2-1} \geq 0$, we conclude that (1.3) is nonnegative and that the monotonicity (1.2) holds. According to Theorem 1.1, this is so if $p/2+1 > n$, whereas if $p/2+1 < n$ (and $p \neq 2, 4, \ldots$) the counterexample of [4] denies (1.2). Thus we have answered Simon's question for the positive minorant property:

THEOREM 1.2. *If $2 < p \neq 4, 6, \ldots$ (1.2) holds for $n \times n$ matrices if and only if $n < p/2 + 1$.*

We now outline our strategy for proving Theorem 1.1. It suffices to consider the case where C is strictly positive definite (otherwise perturb it to $C + \varepsilon I$) and where $n - 2 < q < n - 1$, *i.e.*,

$$q = n - 1 - \alpha, \qquad 0 < \alpha < 1. \tag{1.4}$$

(The case $q \in \mathbb{Z}$ is trivial and the case $q > n - 1$ follows from (1.4) by writing $C^q = C^m C^{q-m}$ where $m = [q - n + 2]$.) By the "resolvent formula" [5, p. 260] for non-integer powers

$$C^q = \frac{\pi}{\sin \pi\alpha} \int_0^\infty C^{n-1}(C+\lambda)^{-1}\lambda^{-\alpha}\, d\lambda. \tag{1.5}$$

Throughout our proof, we subscribe to the power of positive thinking: in order to show that an integral or sum is positive we optimistically inquire whether the integrand or summand itself is positive. Accordingly:

THEOREM 1.3. *If $C \in \mathcal{P}_n^+$ and $\lambda > 0$, then*

$$C^{n-1}(C+\lambda)^{-1} \geq 0. \tag{1.6}$$

Obviously, Theorem 1.1 follows from (1.5) and (1.6). Moreover, it follows from (1.6) that any function on $(0, \infty)$ of the form

$$f(x) = x^m \int_0^\infty (x+\lambda)^{-1} g(\lambda)\, d\lambda, \tag{1.7}$$

where $m \geq n - 1$, $g(\lambda) \geq 0$, and g is suitably regular so that (1.7) converges, will also satisfy

$$C \in \mathcal{P}_n^+ \Rightarrow f(C) \geq 0.$$

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

Let $d(\lambda) = \det(C+\lambda)$. As we compute in Lemma 3.1,

$$C^{n-1}(C+\lambda)^{-1} = d(\lambda)^{-1}Q_n(C,\lambda) \tag{1.8}$$

where Q_n is a polynomial in C and λ, of degree $n-1$ in each,

$$Q_n(C,\lambda) = C^{n-1}\lambda^{n-1} + \sum_{i=0}^{n-2} q_{ni}(C)\lambda^i,$$

where $q_{ni}(C)$ is a ploynomial of degree $n-2$. For example,

$$Q_3 = C^2\lambda^2 + (s_2C - s_3)\lambda + s_3C$$

and

$$Q_4 = C^3\lambda^3 + (s_2C^2 - s_3C + s_4)\lambda^2 + (s_3C^2 - s_4C)\lambda + s_4C^2$$

where s_j is the jth degree symmetric polynomial in the eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \cdots \lambda_n$ of C:

$$s_j = \sum_{1\leq k_1<k_2<\cdots<k_j\leq n} \lambda_{k_1}\lambda_{k_2}\cdots\lambda_{k_j}. \tag{1.9}$$

Thinking positively, we then show that

$$q_{ni}(C) \geq 0. \tag{1.10}$$

For example,

$$\begin{aligned}\big(q_{31}(C)\big)_{kk} &= s_2C_{kk} - s_3 \geq s_2\lambda_1 - s_3\\ &= (\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)\lambda_1 - \lambda_1\lambda_2\lambda_3 > 0.\end{aligned}$$

It turns out that it's relatively easy to show that the diagonal elements of $q_{ni}(C)$ are positive (Lemma 3.2). The off-diagonal elements present more of a challenge. We illustrate our solution in the elementary case of

$$q_{42} = s_2C^2 - s_3C + s_4.$$

For $k_0 \neq k_1$,

$$(C^2)_{k_0k_1} = C_{k_0k_1}(C_{k_0k_0} + C_{k_1k_1}) + \sum_{\beta\neq k_0,k_1} C_{k_0\beta}C_{\beta k_1}. \tag{1.11}$$

It follows that

$$\begin{aligned}(q_{42})_{k_0k_1} &\geq s_2C_{k_0k_1}(C_{k_0k_0} + C_{k_1k_1}) - s_3C_{k_0k_1}\\ &\geq C_{k_0k_1}[s_2(\lambda_1+\lambda_2) - s_3] \geq 0\end{aligned}$$

since

$$C_{k_0k_0} + C_{k_1k_1} \geq \lambda_1 + \lambda_2 \tag{1.12}$$

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

and

$$s_2(\lambda_1 + \lambda_2) \geq s_3. \tag{1.13}$$

The reader ought to be persuaded by this elementary example, as the author was, that the inequality (1.10) is not too far-fetched. However, the case of general n is not so transparent. Technically the heart of the paper is Section 2 where we generalize (1.11) by establishing a curious self-avoiding walk representation for $(C^l)_{k_0k_1}$ in terms of certain determinantal objects $D^l_{\vec{k}}$ (see (2.1) and (2.4)). Inequalities like (1.12) are too simple-minded for the general case and we replace them by a "spectral" representation for $D^l_{\vec{k}}$ (see (3.8)–(3.10)). Finally, the inequality (1.13) is representative of a whole family of inequalities involving the symmetric functions s_j, which we derive in Lemma 3.3.

