

DOI: 10.5281/zenodo.23435404

EXACT PRECEDENCE-CONSTRAINED SCHEDULING SHOWS THAT EUCLID'S ORDER OF PROPOSITIONS IS ECONOMICAL BUT CLOSER TO CHANCE THAN TO THE OPTIMUM IN TOTAL CITATION DISTANCE

Luis Eduardo Muñoz Guerrero^{1*}

¹Universidad Tecnológica de Pereira, Pereira, Colombia; Email: lemuno zg@utp.edu.co; Orcid ID: 0000-0002-9414-6187

Received: 07/09/2023
Accepted: 08/02/2024

Corresponding Author: Luis Eduardo Muñoz Guerrero
(lemuno zg@utp.edu.co)

ABSTRACT

This study measures how close the order of the propositions in Euclid's Elements comes to minimising the total distance between each proposition and those its proof uses. The 465 propositions and 1,676 dependencies marked in Heath's translation define, book by book, a single-machine scheduling problem with precedence constraints, which is NP-hard. Dynamic programming over order ideals and Sidney's decomposition give exactly the expected length under uniformly random admissible orders and the minimum; an optimality index maps them to zero and one. The analysis plan was frozen beforehand. The text is shorter than all 100,000 sampled admissible orders ($p = 1.0 \times 10^{-5}$), but its index is 0.364, below the pre-registered threshold of 0.5 under every reading of the citations (0.31–0.37), although an objective restricted to dependencies inside each scheduling problem raises it to 0.49–0.54. Placebo citations lower it to 0.069. Spinoza's Ethics, the pre-registered replication, reaches 0.353. The largest departures involve widely cited results, such as the triangle's angle sum, stated with their subject, not near their uses. The order is economical, yet it recovers just over a third of the reduction the logic allows, so minimising total citation distance does not alone account for it.

KEYWORDS: The Order Is Economical, Yet It Recovers Just Over a Third of the Reduction the Logic Allows, So Minimising Total Citation Distance Does Not Alone Account for It.

1. INTRODUCTION

The objective of this study is to measure how economical the order of Euclid's *Elements* is, in a precise and testable sense. Each proof in the *Elements* uses propositions proved earlier, so every order of presentation places each proposition at some distance, counted in propositions, from the results it uses. The study asks how close the order that has come down to us comes to the order that minimises the sum of these distances, when only orders that respect the logical dependencies are allowed. It has four specific aims: to formulate the measurement as an optimisation problem and establish its complexity; to compute exactly, for the text as it stands, both the minimum and the expected value under random orders that respect the dependencies; to test, under an analysis plan fixed before the cost of the text's order was computed, whether the order is closer to the optimum than to chance; and to locate the propositions and books in which the text departs most from the optimum.

The arrangement of the *Elements* has been debated for centuries. In the sixteenth century Ramus criticised Euclid's arrangement of the propositions from the point of view of logic, whereas Heath, discussing the placement of I.27 and I.28, defended an arrangement justified by precision of reference (Heath, 1908). Dodgson observed that many propositions have a partly arbitrary place, since a proposition that is not needed until much later could occupy any intermediate position without risk of circular argument, so that a single list excludes many arrangements compatible with the logical sequence (Dodgson, 1879). Mueller argued that I.45 shapes the content and order of Book I, as Schiefsky reports in a study that also draws graphs of deductive relationships for Books I-IV (Schiefsky, 2007). Neuenschwander compiled a census of the propositions used in every proof of Books I-IV (Neuenschwander, 1973) and later published the same analysis for the solid books (Neuenschwander, 1974). Explicit references are themselves an editorial layer: explicit references are an innovation of modern editors, who disagree (Boman et al., 2014).

Computational studies have treated the *Elements* as a network. Wolfram drew a network of the theorems of the *Elements*, each connected to the theorems used in its proof (Wolfram, 2002), and later observed that the dependency graph defines a partial order and that even for Euclid there are things to learn about the layout of theorems and what should count as close to what (Wolfram, 2021). Atkinson examined the dependencies of the Pythagorean theorem and suggested moving I.45 to Book II, where

it is first needed (Atkinson, 2016); Viteri and DeDeo analysed networks of proofs, including Euclid (Viteri and DeDeo, 2022); and Masucci and co-authors studied the deductive network of *Principia Mathematica* (Masucci, 2011). When Beeson and co-authors checked Book I by computer, a fully formal proof required a different order for the early propositions, after which Euclid's order could be followed more closely (Beeson et al., 2019). None of these studies compares the order of a text with the best order its dependencies allow.

The closest precedent comes from formal mathematics, where dependencies extracted from formal libraries are available (Alama et al., 2012): for a Mizar proof, Pał computed the linearization with the smallest sum of distances between premises and their uses (Pał, 2014) and showed that reordering proof steps to bring premises close to their uses is NP-hard (Pał, 2015).

A related question has been studied at scale in linguistics, where words linked by a syntactic dependency tend to be placed close together. Ferrer i Cancho found that the distance between linked words is significantly small (Ferrer i Cancho, 2004); dependency lengths are shorter than random baselines in 37 languages (Futrell et al., 2015); in English, dependency length is much closer to optimal than to random (Gildea and Temperley, 2010); and the evidence supports dependency length minimization as a factor in language structure (Temperley and Gildea, 2018). The optimality score Ω places an observed length on a scale on which the random baseline scores zero and the minimum scores one (Ferrer-i-Cancho et al., 2022). For a deductive text the relevant baseline is different, because an order of presentation must respect the dependencies: the admissible orders are the linear extensions of the dependency order, and a linear extension is a permutation that respects the partial order (Huber, 2006). Random linear extensions already serve as a reference in rank tests that rely on sampling linear extensions to compute p-values (Mondal and Hinrichs, 2016) and in Bayesian analyses of historical lists modelled as random linear extensions (Nicholls et al., 2022), and in learning theory learning sequences are linear extensions of the prerequisite order (Eppstein, 2008).

To our knowledge, the order of exposition of a historical deductive treatise has not been measured against the exact optimum of the corresponding ordering problem, nor against the exact expectation under random orders that respect its dependencies. This study makes both measurements for the *Elements*. Its contributions are four. First, the total

dependency length over admissible orders separates by blocks into single-machine scheduling problems with precedence constraints, which are NP-hard in general and for which approximation guarantees are too weak to measure economy. Second, the expected length under uniformly random admissible orders and the minimum are computed exactly for every block, by dynamic programming over order ideals combined with a decomposition of the scheduling problem. Third, the order of the text is tested against both reference points under a plan frozen before its cost was computed, with validity checks, robustness analyses and an independent re-implementation. Fourth, an exploratory analysis identifies the propositions whose position costs the most. The answer is mixed: the order is far more economical than random admissible orders, but it recovers just over a third of the reduction that the dependencies allow, below the pre-registered threshold, and its economy is concentrated in a few books. When only the dependencies inside each book, or each group of Book X, are counted, it recovers about half, between 0.49 and 0.54.

