Determinants from Homomorphisms

Radu Curticapean
IT University of Copenhagen, Denmark
Basic Algorithms Research Copenhagen, Denmark

Abstract

We give a new combinatorial explanation for well-known relations between determinants and traces of matrix powers. Such relations can be used to obtain polynomial-time and poly-logarithmic space algorithms for the determinant. Our new explanation avoids linear-algebraic arguments and instead exploits a classical connection between subgraph and homomorphism counts.

2012 ACM Subject Classification Computing methodologies → Linear algebra algorithms; Mathematics of computing → Combinatorics; Mathematics of computing → Graph theory; Theory of computation → Parallel algorithms

Keywords and phrases determinant, homomorphisms, matrix trace, Newton identities

Digital Object Identifier 10.4230/LIPIcs.ESA.2022.38

Related Version *arXiv version*: https://arxiv.org/abs/2204.10718

Funding BARC is supported by Villum Foundation grant 16582.

1 Introduction

The determinant of $n \times n$ matrices is, up to scaling, the unique function from $n \times n$ matrices to scalars that is linear and alternating in the rows and columns. It admits the well-known *Leibniz formula*

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i,\pi(i)}, \tag{1}$$

where S_n denotes the set of permutations of $\{1, \dots, n\}$ and $\operatorname{sgn} : S_n \to \{-1, 1\}$ denotes the permutation sign. Writing $\sigma(\pi)$ for the number of cycles in π, the permutation sign can be expressed as $\operatorname{sgn}(\pi) = (-1)^{n+\sigma(\pi)}$.

When presented with only the right-hand side of (1), unaware of the connection to the determinant, one would likely struggle to evaluate this sum of $n!$ terms efficiently. For comparison, it is #P-hard to compute the similarly defined *permanent* [8], which is obtained by omitting the sign factors from the expression.

Yet, determinants can be evaluated efficiently, e.g., via Gaussian elimination in $O(n^3)$ field operations, including divisions. Asymptotically optimal algorithms achieve $O(n^\omega)$ operations, where $\omega < 3$ is the exponent of matrix multiplication [2, Exercise 28.2-3]. Note that $\det(A)$ is defined over any ring containing the entries of A; there are also algorithms computing determinants with a polynomial number of ring operations, i.e., excluding divisions [6, 1, 7].

Determinants from matrix powers

It is classically known in linear algebra that $\det(A)$ for $n \times n$ matrices A can be computed from the matrix traces $\operatorname{tr}(A^k)$ for $1 \le k \le n$. In the following, assume that A is defined over an algebraically closed field $\mathbb{F}$, such that A has eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{F}$. The idea is to express $\det(A)$ and $\operatorname{tr}(A^k)$ for $1 \le k \le n$ as particular polynomials in the eigenvalues $\lambda_1, \dots, \lambda_n$ and then relate these polynomials.

30th Annual European Symposium on Algorithms (ESA 2022).
Editors: Shiri Chechik, Gonzalo Navarro, Eva Rotenberg, and Grzegorz Herman; Article No. 38; pp. 38:1–38:7
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

- The determinant can be expressed as $\det(A) = \lambda_1 \dots \lambda_n$; this is the n-th elementary symmetric polynomial in the eigenvalues. Generally, the k-th elementary symmetric polynomial $e_k(x_1, \dots, x_n)$ in n variables is the sum of monomials $\sum_S \prod_{i \in S} x_i$, where S ranges over all k-subsets $S \subseteq \{1, \dots, n\}$.
- The matrix trace satisfies $\mathrm{tr}(A) = \lambda_1 + \dots + \lambda_n$, and more generally, $\mathrm{tr}(A^k) = \lambda_1^k + \dots + \lambda_n^k$. This is the k-th power-sum polynomial in the eigenvalues. Generally, the k-th power sum polynomial is $p_k(x_1, \dots, x_n) = x_1^k + \dots + x_n^k$.