ACKNOWLEDGEMENTS. I wish to thank Stephen Weissenhofer and John Fournier for introducing me to the positive minorant problem and for many useful discussions, David Brydges for suggesting the use of generating functions, and Greg Lawler for his random walk expertise.

2. **Matrix Identities.** In this section we develop a representation for powers C^ℓ of an $n \times n$ matrix C in terms of the determinantal constructs

$$D^\ell_{\vec{k}} = \sum_{\substack{\ell_1,\ldots,\ell_m=0 \\ \ell_1+\cdots+\ell_m=\ell}}^{\ell} \sum_{\sigma\in S_m} \operatorname{sgn}\sigma \prod_{i=1}^{m} (C^{\ell_i})_{k_i k_{\sigma(i)}}, \tag{2.1}$$

where $\vec{k} = (k_1, \ldots, k_m) \in I^m = \{1, \ldots, n\}^m$ and S_m is the symmetric group of $\{1, \ldots, m\}$. It's easy to see that $D^\ell_{\vec{k}}$ is a symmetric function of $k_1, \ldots, k_m$, and that

$$D^\ell_{\vec{k}} = 0 \text{ unless } \vec{k} \in I^m_0 = \{\vec{k} \in I^m \mid k_i \neq k_j \text{ if } i \neq j\} \tag{2.2}$$

$$D^0_{\vec{k}} = \begin{cases} 1 & \text{if } \vec{k} \in I^m_0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

For example, for $\vec{k} \in I^m_0$,

$$D^2_{\vec{k}} = \sum_i (C^2)_{k_i k_i} + \sum_{i<j} (C_{k_i k_i} C_{k_j k_j} - C_{k_i k_j} C_{k_j k_i}).$$

We let

$$C^j_{k_1 k_2 \cdots k_{j+1}} = C_{k_1 k_2} C_{k_2 k_3} \cdots C_{k_j k_{j+1}},$$

and we adopt the convention that repeated Greek indices are summed from 1 to n.

THEOREM 2.1. *a) For* $k_1 \neq k_2$

$$(C^\ell)_{k_1 k_2} = \sum_{j=1}^{\ell} C^j_{k_1 \alpha_1 \cdots \alpha_{j-1} k_2} D^{\ell-j}_{(k_1, \alpha_1, \ldots, \alpha_{j-1}, k_2)}. \tag{2.4}$$

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

b)

$$(C^{\ell})_{k_1k_1} = \sum_{j=1}^{\ell} C^{j}_{k_1\alpha_1\cdots\alpha_{j-1}k_1} D^{\ell-j}_{(k_1,\alpha_1\cdots\alpha_{j-1})}. \tag{2.5}$$

REMARKS. 1. We are grateful to David Brydges for suggesting the possibility of proving this Theorem using generating functions. Our original proof involved a complicated inductive argument on ℓ.

2. The identities (2.4) and (2.5) have an interesting interpretation in terms of random walks on I. The left side of (2.4), $(C^{\ell})_{k_1k_2} = \sum_{\alpha_1,\ldots,\alpha_{\ell-1}} C^{\ell}_{k_1\alpha_1\cdots\alpha_{\ell-1}k_2}$, can be regarded as a sum over all ℓ-step walks on I from k_1 to k_2. In view of (2.2), the right side is a sum over *self-avoiding* walks with the $D^{\ell-j}$'s representing the contributions of attached loops (see the Appendix). The usefulness of this representation for us is a certain positivity property enjoyed by the D^{ℓ}'s (see (3.10)).