2. METHODS

Data. The text is Heath's English translation of the *Elements*, in which Heath's bracketed annotations refer back to prior propositions (Avigad et al., 2009). The Perseus Digital Library encodes each annotation as a pointer to its target, so that Heath's bracketed citations are machine-readable (Perseus Digital Library, 2023); the file, in the format of the Text Encoding Initiative (TEI), is used at a fixed commit of its public repository and verified by checksum at every run. A proposition v depends on a proposition u when the text of v , outside Heath's notes, points to u ; a pointer to a porism or a lemma counts as a pointer to the proposition that contains it, and

repeated pointers count once. The file contains 465 propositions in thirteen books and 1,677 distinct dependencies between propositions, 913 of them within a book. Exactly one dependency points forward: the proof of X.10 uses X.11, and Heath concluded that X.10 should probably be removed from the *Elements* (Heath, 1908). Because an admissible order must place every proposition after the propositions it uses, this citation is excluded from the primary specification, which keeps 1,676 dependencies; a sensitivity specification keeps it and moves X.10 immediately after X.11.

The *Ethics* of Spinoza is the external replication fixed in the plan: the layout of the *Ethics* imitates Euclid (Steenbakkens, 2009), and Spinoza's geometric method has long invited comparison with the *Elements* (Nadler, 2006). Its dependency graph was compiled by Tredwell, for whom only explicit numbered citations count as edges (Tredwell, 2009); we use Tredwell's dependency table (Autio, 2015) at a fixed commit, merge corollaries with their propositions, and keep the five parts as blocks, which leaves 741 dependencies among 259 propositions.

Problem formulation. The text is a sequence of blocks: Books I to XIII, with Book X divided into its three groups, each introduced by its own definitions. The order of the blocks and their contents are kept as they are; within a block, an order is admissible if every proposition comes after all the propositions it uses, that is, if it is a linear extension of the partial order generated by the dependencies inside the block. Of the 1,676 dependencies, 619 (36.9%) lie inside a block and the others point from a block to a later one (Figure 1); the notation is collected in Table 1. For an order π that places proposition v at position $\pi(v)$, the total dependency length is given by Equation (1), where E is the set of dependencies and p_{uv} the weight of a dependency (one in the primary specification).

$$T(\pi) = \sum_{(u,v) \in E} p_{uv} (\pi(v) - \pi(u)) \quad (1)$$

Writing $\pi(v) = ob + ob(v)$, with ob the number of propositions in the blocks before block b and $ob(v)$ the position of v inside its block, the total length separates into a constant K and one term per block,

$$T(\pi) = K + \sum_{b=1}^m \sum_{v \in B_b} w_v \sigma_b(v) \quad (2)$$

$$w_v = \sum_{u: (u,v) \in E} p_{uv} - \sum_{x: (v,x) \in E} p_{vx} \quad (3)$$

as in Equation (2), where the weight w_v of a proposition, given by Equation (3), is the number of propositions that v uses minus the number that use v when every dependency has unit weight.

Each block is therefore an instance of single-machine scheduling with precedence constraints,

unit processing times and a weighted sum of completion times, written $1|prec,p_j=1|\sum w_j C_j$ in the

standard notation, and the blocks can be optimised independently (Lemma S1 in the supplementary material). A proposition with $w_v < 0$ is cited more often than it cites, so the objective pulls it towards the end of its block, where it is closer to the propositions of later blocks that use it; a proposition with $w_v > 0$ is pulled towards the beginning. The constant K depends only on the order of the blocks. Splitting each weight into the dependencies inside the block and those with other blocks separates the cost of a

block, $T_b = v \in B_b w_{vob}(v)$, into the length T_{bint} of its internal dependencies, given by Equation (1) restricted to the set E_b of dependencies inside block b , and a linear term, Equation (4), where e_v and s_v count the dependencies entering v from earlier blocks and leaving it towards later blocks. The first term alone is the alternative objective of the multiverse; the second pulls widely cited propositions towards the end of their block whatever their internal structure.

$$T_b = T_b^{int} + \sum_{v \in B_b} (e_v - s_v) \sigma_b(v) \quad (4)$$

Table 1: Notation: The Symbols Used in the Paper and Their Meaning.

Symbol	Meaning
$\square, \square_{\square}$	Dependencies, $(\square, \square) \in \square$ when the proof of $\square$ cites $\square$, and their weights
$\square_{\square}, \square_{\square}, \square_{\square}$	Block $\square$, its number of propositions and its internal dependencies
$\square(\square), \square_{\square}(\square)$	Position of $\square$ in the whole text and inside its block
$\square_{\square}$	Weight of $\square$: dependencies entering $\square$ minus dependencies leaving it (Equation (3))
$\square_{\square}, \square_{\square}$	Dependencies entering $\square$ from earlier blocks and leaving it towards later blocks
$\square$	Part of the total length fixed by the order of the blocks (Equation (2))
$\square_0, \square_{min}, \square_{text}$	Expected total length under uniformly random admissible orders, its minimum, and the length of the text
Ω	Optimality index: zero at the random expectation, one at the optimum (Equation (5))
$\square(\square), \square(\square), \square(\square)$	Linear extensions of an ideal $\square$, of its complement, and of a component $\square$
$\square(\square)$	Ratio of a set: mean weight of its propositions
$LB_{\square}$	Certified lower bound on the optimum of block $\square$ (Lemma S3)