The Girard–Newton identities then allow us to relate the power-sum and elementary symmetric polynomials. They state that, for all $1 \leq k \leq n$,

$$k e_k(x_1, \dots, x_n) = \sum_{i=1}^{k} (-1)^{i-1} e_{k-i}(x_1, \dots, x_n) p_i(x_1, \dots, x_n).$$

For $\mathbb{F}$ of characteristic 0, a recursive application of these identities allows us to compute $\det(A) = e_n(\lambda_1, \dots, \lambda_n)$ from the values $p_k(\lambda_1, \dots, \lambda_n) = \mathrm{tr}(A^k)$ for $1 \leq k \leq n$. Csanky's algorithm [4, Chapter 31] implements this approach with an arithmetic circuit of bounded fan-in, $O(\log^2 n)$ depth, and polynomial size. In other words, it shows that determinants can be computed with $O(\log^2 n)$ operations on a polynomial number of parallel processors.

Our result

The main result of this paper is a novel and self-contained derivation of a known and algorithmically useful formula that expresses the determinant of $n \times n$ matrices A as a polynomial combination of traces of matrix powers:

$$\det(A) = (-1)^n \sum_{\lambda \vdash n} (-1)^{|\lambda|} \prod_{\ell=1}^{n} \frac{\mathrm{tr}(A^\ell)^{s_\ell(\lambda)}}{s_\ell(\lambda)! \cdot \ell^{s_\ell(\lambda)}}. \tag{2}$$

Some remarks on the notation are in order. The sum ranges over partitions λ of n, which are multi-sets of positive numbers summing to n. We write $\lambda \vdash n$ to indicate that λ is a partition of n and write $|\lambda|$ for its number of parts. For $\ell \in \mathbb{N}$, we write $s_\ell(\lambda) \in \mathbb{N}$ for the number of occurrences of ℓ in λ.

The previous subsection essentially gives a proof of (2) by appropriately expanding the recursive applications of the Girard–Newton identities. The new proof we present in this paper bypasses notions like eigenvalues, symmetric polynomials, and the Girard–Newton identities, and instead relies on ideas from the theory of graph homomorphism counts.

Our proof of (2) is contained in Section 2. We then sketch in Section 3 how this formula can be used to obtain polynomial-time and parallel algorithms for the determinant.

2 Proof of Equation (2)

In the following, let $[n] = \{1, \dots, n\}$ and let $A = (a_{i,j})_{i,j \in [n]}$ be a matrix. We will study the determinant of A using graph-theoretic language. The graphs G we consider are *directed* and may feature *self-loops*, and some graphs may feature *parallel edges* between the same pair of vertices. We write $V(G)$ and $E(G)$ for the vertices and edges of G.

2.1 Determinants are sums of cycle covers

The matrix A induces an edge-weighted complete directed graph with self-loops on the vertex set $V(A) = [n]$. Abusing notation, we also write A for this weighted graph. In this view, permutations correspond to *cycle covers*, which are edge-sets $C \subseteq E(A)$ inducing vertex-disjoint cycles that cover all vertices of A.

We require the more general notion of a *k-partial cycle cover* for $0 \le k \le n$, which is a collection of vertex-disjoint cycles with k edges in total. We write $\sigma(C)$ for the number of cycles in C, define the sign of C analogously to the permutation sign as $\mathrm{sgn}(C) = (-1)^{|C|+\sigma(C)}$, and define the *format* of C as the partition $\lambda \vdash k$ induced by the multi-set of cycle lengths. Finally, let $\mathcal{C}(n,k)$ be the set of k-partial cycle covers of the complete directed graph on vertex set $[n]$.

Partial cycle covers are connected to *k-partial determinants*. Up to sign, these are the coefficients of characteristic polynomials, and they can be defined as

$$\det_k(A) = \sum_{S \subseteq [n] \text{ of size } k} \det(A[S]),$$

where $A[S]$ is the square sub-matrix of A defined by restricting to the rows and columns contained in S. From the Leibniz formula (1), it follows that

$$\det_k(A) = \sum_{C \in \mathcal{C}(n,k)} \mathrm{sgn}(C) \prod_{uv \in C} a_{u,v}. \tag{3}$$