We introduce the generating function

$$D(t,\vec{\lambda}) = 1 + \sum_{m=1}^{n} \sum_{\vec{k}\in I_0^m} \sum_{\ell=0}^{\infty} D^{\ell}_{\vec{k}} t^{\ell} \lambda_{\vec{k}} \tag{2.6}$$

where $t \in \mathbb{R}$, $\vec{\lambda} \in \mathbb{R}^n$, and $\lambda_{\vec{k}} = \Pi_{i=1}^{m} \lambda_{k_i}$. The infinite sum over ℓ converges for $|t| < ||C||^{-1}$, and we have

$$\sum_{\ell=0}^{\infty} D^{\ell}_{\vec{k}} t^{\ell} = \sum_{\sigma\in S_m} \operatorname{sgn}\sigma \prod_{i=1}^{m} R(t)_{k_ik_{\sigma(i)}} \tag{2.7}$$

where

$$R(t) = (1-tC)^{-1}.$$

Letting Λ be the diagonal matrix with entries $\lambda_1,\ldots,\lambda_n$, we see from (2.6) and (2.7) that

$$D(t,\vec{\lambda}) = 1 + \sum_{m=1}^{n} \sum_{\vec{k}\in I_0^m} \sum_{\sigma\in S_m} \operatorname{sgn}\sigma \prod_{i=1}^{m} (\Lambda R)_{k_ik_{\sigma(i)}}, \tag{2.8}$$

whence:

LEMMA 2.2. *For* $|t| < ||C||^{-1}$,

$$D(t,\vec{\lambda}) = \det\big(1 + \Lambda R(t)\big). \tag{2.9}$$

PROOF. Expanding $\det\big(1 + \Lambda R(t)\big)$ according to the number of elements m of ΛR used, $m = 0, 1, \ldots, n$, we obtain the right hand side of (2.8). ∎

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

For a formal power series $F(\vec{\lambda})$ in $\vec{\lambda} = (\lambda_1, \ldots, \lambda_n)$ and $\vec{\lambda}^{-1} = (\lambda_1^{-1}, \ldots, \lambda_n^{-1})$, and for $\vec{k} \in I^m$, we let $\Lambda_{\vec{k}} F$ denote the coefficient of $\lambda_{\vec{k}}$ in F. To rewrite (2.4) in terms of generating functions, we multiply (2.4) by t^ℓ and sum over $\ell = 1, 2, \ldots$ to obtain

$$R_{k_1k_2} = \sum_{j=1}^{\infty} t^j C^j_{k_1\alpha_1\cdots\alpha_{j-1}k_2} \sum_{\ell=0}^{\infty} t^\ell D^\ell_{(k_1,\alpha_1,\ldots,\alpha_{j-1},k_2)}.$$

The right side is the coefficient of $\lambda_{k_1}\lambda_{k_2}$ in

$$\sum_{j=0}^{\infty} [tC(\Lambda^{-1}tC)^j]_{k_1k_2} \det\big(1 + \Lambda R(t)\big).$$

Hence, in order to establish (2.4) we have to prove that

$$R_{k_1k_2} = \Lambda_{(k_1,k_2)}\Big\{\sum_{j=0}^{\infty}[tC(\Lambda^{-1}tC)^j]_{k_1k_2} \det(1 + \Lambda R)\Big\}. \tag{2.10}$$

We take $\lambda_1, \ldots, \lambda_n$ sufficiently large so that this series converges and the following manipulations are justified. Since $\det(1 + \Lambda R)$ is linear in each λ_i, we can replace each $tC = 1 - R^{-1}$ in (2.10) by $-R^{-1}$; for a factor Λ^{-2} or $\Lambda^{-1}_{k_1i}$ or $\Lambda^{-1}_{ik_2}$ in $[\cdots]_{k_1k_2}$ cannot produce a $\lambda_{k_1}\lambda_{k_2}$ term. Hence (2.10) reads

$$\begin{aligned} R_{k_1k_2} &= \Lambda_{(k_1,k_2)}\Big\{\sum_{j=0}^{\infty}(-1)^{j+1}[R^{-1}(\Lambda^{-1}R^{-1})^j]_{k_1k_2}\det(1 + \Lambda R)\Big\} \\ &= -\Lambda_{(k_1,k_2)}\{(R + \Lambda^{-1})^{-1}_{k_1k_2} \det(1 + \Lambda R)\}. \end{aligned} \tag{2.11}$$

Similarly, the generating function version of (2.5) is

$$R_{k_1k_1} - 1 = \Lambda_{k_1}\{[1 - (R + \Lambda^{-1})^{-1}_{k_1k_1}]\det(1 + \Lambda R)\}. \tag{2.12}$$

PROOF OF THEOREM 2.1. Both (2.11) and (2.12) follow easily from Cramer's Rule. We write out the proof of (2.11), taking without loss of generality $k_1 = 1, k_2 = 2$. Let $\omega_i = \lambda_i^{-1}$ and $\Omega = \Lambda^{-1}$. Pulling out a factor of $\det\Lambda$ from $\det(1 + \Lambda R)$, we rewrite (2.11) as

$$R_{12} = -\Omega_{(3,\ldots,n)}\{(R + \Omega)^{-1}_{12} \det(R + \Omega)\} \tag{2.13}$$

where $\Omega_{(3,\ldots,n)}$ extracts the coefficient of $\omega_3 \cdots \omega_n$. By Cramer's Rule,

$$\begin{aligned} \text{R.S. of (2.13)} &= \Omega_{(3,\ldots,n)}\left\{R_{12}\det\begin{pmatrix} R_{33} + \omega_3 & R_{34} & \cdots & \\ R_{43} & R_{44} + \omega_4 & \cdots & \\ \vdots & \vdots & \ddots & R_{nn} + \omega_n \end{pmatrix}\right\} \\ &= R_{12}. \end{aligned}$$

■

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

3. **Proof.** We prove Theorem 1.3, beginning with the formula (1.8) for $(C+\lambda)^{-1}$.