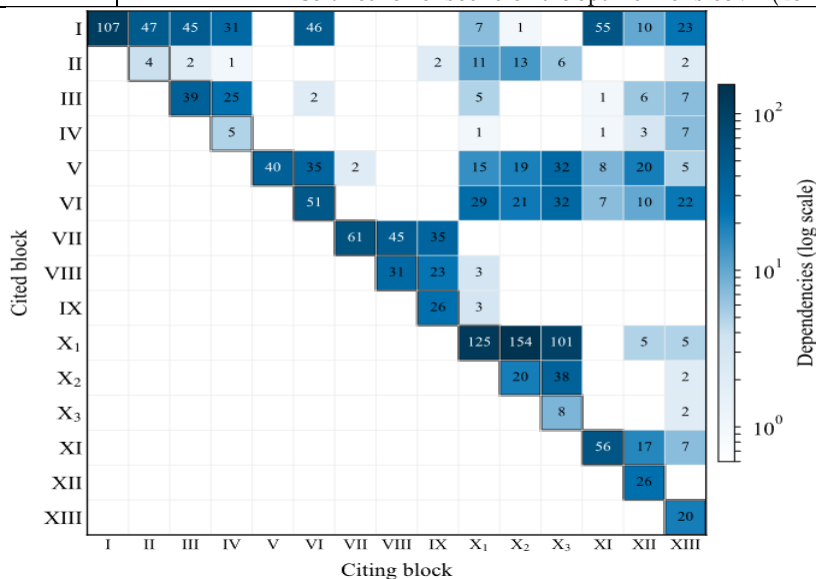


Figure 1: Dependencies Between the Blocks of the Elements (Primary Specification). Row: Block Of the Cited Proposition; Column: Block of the Citing Proposition; Outlined Diagonal: Dependencies Inside a Block, The Only Ones Whose Length the Order Inside the Blocks Can Change. Every Dependency Points from a Block to Itself or to a Later Block. X1 To X3: The Three Groups of Book X.

Let E_0 be the expected total length under uniformly random admissible orders, drawn independently in each block, and T_{min} its minimum over admissible orders. The optimality index in Equation (5) places the text on a linear scale on which a random admissible order scores zero and an optimal order scores one; negative values mean that

the text is longer than the average admissible order. It is the constrained analogue of the optimality score Ω used for syntactic dependencies (Ferrer-i-Cancho et al., 2022), whose baseline allows every permutation of the words; here the baseline respects the logical dependencies.

The same index is computed for each block,

because mixing lengths from sequences of different length is misleading (Ferrer-i-Cancho and Liu, 2014)

$$\Omega = \frac{E_0 - T_{\text{text}}}{E_0 - T_{\text{min}}} \quad (5)$$

Complexity. Minimising the total dependency length over admissible orders is NP-hard, even for a single block with no dependencies from other blocks. With unit weights, that restricted problem is the directed minimum linear arrangement problem, and the directed version on acyclic graphs is NP-complete (Shiloach, 1979); the reduction, which maps an instance of that problem to a text with one block, is given in the supplementary material (Proposition S1).

The problem is also a special case of $1|prec, p_j=1|\sum w_j C_j$, which is NP-hard even when all processing times are equal (Graham et al., 1979); it becomes polynomial for tree-like and series-parallel precedence constraints (Graham et al., 1979), but the methods used here do not rely on such structure.

Approximation does not help here: any order

and blocks of different sizes can only be compared on a normalised scale.

consistent with the Sidney decomposition is within a factor of two of the optimum (Margot et al., 2003), virtually all 2-approximations respect the Sidney decomposition (Correa and Schulz, 2005), the problem is not approximable within $2 - \varepsilon$ under a strengthened unique games conjecture (Bansal and Khot, 2009), and the unit-time case is as hard to approximate as the general problem (Woeginger, 2003). A guarantee within a factor of two leaves the index undetermined, so T_{min} must be exact. Computing E_0 involves the uniform distribution over linear extensions, and counting linear extensions is #P-complete (Brightwell and Winkler, 1991). Neither quantity can therefore be expected to have an efficient algorithm for arbitrary partial orders; both are computed exactly here by exploiting the structure of each block (Table 2).

Table 2: Structure Of the Blocks That Makes the Exact Computation Feasible.

Block	□	Links	Comp.	Max.	Ideals	Sidney
I	48	107	2	47	201,978	5 / 42
II	14	4	10	2	3	10 / 2
III	37	39	5	33	555,182	16 / 10
IV	16	5	11	6	12	11 / 6
V	25	40	1	25	6,464	8 / 8
VI	33	51	2	32	358,258	4 / 30
VII	39	61	3	36	44,160	11 / 17
VIII	27	31	4	23	5,872	9 / 7
IX	36	26	13	13	364	15 / 9
X (first group)	47	125	3	43	169,988	5 / 40
X (second group)	37	20	21	6	17	24 / 6
X (third group)	31	8	23	3	5	30 / 2
XI	39	56	3	36	99,240	9 / 19
XII	18	26	3	16	61	7 / 9
XIII	18	20	2	17	6,487	6 / 10
Ethics I	36	69	2	35	35,916	8 / 20
Ethics II	49	114	4	46	12,067,264	10 / 37
Ethics III	59	140	4	56	32,661,272	7 / 53
Ethics IV	73	98	13	61	above limit	22 / 36
Ethics V	42	43	13	28	18,816	16 / 15

n: propositions; links: dependencies inside the block; comp.: weakly connected components; max.: propositions in the largest component; ideals: order ideals of the largest component enumerated by the null

algorithm; Sidney: number of pieces of the Sidney decomposition and size of the largest piece.

Null expectation. By linearity, E_0 only requires the expected position of every proposition, Equation (6).

$$E_0 = K + \sum_{b=1}^m \sum_{v \in B_b} w_v \mathbb{E}[\sigma_b(v)] \quad (6)$$

Algorithm 1 obtains these expectations exactly. Each block is split into its weakly connected components; for a component C , the algorithm enumerates its order ideals (sets closed under

predecessors) level by level and counts, for every ideal I , the linear extensions of I , $f(I)$, and of its complement, $g(I)$. In the language of learning theory, the ideals are knowledge states and each admissible

order is a learning path, since a learning path is a maximal chain of knowledge states (Doignon and Falmagne, 2015) and learning spaces generalize partial orders (Falmagne and Doignon, 2011). Every linear extension of C is a maximal chain of ideals, so the expected rank of an element is given by Equation

$$\mathbb{E}[r(x)] = \frac{1}{e(C)} \sum_{I \in \mathcal{I}_x} (|I| + 1) h_x(I) \quad (7)$$

In a uniform linear extension of a disjoint union of components, the order inside each component is uniform and independent of how the components are

$$\mathbb{E}[\sigma(x)] = \mathbb{E}[r_i(x)] \frac{n+1}{a_i+1} \quad (8)$$

For the same reason, the number of linear extensions of the union is the multinomial product Equation (9), so a block never has to be enumerated as a whole. The algorithm uses exact rational

$$e(C_1 \sqcup \dots \sqcup C_m) = \frac{n!}{a_1! \dots a_m!} \prod_{i=1}^m e(C_i) \quad (9)$$