Given $\lambda \vdash k$, let $C_\lambda \in \mathcal{C}(n,k)$ be any fixed cycle cover of format λ. We can regroup terms in (3) to obtain

$$\det_k(A) = \sum_{\lambda \vdash k} \mathrm{sgn}(C_\lambda) \underbrace{\sum_{\substack{C \in \mathcal{C}(n,k) \text{ of} \\ \text{format } \lambda}} \prod_{uv \in C} a_{u,v}}_{=:\mathrm{sub}(C_\lambda \to A)}. \tag{4}$$

Note that the quantity $\mathrm{sub}(C_\lambda \to A)$ defined above is a weighted sum over the cycle covers C isomorphic to C_λ, weighted by the product of the edge-weights in C.

2.2 Relating subgraphs, embeddings and homomorphisms

Let L be a graph, possibly containing parallel edges. The weighted *homomorphism* and *embedding* counts from L into A are defined as

$$\mathrm{hom}(L \to A) = \sum_{f : V(L) \to [n]} \prod_{\substack{e \in E(L) \\ \text{with } e = uv}} a_{f(u),f(v)}, \tag{5}$$

$$\mathrm{emb}(L \to A) = \sum_{\substack{f : V(L) \to [n] \\ \text{injective}}} \prod_{\substack{e \in E(L) \\ \text{with } e = uv}} a_{f(u),f(v)}. \tag{6}$$

For example, if C_ℓ denotes the directed ℓ-cycle for $\ell \in \mathbb{N}$, then $\mathrm{hom}(C_\ell \to A) = \mathrm{tr}(A^\ell)$. Moreover, for $\lambda \vdash k$, recall that $s_\ell(\lambda)$ counts the occurrences of part ℓ in λ. We have

$$\mathrm{hom}(C_\lambda \to A) = \prod_{\ell=1}^{k} \mathrm{tr}(A^\ell)^{s_\ell(\lambda)}, \tag{7}$$

since homomorphisms from a disjoint union of graphs can be chosen independently for the individual components; this implies that $\mathrm{hom}(C_\lambda \to A)$ is the product of homomorphism counts for the individual cycles in C_λ.

Given a graph P without parallel edges, an *automorphism* of P is an isomorphism into itself. We write $\mathrm{aut}(P)$ for the number of automorphisms of P. For example,

$$\mathrm{aut}(C_\lambda) = \prod_{\ell=1}^{k} s_\ell(\lambda)! \cdot \ell^{s_\ell(\lambda)}, \tag{8}$$

since any automorphism of C_λ (i) permutes the set of $s_\ell(\lambda)$ cycles for any fixed length ℓ, which gives rise to a factor of $s_\ell(\lambda)!$ in the above product, and (ii) independently applies an automorphism to each cycle, giving rise to a factor of ℓ for every cycle of length ℓ.

With these notions set up, we can successively express subgraph counts from cycle covers, as defined in (4), via homomorphism counts. First, we transition from subgraph to embedding counts: As every subgraph isomorphic to C_λ gives rise to $\mathrm{aut}(C_\lambda)$ many embeddings with the same image, we obtain

$$\mathrm{sub}(C_\lambda \to A) = \frac{\mathrm{emb}(C_\lambda \to A)}{\mathrm{aut}(C_\lambda)}. \tag{9}$$

Next, we transition from embedding to homomorphism counts. Roughly speaking, embedding counts from a graph H are equal to homomorphism counts from H plus "lower-order terms" involving only homomorphism counts from graphs F with strictly less vertices than H. This follows directly from [5, (5.18)] and we include a simple proof for completeness, also contained in [3].

▶ **Lemma 1.** *For any fixed graph H, there are coefficients $\beta_F \in \mathbb{Z}$ for all graphs F with $|V(F)| < |V(H)|$ such that*

$$\mathrm{emb}(H \to A) = \mathrm{hom}(H \to A) + \sum_{\substack{\text{graphs } F \text{ with} \\ |V(F)|<|V(H)|}} \beta_F \, \mathrm{hom}(F \to A).$$

Proof. Given a partition ρ of the set[1] $V(H)$, the *quotient* H/ρ is the multigraph obtained by identifying the vertices within each block of ρ while keeping all possibly emerging self-loops and multi-edges. We have

$$\mathrm{hom}(H \to A) = \sum_{\substack{\text{partition } \rho \\ \text{of } V(H)}} \mathrm{emb}(H/\rho \to A), \tag{10}$$

since any homomorphism $f : V(H) \to [n]$ induces a partition $\rho = \{f^{-1}(i) \mid i \in [n]\}$ by which f may be viewed as an embedding from H/ρ to A.