LEMMA 3.1. *If A is an $n\times n$ matrix and $-\lambda$ is not an eigenvalue of A, then*

$$A^{n-1}(A+\lambda)^{-1} = d(\lambda)^{-1}Q_n(A,\lambda) \tag{3.1}$$

where $d(\lambda) = \det(A+\lambda)$,

$$Q_n(A,\lambda) = A^{n-1}\lambda^{n-1} + \sum_{i=0}^{n-2} q_{ni}(A)\lambda^i, \tag{3.2}$$

and

$$q_{ni}(A) = \sum_{l=n-2-i}^{n-2} (-1)^{n-l}s_{2n-2-i-l}A^l, \tag{3.3}$$

where s_j is the jth degree symmetric polynomial in the eigenvalues of A (see (1.9)).

PROOF. Since

$$Aq_{ni} = -q_{n,i-1} + s_{n-i}A^{n-1}$$

where $q_{n,-1} = 0$, we have

$$A\sum_{i=0}^{n-2} q_{ni}\lambda^i = -\sum_{i=0}^{n-3} q_{ni}\lambda^{i+1} + A^{n-1}\sum_{i=0}^{n-2} s_{n-i}\lambda^i. \tag{3.4}$$

Now $d(\lambda) = \sum_{i=0}^n s_{n-i}\lambda^i$ where $s_0 = 1$, and, by the Cayley-Hamilton Theorem,

$$q_{n,n-2} = \sum_{l=0}^{n-2}(-1)^{n-l}s_{n-l}A^l = -A^n + s_1A^{n-1},$$

so that (3.4) becomes

$$\begin{aligned}(A+\lambda)\sum_{i=0}^{n-2} q_{ni}\lambda^i &= q_{n,n-2}\lambda^{n-1} + A^{n-1}\big(d(\lambda) - \lambda^n - s_1\lambda^{n-1}\big)\\ &= -A^n\lambda^{n-1} - A^{n-1}\lambda^n + A^{n-1}d(\lambda)\end{aligned}$$

or

$$(A+\lambda)Q_n = A^{n-1}d(\lambda).$$

■

The proof of Theorem 1.3 has now been reduced to showing that for $0\le i\le n-2$

$$q_{ni}(C) \ge 0. \tag{3.5}$$

As we remarked in Section 1, the easy part of (3.5) is the case of the diagonal elements of q_{ni}. In fact, (3.5) holds in the operator sense, *i.e.*, $\vec{x}\cdot q_{ni}\vec{x} \ge 0$ for any $\vec{x}\in\mathbb{R}^n$:

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

LEMMA 3.2. *If $C \in \mathcal{P}_n$ and $0 \leq i \leq n-2$, then $q_{ni}(C) \geq 0$ in the operator sense and thus along the diagonal.*

PROOF. Diagonalizing (3.3), we need to show that for any eigenvalue λ_l,

$$q_{ni}(\lambda_l) = \sum_{j=n-2-i}^{n-2} (-1)^{n-j} s_{2n-2-i-j} \lambda_l^j \geq 0. \tag{3.6}$$

Let $s_{j\backslash l}$ denote the jth degree symmetric polynomial in the eigenvalues of C but with λ_l set equal to 0, and let $d_j = s_{n-i+j\backslash l}$. Note that $d_i = s_{n\backslash l} = 0$ and that $s_{n-i+j} = d_j + \lambda_l d_{j-1}$. Then

$$\begin{aligned} q_{ni}(\lambda_l) &= \sum_{k=0}^{[i/2]} s_{n-i+2k} \lambda_l^{n-2-2k} - \sum_{k=0}^{[(i-1)/2]} s_{n-i+1+2k} \lambda_l^{n-3-2k} \\ &= \sum_{k=0}^{[i/2]} (d_{2k} + \lambda_l d_{2k-1}) \lambda_l^{n-2-2k} - \sum_{k=0}^{[(i-1)/2]} (d_{2k+1} + \lambda_l d_{2k}) \lambda_l^{n-3-2k} \\ &= d_{-1} \lambda_l^{n-1} \geq 0. \end{aligned}$$

∎

We are left with the off-diagonal elements of q_{ni}. Without loss of generality we consider the 12 element and we apply the representation (2.4):

$$\big(q_{ni}(C)\big)_{12} = \sum_{\substack{l=n-2-i \\ l>0}}^{n-2} (-1)^{n-l} s_{2n-2-i-l} \sum_{j=1}^{l} C^j_{1\alpha_1\cdots\alpha_{j-1}2} D^{l-j}_{12\alpha_1\cdots\alpha_{j-1}}. \tag{3.7}$$