Algorithm 1 Exact expected positions under uniformly random admissible orders

- 1: **Input:** partial order P of one block (its propositions and internal dependencies)
 - 2: **Output:** $\mathbb{E}[\sigma(v)]$ for every proposition v of the block
 - 3: split P into weakly connected components $C_1, \dots, C_m$; $n \leftarrow |P|$
 - 4: **for each** component C with a elements **do**
 - 5: enumerate the ideals of C by size; $f(\emptyset) \leftarrow 1$
 - 6: **for each** ideal I and minimal x of $C \setminus I$ **do** $f(I \cup \{x\}) \leftarrow f(I \cup \{x\}) + f(I)$
 - 7: $g(C) \leftarrow 1$; **for each** ideal I by decreasing size **do** $g(I) \leftarrow \sum_x g(I \cup \{x\})$
 - 8: **assert** $f(C) = g(\emptyset)$
 - 9: **for each** $x \in C$ **do** $r(x) \leftarrow$ expected rank by Equation (7)
 - 10: **for each** $x \in C$ **do** $\mathbb{E}[\sigma(x)] \leftarrow r(x) (n+1)/(a+1)$
 - 11: **assert** $\sum_v \mathbb{E}[\sigma(v)] = n(n+1)/2$ and $\mathbb{E}[\sigma(u)] < \mathbb{E}[\sigma(v)]$ for every dependency
 - 12: **return** $\mathbb{E}[\sigma(\cdot)]$
-

Exact sampling. The same counts give exact uniform samples, which the randomization test needs. Starting from the empty ideal, the next element is drawn with the probability in Equation (10), which makes every linear extension of the component equally likely, and the components are interleaved by a uniformly random arrangement of their labels. Sampling linear extensions is usually approximate: the Karzanov-Khachiyan chain (Karzanov and Khachiyan, 1991) mixes, by a path-

(7), where $\mathcal{I}_x$ is the set of ideals I such that x is minimal in $C \setminus I$, $h_x(I) = f(I) g(I \cup \{x\})$ counts the linear extensions in which x comes right after the elements of I , and $e(C) = f(C) = g(\emptyset)$ is the number of linear extensions.

interleaved, which gives the interleaving identity Equation (8) for an element of a component with a_i of the n elements (Proposition S2).

arithmetic for the primary analysis and a vectorised floating-point version for the replicates; the two agree to a relative error below 10^{-9} on every block.

coupling analysis of Bubley and Dyer, in $O(n^3 \log n)$ steps (Wilson, 2004); perfect samplers built on coupling from the past (Propp and Wilson, 1996)

Remove the approximation, and an algorithm for perfectly generating samples of linear extensions runs in the same expected time (Huber, 2006); in practice, actual mixing times are far below the worst-case bounds (Talvitie et al., 2017). Exact enumeration of the ideals avoids these approximations wherever the components are small enough.

$$\Pr[x \mid I] = \frac{g(I \cup \{x\})}{g(I)} \quad (10)$$

Exact optimum. For a set I of propositions define the ratio Equation (11).

$$\rho(I) = \frac{1}{|I|} \sum_{v \in I} w_v \quad (11)$$

Let $J(P)$ be the family of non-empty initial sets (ideals) of the partial order P of a block. The first Sidney piece is the inclusion-minimal initial set of

$$S_1 = \min_{\subseteq} \left\{ I \in J(P) : \rho(I) = \max_{J \in J(P)} \rho(J) \right\} \quad (12)$$

By Sidney's decomposition theorem (Sidney, 1975), some optimal order processes, one after another, initial sets of maximum ratio obtained by repeatedly removing such a set; we compute the finest decomposition by a parametric minimum cut (Margot et al., 2003), and optimal schedules of the Sidney pieces can be concatenated (Woeginger, 2003). Inside each piece, the optimum follows from the recursion Equation (13) over its ideals, because the last element of any prefix J of an admissible order

largest ratio, Equation (12); the next pieces are obtained in the same way from what remains.

is maximal in J and occupies position $|J|$; this combination of dynamic programming over the partial-order decomposition (Sidney and Steiner, 1986) is exact and needs no solver. Figure 2 shows the two computations on the first 8 propositions of Book I, with only the dependencies among them: the 11 ideals of the component of 7 propositions, the probabilities with which the exact sampler moves between them, and the path that the recursion selects.

$$\text{OPT}(J) = \min_{x \in \max(J)} [\text{OPT}(J \setminus \{x\}) + w_x |J|] \quad (13)$$

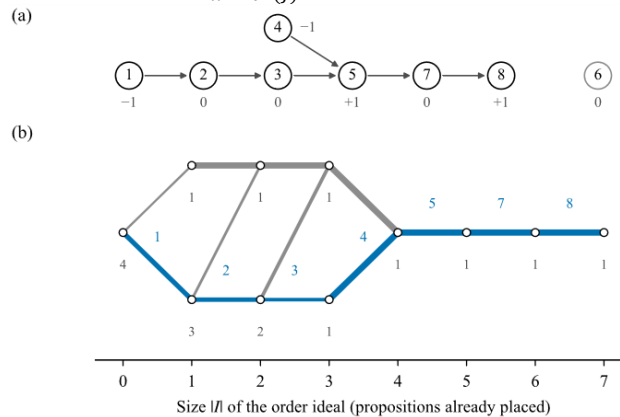


Figure 2: The Null Algorithm and the Optimum Recursion on a Small Instance: The First 8 Propositions of Book I, With Only the Dependencies Among Them. (A) Dependency Graph (Hasse Diagram); Below Each Proposition, Its Weight w_v in This Instance; 1,6 Forms A Component of Its Own. (B) Lattice Of the Order Ideals of the Larger Component, From the Empty Set (Left) To the Whole Component (Right).

Each linear extension is a path from left to right; the number below an ideal is $g(I)$, the number of ways to complete it; the width of an edge is the probability Equation (10) with which the exact sampler takes it, and the highlighted path, labelled with the proposition added at each step, is the optimal order found by the recursion Equation (13).

As an independent check, each block is also solved as an integer programme with assignment variables x_{vt} , equal to one when proposition v occupies position t : the objective Equation (14) is minimised subject to the assignment constraints

Equation (15) and to the disaggregated precedence constraints Equation (16) for every t and every dependency in the transitive reduction R of the block.

The model is solved with HiGHS (Huangfu and Hall, 2018) at a relative gap of 10^{-9} , and every solution is checked for integrality, feasibility and bounds before it is accepted.