Write $\perp$ for the finest partition of the set $V(H)$, that is, the partition consisting of $|V(H)|$ singleton parts. By rearranging (10) and using $H/\perp = H$, we obtain

$$\mathrm{emb}(H \to A) = \mathrm{hom}(H \to A) - \sum_{\substack{\text{partition } \rho \neq \perp \\ \text{of } V(H)}} \mathrm{emb}(H/\rho \to A). \tag{11}$$

Note that all graphs H/ρ with $\rho \neq \perp$ have strictly less vertices than H. We can therefore apply (11) again to express each term $\mathrm{emb}(H/\rho \to A)$ on the right-hand side as $\mathrm{hom}(H/\rho \to A)$ minus embedding counts from smaller graphs. This process can be iterated until reaching single-vertex graphs, from which homomorphism and embedding counts coincide trivially. Upon termination of this process, all occurrences of embedding counts have been replaced by homomorphism counts. ◀

Combining (4), (9), and Lemma 1, it follows that the k-partial determinant is a linear combination of homomorphism counts from k-partial cycle covers plus "lower-order terms".

[1] Note that ρ here is a partition of a *set*, not a partition of a *number*.

▶ **Corollary 2.** *For any fixed $k \in \mathbb{N}$, there are coefficients $\alpha_F \in \mathbb{Q}$ for all graphs F with $|V(F)| < k$ such that, for any $n \times n$ matrix A,*

$$\det_k(A) = \left(\sum_{\lambda \vdash k} \frac{\mathrm{sgn}(C_\lambda)}{\mathrm{aut}(C_\lambda)} \hom(C_\lambda \to A) \right) + \sum_{\substack{F \text{ with} \\ |V(F)|<k}} \alpha_F \hom(F \to A). \tag{12}$$

2.3 Lower-order terms vanish

As it turns out, the "lower-order terms" in (12) vanish. To show this, we use Kronecker products to lift this equality to a polynomial identity and then compare coefficients. For $t \in \mathbb{N}$, the *Kronecker product*[2] $A \otimes J_t$ of A with the $t \times t$ all-ones matrix J_t is an $nt \times nt$ matrix with row and column indices from $[n] \times [t]$, such that the entry at row (i, r) and column (j, r') equals $a_{i,j}$. In other words, each entry $a_{i,j}$ of A is replaced in $A \otimes J_t$ by a $t \times t$ matrix that contains only $a_{i,j}$.

It turns out that (12) "reacts polynomially" to this operation: When fixing $k = n$ and replacing A by $A \otimes J_t$ for varying t in (12), the homomorphism counts $\hom(S \to A \otimes J_t)$ for graphs S on the right-hand side become polynomials in t. In fact, each such homomorphism count contains only a single monomial:

▷ Claim 3. For any graph S, we have $\hom(S \to A \otimes J_t) = t^{|V(S)|} \hom(S \to A)$.

Proof. Every function $f : V(S) \to [n]$ induces $t^{|V(S)|}$ functions $f' : V(S) \to [n] \times [t]$, all of the same edge-weight product, by choosing an index $r_v \in [t]$ for each vertex $v \in V(S)$. Conversely, every such function is induced by the function that forgets the second component of the images. ◁

Applying Claim 3 on (12) with $k = n$, and with $A \otimes J_t$ instead of A, we obtain

$$\det_n(A \otimes J_t) = t^n \left(\sum_{\lambda \vdash n} \frac{\mathrm{sgn}(C_\lambda)}{\mathrm{aut}(C_\lambda)} \hom(C_\lambda \to A) \right) + \sum_{\substack{F \text{ with} \\ |V(F)|<n}} t^{|V(F)|} \alpha_F \hom(F \to A). \tag{13}$$

We now observe that $\det_n(A \otimes J_t)$ is proportional to t^n. This will allow us to ignore the lower-order graphs F in (13).