(Recall that Greek indices are summed from 1 to n.) Let

$$C_{jk} = R_{\beta j} \lambda_\beta R_{\beta k}$$

where R is the orthogonal matrix diagonalizing C. Inserting

$$(C^{l_i})_{k_i k_j} = R_{\beta_i k_i} \lambda_{\beta_i}^{l_i} R_{\beta_i k_j}$$

into the definition (2.1), we obtain ($k_1 = 1, k_2 = 2$)

$$\begin{aligned} D^{l-j}_{12k_3\cdots k_{j+1}} &= \sum_{l_1+\cdots+l_{j+1}=l-j} \sum_{\sigma\in S_{j+1}} \operatorname{sgn}\sigma \prod_{i=1}^{j+1} R_{\beta_i k_i} \lambda_{\beta_i}^{l_i} R_{\beta_i k_{\sigma(i)}} \\ &= a^{\beta_1\cdots\beta_{j+1}}_{12k_3\cdots k_{j+1}} p^{l-j}(\lambda_{\beta_1}, \ldots, \lambda_{\beta_{j+1}}) \end{aligned} \tag{3.8}$$

where, for $\vec{k}, \vec{b} \in I^m$

$$a^{\vec{b}}_{\vec{k}} = \sum_{\sigma\in S_m} \operatorname{sgn}\sigma \prod_{i=1}^{m} R_{b_i k_i} R_{b_i k_{\sigma(i)}}, \tag{3.9}$$

and

$$p^l(x_1, \ldots, x_m) = \sum_{\substack{l_1,\ldots,l_m=0 \\ l_1+\cdots+l_m=l}}^{l} \prod_{i=1}^{m} x_i^{l_i}$$

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

is the "generalized l-th power of $\vec{x} = (x_1, \ldots, x_m)$". We adopt the convention that $p^l = 0$ if $l < 0$.

It's easy to see that $a_{\vec{k}}^{\vec{b}} = 0$ unless $\vec{b}, \vec{k} \in I_0^m$. Moreover, since $p^{l-j}(\lambda_{\beta_1}, \ldots, \lambda_{\beta_{j+1}})$ is symmetric in the β_i's, we can symmetrize $a_{\vec{k}}^{\vec{b}}$, replacing it by

$$\begin{aligned} d_{\vec{k}}^{\vec{b}} &= \frac{1}{(j+1)!} \sum_{\tau \in S_{j+1}} a_{\vec{k}}^{b_{\tau(1)} \cdots b_{\tau(j+1)}} \\ &= \frac{1}{(j+1)!} \sum_{\tau,\sigma} \operatorname{sgn} \sigma \prod_{i=1}^{j+1} R_{b_{\tau(i)} k_i} R_{b_{\tau(i)} k_{\sigma(i)}}. \end{aligned}$$

Making the change of variables $i \longrightarrow \tau^{-1}(i)$ and $\sigma \longrightarrow \sigma \circ \tau$, we obtain

$$\begin{aligned} d_{\vec{k}}^{\vec{b}} &= \frac{1}{(j+1)!} \sum_{\tau,\sigma} \operatorname{sgn} \tau \operatorname{sgn} \sigma \prod_{i=1}^{j+1} R_{b_i k_{\tau^{-1}(i)}} R_{b_i k_{\sigma(i)}} \\ &= \frac{1}{(j+1)!} \Big[\sum_{\sigma} \operatorname{sgn} \sigma \prod_{i=1}^{j+1} R_{b_i k_{\sigma(i)}} \Big]^2 \geq 0. \end{aligned} \tag{3.10}$$

Substitution back into (3.7) gives

$$\big(q_{ni}(C)\big)_{12} = \sum_{j=1}^{n-2} C^j_{1\alpha_1 \cdots \alpha_{j-1} 2} d^{\beta_1 \cdots \beta_{j+1}}_{12\alpha_1 \cdots \alpha_{j-1}} p_{nij}(\lambda_{\beta_1}, \ldots, \lambda_{\beta_{j+1}}) \tag{3.11}$$

where

$$p_{nij}(x_1, \ldots, x_{j+1}) = \sum_{l=n-2-i}^{n-2} (-1)^{n-l} s_{2n-2-i-l} p^{l-j}(x_1, \ldots, x_{j+1}). \tag{3.12}$$

Letting $n = m + j$ and $l = k + j$ we can rewrite $p_{nij}(x_1, \ldots, x_r)$ as

$$q^j_{mi}(x_1, \ldots, x_r) = \sum_{k=m-2-i}^{m-2} (-1)^{m-k} s_{2m+j-2-i-k} p^k(x_1, \ldots, x_r). \tag{3.13}$$

Given that in (3.11) $C^j_{1\ldots} \geq 0$ and $d^{\beta_1 \cdots}_{1 \ldots} \geq 0$, the proof of (3.5) is completed by:

LEMMA 3.3. *For $0 \leq i, j \leq n-2$, $n \leq m + r - 1$, and $\vec{k} \in I_0^r$,*

$$q^j_{mi}(\lambda_{k_1}, \ldots, \lambda_{k_r}) \geq 0. \tag{3.14}$$

REMARKS. 1. We of course wish to apply the Lemma in the case $m = n - j$ and $r = j + 1$ to conclude that (3.12) ≥ 0, but we have rewritten (3.12) in the form (3.13) to facilitate the proof by induction and to bring p_{nij} into line with the form of q_{ni} in (3.3). When $j = 0$ and $r = 1$, $q^0_{mi}(x_1) = q_{mi}(x_1)$ of (3.3).