Exact methods for the linear ordering problem, which includes single-machine weighted completion time (Grötschel et al., 1984), also include branch-and-bound with a Lagrangean-type lower bound (Potts, 1980); here the integer programme serves only as a

baseline, and within its time limit it reached the same certified optimum as the decomposition in 14 of the 15 blocks.

$$\min \sum_v \sum_{t=1}^n t w_v x_{vt} \quad (14)$$

$$\sum_{t=1}^n x_{vt} = 1, \quad \sum_v x_{vt} = 1 \quad (15)$$

$$\sum_{s \leq t} x_{vs} \leq \sum_{s < t} x_{us}, \quad (u, v) \in R \quad (16)$$

Certified bound. Independently of any solver, each block admits the lower bound Equation (17) (Lemma S3), where $I(u, v)$ is the set of propositions between u and v in the partial order.

$$LB_b = \sum_{(u,v)} p_{uv} (1 + |I(u, v)|) + \sum_v c_v \quad (17)$$

The term c_v of each proposition is given by Equation (18), where e_v and s_v are the dependencies entering v from earlier blocks and leaving it towards later blocks, and α_v and δ_v are the numbers of ancestors and descendants of v inside the block.

$$c_v = e_v(1 + \alpha_v) - s_v(n_b - \delta_v) \quad (18)$$

The bound serves as a check on every optimum and, because it does not depend on any solver, it also certifies a lower bound on the index, Equation (19), which for the Elements is 0.145.

$$\Omega \geq \Omega_{LB} = \frac{E_0 - T_{\text{text}}}{E_0 - K - \sum_b LB_b} \quad (19)$$

Monte Carlo null for one component. In Part IV of the Ethics, one component of 61 propositions exceeds the limit of 50,000,000 ideals fixed for the exact method.

For that component only, the plan replaces the exact expectation by the average of 10,000 independent chains of the slightly modified chain of

Bubley and Dyer (Huber, 2006), each run for $3n \ln n$ steps from a fixed linear extension.

Before the plan was frozen, this rule reproduced the exact expected ranks of a component of 46 propositions (joint $p = 0.82$) and of one of 56 propositions (joint $p = 0.86$). The Monte Carlo error of E_0 is propagated to the index by Equation (20).

$$se(\Omega) = se(E_0) \frac{|T_{\text{text}} - T_{\min}|}{(E_0 - T_{\min})^2} \quad (20)$$

Pre-registered analysis. The analysis plan, including the hypotheses, the decision rules, the validity checks and the robustness analyses, was fixed and versioned in the repository before any cost of the text's order was computed; the code that computes that cost refuses to run without the tag of the frozen plan and logs every access.

The only quantities computed beforehand depend on the dependency graph alone. The question is descriptive: the analysis measures a property of the text and makes no claim about the intentions of its author or its editors.

The primary hypothesis is that the order of the Elements is closer to the optimum than to a random admissible order. It is supported if Ω exceeds 0.5 and

the text's length is smaller than that of uniformly random admissible orders, with the one-sided randomization p value Equation (21) computed from $M = 100,000$ orders drawn exactly and independently in each block, at the 5% level; this form of the p value never returns zero (Phipson and Smyth, 2010). The threshold of 0.5 is the midpoint of the scale: the smallest effect at which the order is closer to the optimum than to chance. Because E_0 , $T_{\min}$ and T_{text} are exact, Ω has no sampling error. A calibration on the dependency graph alone, made before the plan was frozen, showed that one null standard deviation corresponds to 0.047 on the scale of Ω , so that the smallest effect the test detects at the same level is about 0.077, far below the threshold.

$$p = \frac{1 + \#\{s: T_s \leq T_{\text{text}}\}}{M + 1} \quad (21)$$

A positive result had to pass four checks fixed in the plan. First (V1), admissible orders drawn at

random and treated as if they were the text must trigger the decision rule in fewer than 5% of 1,000 draws. Second (V2), in 100 sets of placebo citations, each citation inside a block is redirected to a random earlier proposition of the same block, keeping the number of prerequisites of every proposition and all the dependencies between blocks; the text's index must exceed the 95th percentile of the placebo indices, so that the economy comes from what is cited and not merely from citations pointing backwards. Third (V3), the index must remain above 0.5 when

$$\frac{1}{|\mathcal{S}|} \# \left\{ s \in \mathcal{S} : \Omega_s \leq \frac{1}{2} \right\} \geq c \quad (22)$$

The secondary analyses fixed in the plan are the replication on the Ethics, with the same decision rule and 10,000 random orders; per-block tests, a family of 15 one-sided randomization tests adjusted with Holm's step-down procedure Equation (23), where $p(1) \leq \dots \leq p(m)$ are the ordered p values; the removal of one block at a time.

200 replicates that delete a random 10% of the

$$\tilde{p}_{(k)} = \max_{j \leq k} \min \{1, (m - j + 1) p_{(j)}\} \quad (23)$$

Exploratory analyses. Three analyses were declared exploratory in the plan, and their p values are adjusted with the Benjamini-Hochberg procedure. For every proposition, the relocation gain Equation (24) is the largest reduction of its block's

$$\Delta_v = T_b(\sigma) - \min_{\ell_v \leq q \leq h_v} T_b(\sigma_{v \rightarrow q}) \quad (24)$$

For every block, the compatibility with the Sidney decomposition Equation (25) is the fraction of pairs in different Sidney pieces, $s(x) < s(y)$, that the text presents in the same order as the decomposition.

$$\kappa_b = \frac{\#\{(x, y) : s(x) < s(y), \sigma(x) < \sigma(y)\}}{\#\{(x, y) : s(x) < s(y)\}} \quad (25)$$

Deviations from the plan. The primary analysis was run first and recorded as null; the other analyses were run afterwards, with the same seeds, under the same frozen plan.

The check V2, which the plan requires only for a positive result, was run as an informative control and enters no decision. After the primary analysis, an error was found in the implementation of the Sidney decomposition: when several initial sets tied at the maximum ratio, it merged them into one piece instead of returning the finest decomposition fixed in the plan.

A merged piece is still an initial set of maximum ratios, so every optimum computed before the correction was exact and no index changed, but the

Heath's notes are included, when porisms and lemmas are not attributed to their propositions, when repeated citations are weighted and when citations of definitions, postulates and common notions are added. Fourth (V4), every optimum must be certified and every bound respected. The result is null if the rule fails, and an informative null if, in addition, Ω stays at or below 0.5 in a fraction of at least $c = 95\%$ of the set $\mathcal{S}$ of specifications of the multiverse, Equation (22).

dependencies; and a multiverse of the 64 combinations of six choices: Heath's notes included or not, porisms and lemmas attributed or not, repeated citations weighted or not, citations of fixed elements added or not, all dependencies or only those inside the blocks in the objective, and the forward dependency of X.10 excluded or repaired.

cost obtained by moving that proposition alone to another admissible position q , between the last proposition it uses, l_v , and the first one that uses it, h_v , with the rest of the order fixed.