▷ Claim 4. We have $\det_n(A \otimes J_t) = t^n \det(A)$.

Proof. By definition of the partial determinant, we have

$$\det_n(A \otimes J_t) = \sum_{\substack{S \subseteq [n]\times[t] \\ \text{with } |S|=n}} \det((A \otimes J_t)[S]).$$

If S contains two pairs that agree in the first component, then the $n \times n$ matrix $(A \otimes J_t)[S]$ contains two equal columns and its determinant vanishes. We can therefore restrict the summation to index sets of the form $S = \{(1, r_1), \ldots, (n, r_n)\}$ for $r_1, \ldots, r_n \in [t]$. There are t^n sets S of this form, each with $(A \otimes J_t)[S] = A$. The claim follows. ◁

[2] The Kronecker product $A \otimes B$ can be defined for general A and B, but we only require $B = J_t$.

Now consider the polynomial identity (13) again. By Claim 4, the left-hand side equals $t^n\det(A)$. Then comparing the coefficients of t^n on both sides yields

$$\det(A) = \sum_{\lambda\vdash n} \frac{\mathrm{sgn}(C_\lambda)}{\mathrm{aut}(C_\lambda)} \hom(C_\lambda \to A) \tag{14}$$

$$= \sum_{\lambda\vdash n} (-1)^{n+|\lambda|} \prod_{\ell=1}^{n} \frac{\mathrm{tr}(A^\ell)^{s_\ell(\lambda)}}{s_\ell(\lambda)! \cdot \ell^{s_\ell(\lambda)}}. \tag{15}$$

For the last equation, we expanded $\hom(C_\lambda \to A)$ via (7) and $\mathrm{aut}(C_\lambda)$ via (8), and we used the definition of $\mathrm{sgn}(C_\lambda)$. This proves (2).

▶ Remark. The above argument applies more generally. Consider any function $F : \mathbb{Q}^{k\times k} \to \mathbb{Q}$ that admits a set $\mathcal{H}$ of k-vertex graphs and coefficients $\alpha_H \in \mathbb{Q}$ for $H \in \mathcal{H}$ such that

$$F(A) = \sum_{H\in\mathcal{H}} \alpha_H \,\mathrm{emb}(H \to A).$$

By Lemma 1, every $\mathrm{emb}(H \to A)$ is a sum of $\hom(H \to A)$ and lower-order homomorphism counts; this yields an analogue of (12). If F vanishes on $k \times k$ matrices with two identical rows/columns, then an analogue of Claim 4 holds, and it follows as above that the lower-order homomorphism counts vanish.

3 Algorithmic applications

Equation (2) does not directly imply a polynomial-time algorithm for the determinant, as the sum over partitions $\lambda \vdash n$ involves a super-polynomial number of terms. Nevertheless, this sum can be computed in polynomial time via dynamic programming or polynomial multiplication, as shown below.

▶ **Lemma 5.** *Given* $\mathrm{tr}(A^\ell)$ *for all* $1 \le \ell \le n$*, we can compute* $\det(A)$ *with* $O(n^3)$ *operations.*

Proof. Let X be a formal indeterminate. For $1 \le \ell \le n$, define the polynomial

$$p_\ell(X) = \sum_{i=0}^{\lfloor n/\ell \rfloor} (-1)^i \frac{\mathrm{tr}(A^\ell)^i}{i! \cdot \ell^i} X^{\ell i}.$$

We observe first that the coefficient of X^n in the product $(-1)^n p_1 \dots p_n$ is the desired sum in (2), and then focus on computing that coefficient.