2. (3.14) systematizes the inequalities among the s_j's, like (1.13), that we require for the proof of Theorem 1.3. For example, when $n = 4$, $r = 2$, $m = 3$, $i = 1$, and $j = 0$, (3.14) reads

$$q^0_{31}(\lambda_{k_1}, \lambda_{k_2}) = s_2(\lambda_{k_1} + \lambda_{k_2}) - s_3 \geq 0. \tag{3.15}$$

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

This is basically (1.13); for $\lambda_{k_1} + \lambda_{k_2} \geq \lambda_1 + \lambda_2$ where λ_1 and λ_2 are the two smallest eigenvalues, and so (3.15) amounts to

$$(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\lambda_4 + \lambda_2\lambda_3 + \lambda_2\lambda_4 + \lambda_3\lambda_4)(\lambda_1 + \lambda_2) \\ \geq \lambda_1\lambda_2\lambda_3 + \lambda_1\lambda_2\lambda_4 + \lambda_1\lambda_3\lambda_4 + \lambda_2\lambda_3\lambda_4.$$

3. We need only prove (3.14) for $n = m + r - 1$. The case of $n < m + r - 1$ then follows by setting λ_j's equal to 0.

PROOF. We prove (3.14) for $n = m + r - 1$ by induction on $r = 1, 2, \ldots$. When $r = 1$, q^j_{mi} is almost the same as q_{mi} (see (3.6)), and the proof of (3.14) is almost identical to that of (3.6). We do not repeat it here.

So we assume that (3.14) holds for r and increase r (and n) by 1, taking $\lambda_{k_{r+1}} = \lambda_{n+1}$ without loss of generality and writing $(\lambda_{k_1}, \ldots, \lambda_{k_r}) = \vec{\lambda}$. Let s_j and p^j denote the symmetric function and generalized power for $\vec{\lambda}$. Then those for $(\vec{\lambda}, \lambda_{n+1})$ are

$$s_j + s_{j-1}\lambda_{n+1} \text{ and } \sum_{l=0}^{j} p^{j-l}\lambda^l_{n+1},$$

and we need to show that ($\lambda = \lambda_{n+1}$, $t = 2m + j - 2 - i$)

$$\sum_{k=m-2-i}^{m-2} (-1)^{m-k}(s_{t-k} + s_{t-1-k}\lambda) \sum_{l=0}^{k} p^{k-l}\lambda^l \geq 0. \tag{3.16}$$

$$\begin{aligned} \text{L.S.} &= \sum_{k=m-2-i}^{m-2} (-1)^{m-k}\Big[s_{t-k}p^k + \sum_{l=1}^{k+1}(s_{t-k}p^{k-l} + s_{t-1-k}p^{k-l+1})\lambda^l\Big] \\ &= \sum_{k=m-2-i}^{m-2} (-1)^{m-k}s_{t-k}p^k + \sum_{l=1}^{m-1} \lambda^l \sum_{k=k_0}^{m-2}(-1)^{m-k}(s_{t-k}p^{k-l} + s_{t-1-k}p^{k-l+1}) \end{aligned}$$

where $k_0 = \max(m-2-i, l-1)$. The first term is nonnegative by the inductive hypothesis. The sum over k telescopes to

$$s_{t-m+1}p^{m-1-l} + (-1)^{m-k_0}s_{t-k_0}p^{k_0-l}.$$

Since $k_0 < m - 1$, (3.16) reduces to showing that

$$s_{t-k}p^k \leq s_{t-l}p^l \text{ if } k \leq l.$$

This is obviously true since both sides consist of terms from p^t, but on the L.S. each term has at most k repetitions of a λ_j whereas on the R.S. there can be l repetitions. ∎

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

Appendix: Self-Avoiding Walks. On the basis of some observations of Greg Lawler's, we here elucidate the random walk formulas of Theorem 2.1 using the language of self-avoiding walks (SAW's).