Finally, the density of each block is compared with its index, and the relocation gains with the fraction of each proposition that Heiberg's critical edition of the Greek text marks as suspect (Heiberg, 1883).

merged pieces made the specifications of the multiverse that count only the dependencies inside the blocks fall back on the integer programme.

The corrected code, checked against exhaustive enumeration, was used for the multiverse, the exact baseline, the benchmark and the exploratory comparison with the decomposition. The replication on the Ethics, the multiverse, the exact baseline, the benchmark and the exploratory analyses were run on a virtual machine with 4 virtual cores at 2.1 GHz and 15.7 GiB of memory, not on the desktop computer that ran the other analyses; a first run of the replication on the Ethics on the desktop computer had been stopped for lack of memory. The scaling curve of the benchmark on synthetic instances and

the integer programme of the exact baseline for the parts of the Ethics were not run for lack of time.

Software and reproducibility. The pipeline is written in Python and uses NumPy, SciPy and HiGHS. It downloads the pinned source files, verifies their checksums, runs the analysis in the order fixed by the plan and writes every result to a machine-readable file that records the commit, the data checksum, the plan tag, the random seed, the library versions and the detected hardware.

The document is compiled from a single content

model, and every number in the text is read from those files rather than transcribed (Figure 3).

Automated tests compare every algorithm with exhaustive enumeration on random small instances, and quality gates block the build when a check fails.

The primary analysis, with exact rational arithmetic for E_0 and the 100,000 random orders, ran in 57 s on a desktop computer with 8 processor cores and 15.6 GiB of memory, with the numerical libraries restricted to one thread.

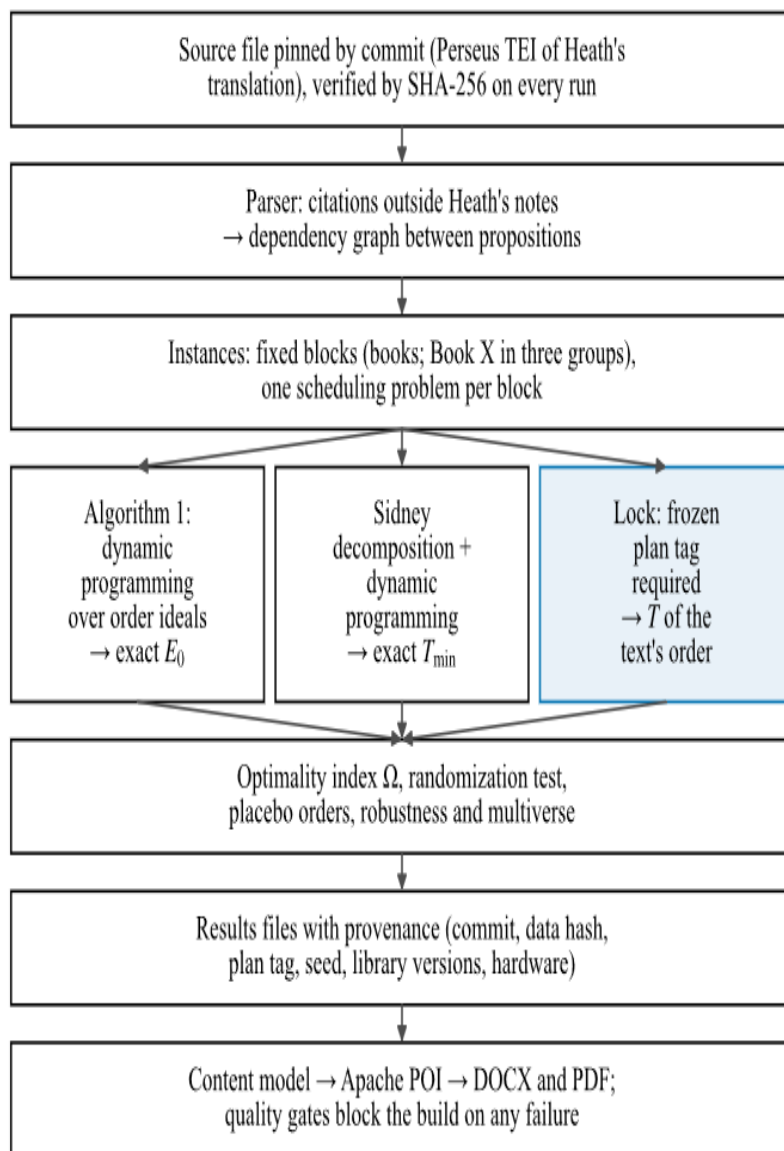


Figure 3: Architecture Of the Pipeline. The Code That Computes the Cost of the Text's Order Runs Only Under the Tag of the Frozen Analysis Plan.

3. RESULTS

Primary analysis. The total dependency length of the text is 151,743. Uniformly random admissible orders have an expected length of 156,170.2, and the

shortest admissible order has length 144,005; the constant $K = 149,911$, fixed by the order of the blocks, is part of all three.

The text therefore removes 4,427.2 of the 12,165.2

units of length that separate the random expectation from the optimum (Figure 4), and $\Omega = 0.364$.

None of the 100,000 random admissible orders was as short as the text, whose length lies 7.8 null standard deviations below the expectation, so the randomization p value is 1.0×10^{-5} , the smallest value the test can return. In relative terms, the text

exceeds the minimum by 5.4%, and the random expectation exceeds it by 8.4%.

The order is thus far more economical than chance, but Ω falls short of the threshold of 0.5 by 0.136, well beyond the minimum detectable effect, 0.077. By the rule fixed in the plan, the primary hypothesis is not supported and the result is null.