For the first part, note that the coefficient of X^n can be viewed as a weighted count of all ways to choose a power X^{i_ℓ} from each polynomial p_ℓ, subject to $\sum_\ell \ell \cdot i_\ell = n$. These choices yield a partition $(1^{i_1}, \dots n^{i_n}) \vdash n$ that is weighted by $\prod_{\ell=1}^n (-1)^{i_\ell} \frac{\mathrm{tr}(A^\ell)^{i_\ell}}{i_\ell! \cdot \ell^{i_\ell}}$. Thus, the coefficient of X^n in $(-1)^n p_1 \dots p_n$ can be viewed as a sum over partitions $\lambda \vdash n$ whose terms correspond to those in (2).

We compute the first $n+1$ coefficients of $(-1)^n p_1 \dots p_n$, including the coefficient of X^n, by iteratively multiplying p_t onto $p_1 \dots p_{t-1}$ and truncating the intermediate result to the first $n+1$ coefficients. Using standard polynomial multiplication, each of the n iterations takes $O(n^2)$ operations. Overall, this procedure requires $O(n^3)$ operations. ◀

By naively iterating matrix multiplication, we can compute $\mathrm{tr}(A^\ell)$ for all $1 \le \ell \le n$ with $O(n^{\omega+1})$ overall operations, where ω is the exponent of matrix multiplication. This implies:

▶ **Theorem 6.** *The determinant* $\det(A)$ *can be computed with* $O(n^{\omega+1})$ *operations.*

The traces and subsequent application of (2) can also be computed with arithmetic circuits of constant fan-in and poly-logarithmic depth: Any product of two matrices can be computed trivially in $O(\log n)$ depth and $O(n^3)$ size. Repeated squaring allows us to compute all matrix powers A^ℓ and their traces $\mathrm{tr}(A^\ell)$ for $1 \le \ell \le n$ in $O(\log^2 n)$ depth and $\tilde{O}(n^4)$ overall size.

After all traces are computed, the polynomial multiplications from the proof of Lemma 5 can be performed in $O(\log^2 n)$ depth and $\tilde{O}(n^3)$ size: A single polynomial multiplication can be computed in $O(\log n)$ depth and $\tilde{O}(n^2)$ size. The truncation of the product $(-1)^n p_1 \ldots p_n$ to the first $n+1$ coefficients can then be computed as an $O(\log n)$-depth binary tree, with $p_1, \ldots, p_n$ at the leaves, and each internal node performing a polynomial multiplication followed by truncating to the first $n+1$ coefficients. This implies:

▶ **Theorem 7.** *The determinant* $\det(A)$ *can be computed with an arithmetic circuit of constant fan-in,* $O(\log^2 n)$ *depth, and* $\tilde{O}(n^4)$ *size.*

References

1 Stuart J. Berkowitz. On computing the determinant in small parallel time using a small number of processors. *Inf. Process. Lett.*, 18(3):147–150, 1984. doi:10.1016/0020-0190(84)90018-8.

2 Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, 3rd Edition*. MIT Press, 2009. URL: http://mitpress.mit.edu/books/introduction-algorithms.

3 Radu Curticapean, Holger Dell, and Dániel Marx. Homomorphisms are a good basis for counting small subgraphs. In Hamed Hatami, Pierre McKenzie, and Valerie King, editors, *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017.*, pages 210–223. ACM, 2017. doi:10.1145/3055399.3055502.

4 Dexter C. Kozen. *Csanky's Algorithm*, pages 166–170. Springer New York, New York, NY, 1992. doi:10.1007/978-1-4612-4400-4_31.

5 László Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012. URL: http://www.ams.org/bookstore-getitem/item=COLL-60.

6 Meena Mahajan and V. Vinay. A combinatorial algorithm for the determinant. In Michael E. Saks, editor, *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, 5-7 January 1997, New Orleans, Louisiana, USA*, pages 730–738. ACM/SIAM, 1997. URL: http://dl.acm.org/citation.cfm?id=314161.314429.

7 P. A. Samuelson. A method of determining explicitly the coefficients of the characteristic equation. *The Annals of Mathematical Statistics*, 13(4):424–429, 1942. URL: http://www.jstor.org/stable/2235845.

8 Leslie G. Valiant. The complexity of computing the permanent. *Theor. Comput. Sci.*, 8:189–201, 1979. doi:10.1016/0304-3975(79)90044-6.