Let $\mathcal{W}_\ell(k_1, k_2)$ be the set of ℓ-step walks on the state space $I = \{1, 2, \ldots, n\}$ with initial point k_1 and final point k_2:

$$\mathcal{W}_\ell(k_1, k_2) = \{w = (w_0, w_1, \ldots, w_\ell) \mid w_j \in I, w_0 = k_1, w_\ell = k_2\}.$$

Given an $n \times n$ transition matrix C, let

$$C(w) = C_{w_0 w_1} C_{w_1 w_2} \cdots C_{w_{\ell-1} w_\ell}.$$

With this notation

$$(A.1) \qquad (C^\ell)_{k_1 k_2} = \sum_{w \in \mathcal{W}_\ell(k_1, k_2)} C(w).$$

Let $\mathcal{S}_\ell(k_1, k_2) \subset \mathcal{W}_\ell(k_1, k_2)$ be the subset of ℓ-step self-avoiding walks, *i.e.*, $w \in \mathcal{S}_\ell$ if and only if $w_j \neq w_k$ for $j \neq k$. The set of all SAW's from k_1 to k_2 is $\bigcup_{\ell=1}^{n-1} \mathcal{S}_\ell(k_1, k_2)$. Given a $w \in \mathcal{W}_\ell(k_1, k_2)$ with $k_1 \neq k_2$, we can map w onto a SAW $\omega = E(w)$ by erasing loops, where the chronological loop erasure operator E is defined recursively as follows. If $w \in \mathcal{S}_\ell$, then $E(w) = w$. Otherwise, let i be the smallest integer for which $w_i = w_j$ for some $j > i$, and let j be the largest such integer. Set

$$w' = (w_0, w_1, \ldots, w_i, w_{j+1}, \ldots, w_\ell).$$

We iterate this procedure until we obtain a SAW $\omega \in \mathcal{S}_m(k_1, k_2)$ where $m < \ell$.

Conversely, given a SAW $\omega = (\omega_0, \ldots, \omega_m)$ and a size $\ell > m$, the walks $w \in \mathcal{W}_\ell$ such that $E(w) = \omega$ are obtained by inserting a loop after each ω_j that does not meet any of the previous ω_i's:

$$w = (\omega_0, w_0^1, \ldots, w_0^{\ell_0}, \omega_1, w_1^1, \ldots, w_1^{\ell_1}, \ldots, \omega_m, w_m^1, \ldots, w_m^{\ell_m})$$

where

$$\ell_0 \geq 0, \ldots, \ell_m \geq 0, \qquad \ell_0 + \cdots + \ell_m = \ell - m,$$
$$w_i^{\ell_i} = \omega_i, \qquad i = 0, \ldots, m,$$

and

$$w_j^k \neq \omega_i \text{ if } i < j.$$

Using this correspondence, we can rewrite the walk representation (A.1) when $k_1 \neq k_2$ as

$$(A.2) \qquad (C^\ell)_{k_1 k_2} = \sum_{m=1}^{\ell} \sum_{\omega \in \mathcal{S}_m(k_1, k_2)} C(\omega) L^{\ell-m}(\omega)$$

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

where the factor $L^{\ell-m}(\omega)$ comes from the attached loops: for $\ell > 0$,

$$(A.3)\qquad L^{\ell}(\omega) = \sum_{\substack{\ell_0,\ldots,\ell_m \\ \ell_0+\cdots+\ell_m=\ell}} C^{\ell_0}(\omega_0)C^{\ell_1}(\omega_1 \mid \omega_0)\cdots C^{\ell_m}(\omega_m \mid \omega_0,\ldots,\omega_{m-1})$$

where, for distinct $k, k_1, \ldots, k_m$ in I,

$$(A.4)\qquad C^{\ell}(k \mid k_1,\ldots,k_m) = \sum_{\substack{w\in \mathcal{W}_\ell(k,k) \\ w_j\neq k_i}} C(w).$$

For $\ell = 0$, $L^0(\omega) = C^0(k \mid k_1,\ldots,k_m) = 1$. Another way of writing (A.4) is to let $P_{\vec{k}}$ be the $n \times n$ matrix that projects onto the standard basis vectors $e_{k_1},\ldots,e_{k_m}$, and let $Q_{\vec{k}} = 1 - P_{\vec{k}}$; then

$$(A.5)\qquad C^{\ell}(k \mid k_1,\ldots,k_m) = [(Q_{\vec{k}}CQ_{\vec{k}})^{\ell}]_{kk}.$$

A comparison of (A.2) with (2.4) reveals that the determinantal construct (2.1) must be the same as the loop contribution (A.3). This identification is the main point of this Appendix:

THEOREM A.1. *For $\omega \in \mathcal{S}_m$ or, equivalently, $\vec{\omega} \in I_0^{m+1}$ (see (2.2)),*

$$(A.6)\qquad D^{\ell}_{\vec{\omega}} = L^{\ell}(\omega).$$

To prove (A.6) we shall establish that both sides satisfy the same recursion relation. Given a SAW $\omega = (\omega_0,\ldots,\omega_{m-1})$, let $\omega' = (\omega_0,\ldots,\omega_{m-1},\omega_m)$ be a SAW with one more step. From the definition (A.3) it is obvious that

$$(A.7)\qquad L^{\ell}(\omega') = \sum_{j=0}^{\ell} C^{j}(\omega_m \mid \omega)L^{\ell-j}(\omega)$$

where, for $m = 1$,

$$(A.8)\qquad L^{\ell}(\omega_0) = C^{\ell}(\omega_0) = (C^{\ell})_{\omega_0\omega_0}.$$

By (A.10) below, $D^{\ell}_{\vec{\omega}}$ satisfies the same recursion. Since it satisfies the same initial condition, namely $D^{\ell}_{\omega_0} = (C^{\ell})_{\omega_0\omega_0}$, the desired identity (A.6) follows by induction on the size of ω.