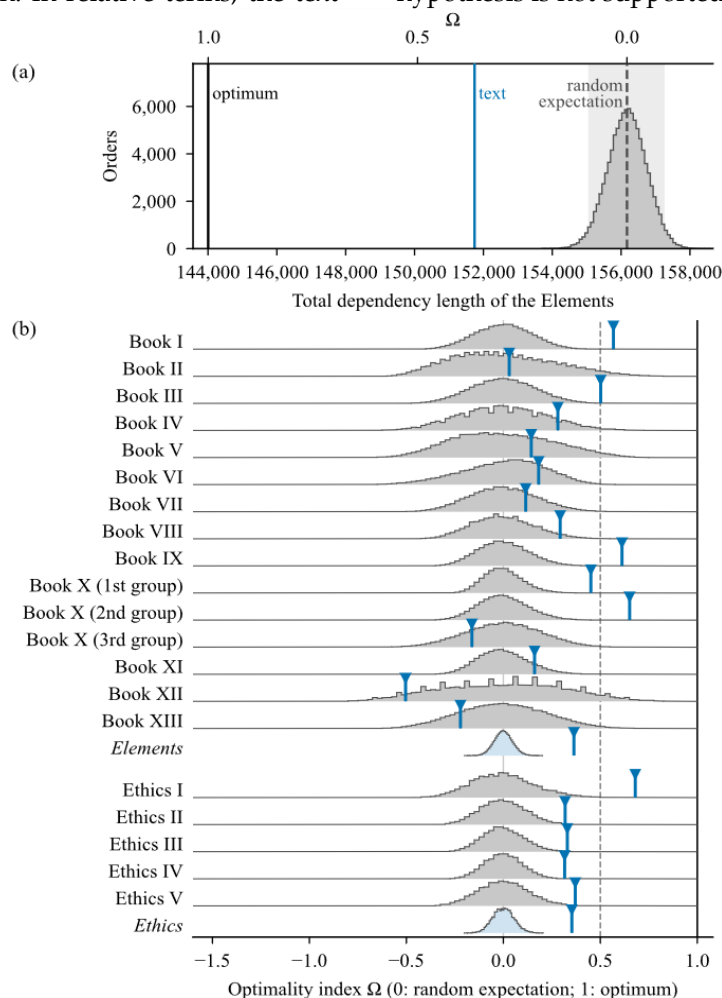


Figure 4: Null Distributions and the Position of The Text. (A) Total Dependency Length of The 100,000 Uniformly Random Admissible Orders of The Elements (Histogram; Shaded: Central 95%), Of the Text and Of the Optimum; Top Axis: The Same Scale in Units Of Ω . (B) Distribution Of Ω Under the Null in Each Block of the Elements, In Each Part Of The Ethics And In Each Whole Text (Blue Fill), From The Samples Of The Randomization Tests; Each Distribution Is Scaled To Its Own Maximum. Marker: Ω Of the Text; Dashed Line: Pre-Registered Threshold; Solid Line at One: The Optimum. Mass Below $\Omega = -1.6$ Is Not Drawn.

Heterogeneity across blocks. The economy is uneven (Table 3; Figure 4b): the index ranges from -0.50 to 0.65 across the 15 blocks. The threshold is exceeded in 4 blocks, Books I, III and IX and the second group of Book X, with Book III barely above it. A negative index, which means that the text is longer than the average random admissible order, appears in 3 blocks: the third group of Book X and Books XII and XIII. After Holm correction, the text is significantly shorter than random admissible orders

in 5 blocks, Books I, III and IX and the first two groups of Book X.

The null distributions show why the per-block tests differ in power: the central 95% of the null index spans 0.99 units of Ω in Book II but only 0.39 in the first group of Book X, so a given index is far stronger evidence in the larger blocks.

Removing one block at a time keeps Ω between 0.315 and 0.400, so no single block decides the overall result.

Table 3: Reduction In Total Dependency Length Achieved by the Text and Achievable by the Optimum in Each Block, Relative to the Expectation Under Uniformly Random Admissible Orders.

Block	n	Links	Achieved	Achievable	Holm p
I	48	107	882.5	1,554.5	2.8×10^{-4}
II	14	4	8.0	262.0	1.0
III	37	39	400.3	797.3	0.0086
IV	16	5	42.6	151.6	1.0
V	25	40	113.5	789.5	1.0
VI	33	51	150.7	828.7	1.0
VII	39	61	72.6	629.6	1.0
VIII	27	31	84.6	288.6	0.42
IX	36	26	527.9	862.9	2.8×10^{-4}
X (first group)	47	125	1,122.1	2,486.1	0.0014
X (second group)	37	20	1,144.5	1,757.5	1.5×10^{-4}
X (third group)	31	8	-127.7	787.3	1.0
XI	39	56	103.2	641.2	1.0
XII	18	26	-44.6	88.4	1.0
XIII	18	20	-53.0	240.0	1.0
All blocks	465	619	4,427.2	12,165.2	1.0×10^{-5}

n: propositions in the block; links: dependencies inside the block; achieved: $E_0 - T_{\text{text}}$; achievable: $E_0 - T_{\text{min}}$. The index of each block is the ratio of the two reductions. Holm p: one-sided randomization tests adjusted with Holm's procedure; for all blocks together, the p value of the primary test.

Where the text departs from the optimum. These analyses are exploratory. Of the 465 propositions, 312 (67.1%) could be moved on their own to another admissible position that shortens their block, but the large gains are concentrated in a few of them (Table 4; Figure 5). The largest belongs to VI.17, "If three straight lines be proportional, the rectangle contained by the extremes is equal to the square on the mean", which the text states at position 17 of Book VI and whose best position is the last one, 33.

The next are I.32, on the exterior angle and the angle sum of a triangle, and I.6, the converse of the

theorem on the base angles of an isosceles triangle, which could both move to the end of Book I. These three propositions are cited far more often than they cite, and only by propositions of later books, so the objective pulls them towards the end of their own book, next to the propositions that use them, whereas the text states them early.

The opposite happens to X.112 and X.113, which use many earlier results and whose best position is at the start of their group.

The same pattern shapes an optimal order of Book I as a whole (Figure 6): most propositions keep nearly their place, while I.6, I.14 and I.32 move to the last positions; the internal dependencies of the two orders have similar total lengths (973 and 937), so most of what the optimum gains comes from the dependencies that leave the book.

Table 4: Propositions Whose Individual Relocation Shortens Their Block the Most (Exploratory).

Proposition	Position in text	Best position	Gain $\square_{\square}$
VI.17	17	33	311
I.32	32	48	272
I.6	6	48	233
X.112	28	1	175
X.17	17	47	164
X.113	29	1	147
V.11	11	15	139
VII.14	14	39	119
XI.38	38	10	116
V.9	9	25	105

Positions are counted inside the block (for Book X, inside its group); the gain is the reduction of the block's cost when only that proposition moves. The

enunciations are listed in the supplementary material.