Let $\vec{k} \in I_0^m$ and $\vec{k}' \in I_0^{m'}$ so that $(\vec{k},\vec{k}') \in I_0^{m+m'}$. As in (A.5), we let $P' = P_{\vec{k}'}$, $Q' = 1-P'$, $C' = Q'CQ'$, $R' = (1-tC')^{-1}$, and $D^{\ell}_{\vec{k}|\vec{k}'}$ be given by (2.1) with C replaced by C'.

LEMMA A.2. *For $(\vec{k},\vec{k}') \in I_0^{m+m'}$,*

$$(A.9)\qquad D^{\ell}_{(\vec{k},\vec{k}')} = \sum_{j=0}^{\ell} D^{j}_{\vec{k}|\vec{k}'}D^{\ell-j}_{\vec{k}'}.$$

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

REMARK. When $m = 1$, in which case $D^{j}_{\vec{k}|\vec{k}'} = (C')^{j}_{k_1 k_1}$, the reduction formula (A.9) becomes

$$(A.10) \qquad D^{\ell}_{(\vec{k}',k_1)} = \sum_{j=0}^{\ell} (C')^{j}_{k_1 k_1} D^{\ell-j}_{\vec{k}'}$$

which is the same recursion relation as (A.7).

PROOF. We base our proof of (A.9) on the generating function (2.9). By the resolvent identity and the fact that Q' commutes with R',

$$\begin{aligned} Q'(R-R') &= tQ'R'(C-C')R \\ &= tR'Q'(P'C + CP' + P'CP')R \\ &= tQ'R'CP'R. \end{aligned}$$

Hence

$$R = P'R + Q'R = P'R + Q'R' + tQ'R'CP'R.$$

We wish to extract the $\lambda_{\vec{k}'} = \prod_i \lambda_{k'_i}$ term from

$$\det(1+\Lambda R) = \det(R^{-1} + \Lambda P' + \Lambda Q'R'R^{-1} + t\Lambda Q'R'CP') \det R.$$

In each of the nonzero columns of P' we clearly have to choose the element $\lambda_{k'_i}$ of $\Lambda P'$, and so the matrix $t\Lambda Q'R'CP'$ doesn't contribute. Therefore,

$$\begin{aligned} \partial_{\lambda_{\vec{k}'}} \det(1+\Lambda R) &= \partial_{\lambda_{\vec{k}'}} \det(1 + \Lambda P'R + \Lambda Q'R') \\ &= \partial_{\lambda_{\vec{k}'}} \det[(1+\Lambda Q'R')(1+\Lambda P'R)] \\ &= \det(1+\Lambda Q'R')\partial_{\lambda_{\vec{k}'}} \det(1+\Lambda P'R) \end{aligned}$$

since $(\Lambda Q'R')(\Lambda P'R) = 0$. Taking the derivatives $\partial_{\lambda_{\vec{k}}}$ and ∂_t^{ℓ}, we obtain (A.9). ∎

The second identity in Theorem 2.1 has a similar interpretation in terms of a walk being decomposed into a SAW with attached loops. The only difference is that for the contributions to $(C^{\ell})_{k_1 k_1}$ each SAW is actually a SAL, a self-avoiding loop that avoids itself except for its equal initial and final points. The loop erasure mapping from $\mathcal{W}_{\ell}(k_1, k_1)$ to $\mathcal{S}_m(k_1, k_1)$ is constructed as before except that when a walk w has more than one loop at its initial point w_0 we do not erase the *last* loop at ω_0.

REFERENCES

1. M. Déchamps-Gondim, F. Lust-Piquard and H. Queffelec, *On the Minorant Properties in $C_p(H)$*, Pac. J. Math. **119**(1985), 89–101.

2. V. V. Peller, Dokl. Acad. Nauk. Math. **252**(1980), 43–47.

3. B. Simon, *Pointwise Domination of Matrices and Comparison of I_p Norms*, Pac. J. Math. **97**(1981), 471–475.

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press

4. S. Weissenhofer, University of British Columbia Ph.D. Thesis, 1993.
5. K. Yosida, *Functional Analysis*, Springer-Verlag, New York, 1968.

Mathematics Department
University of British Columbia
Vancouver, B.C.
V6T 1Z2
e-mail: rosen@math.ubc.ca

https://doi.org/10.4153/CJM-1996-009-9 Published online by Cambridge University Press