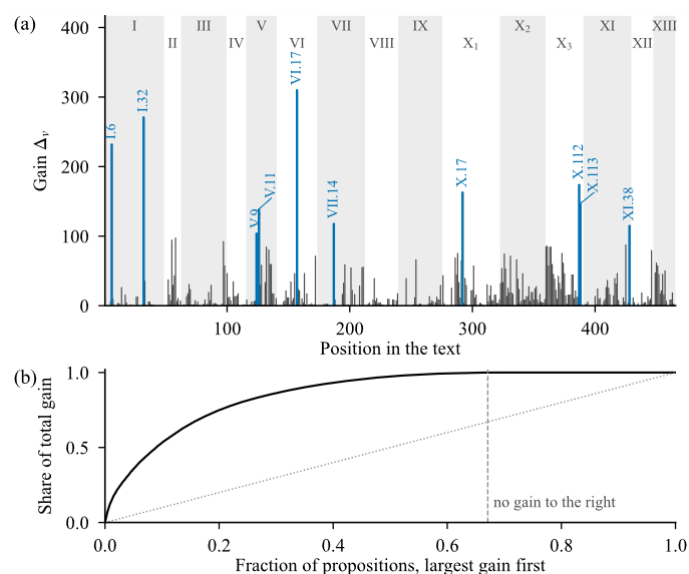


Figure 5: Relocation Gains Along The Text (Exploratory). (A) Gain Δv Of Each Proposition, In the Order of the Text; Shading Separates the Blocks; Highlighted and Labelled: The Ten Largest Gains. (B) Share Of the Total Gain Accounted for by the Propositions with the Largest Gains; Dotted: Equal Gains; Dashed: The Last Proposition with a Positive Gain.

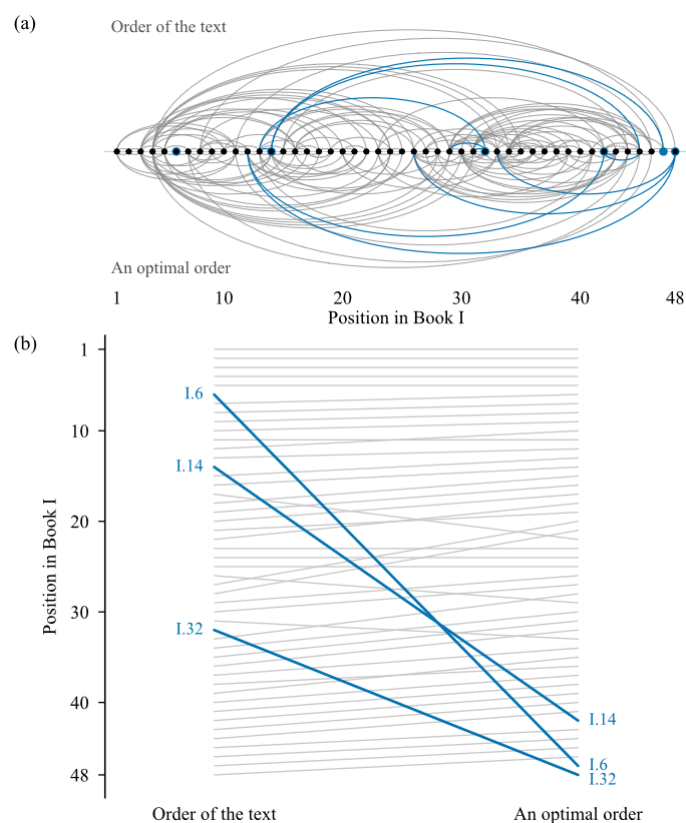


Figure 6: Book I In the Order of The Text and In an Optimal Order. (A) Internal Dependencies Drawn as Arcs, Above the Axis in the Text and Below It in the Optimal Order; The Height of an Arc Is Proportional to Its Length; Highlighted: Propositions That Move Eight Positions or More and Their Dependencies. (B) Position Of Each Proposition in the Two Orders; Highlighted: The Same Moves.

Neither of the two editorial signals examined is associated with these gains. The relocation gain is unrelated to the fraction of each proposition that

Heiberg brackets (Spearman $\rho = 0.056$ over 457 propositions, adjusted $p = 0.23$), and of the 7 propositions whose notes in Heath's commentary

mention interpolation or spurious material, the best placed ranks 66th by gain. X.10, the proposition with the forward citation, has a gain of 0: its place is not costly in this sense. Across blocks, the index is not associated with the density of dependencies (Spearman $\rho = -0.37$, adjusted $p = 0.23$). The text presents in the order of the Sidney decomposition between 48% and 81% of the pairs of propositions in different pieces; the lowest values occur in the third group of Book X and in Book XIII, two of the blocks with a negative index, but the highest occurs in Book XII, the third one, so this compatibility does not track the index.

Validity checks and robustness. The checks were designed for a positive result, but they also bear on the null (Table 5). The decision rule fired in 0% of the 1,000 placebo orders, whose largest index was 0.137. With placebo citations (V2), which keep every citation pointing backwards inside its block but change its target, the index of the text order falls to a mean of 0.069, and the largest of the 100 placebo indices, 0.150, stays far below the real one: the economy depends on which propositions are cited, not merely on citations pointing backwards. Under the four alternative readings of the citations (V3), the index was 0.365 with Heath's notes, 0.320 without attributing porisms and lemmas, 0.372 with repeated citations weighted and 0.352 with citations of fixed

elements. Every optimum was certified and every block respected its lower bound (V4). In the 32 specifications of the multiverse that count every dependency, Ω ranges from 0.307 to 0.372 and never exceeds the threshold (Figure 7); the lowest values come from not attributing porisms and lemmas to their propositions. When the objective counts only the dependencies inside each block, it ranges from 0.492 to 0.535 and exceeds the threshold in 24 of the 32 specifications; it stays at or below it only when porisms and lemmas are not attributed and repeated citations are not weighted. The objective is by far the most consequential choice (Table 6). Block by block, the change is not uniform (Figure 7c): counting only the internal dependencies raises the index in 9 of the 15 blocks and lowers it in the others, most in Books II and IV, so the gap between the two objectives is the net contribution of the linear term of Equation (4), which rewards stating widely cited propositions late. Deleting a random 10% of the dependencies lowers the index slightly, to a mean of 0.343, and no replicate reaches the threshold (Figure 8). Because Ω stays at or below the threshold in only 62.5% of the specifications, short of the 95% fixed in the plan, the null is not informative: it holds for the total dependency length under every reading of the citations, but not when the objective counts only the dependencies inside the blocks.

Table 5: Pre-Registered Decision Rules and Validity Checks with Their Outcome, And Robustness Analyses of the Index of the Elements.

Analysis	Rule or design	Observed	Outcome
H1 (Elements)	$\square > 0.5$ and $\square < 0.05$	$\square = 0.364$; $\square = 1.0 \times 10^{-5}$	not supported
H2 (Ethics)	$\Omega > 0.5$ and $p < 0.05$	$\Omega = 0.353$; $p = 1.0 \times 10^{-4}$	not supported
H3 (blocks)	Holm-adjusted $p < 0.05$	5 of 15 blocks	secondary
V1	placebo orders: rule fires in under 5% of 1,000	fired in 0%; largest Ω 0.137	passed
V2	placebo citations: text above the 95th percentile of 100	placebo mean 0.069, largest 0.150; text above all	passed (control)
V3	$\Omega > 0.5$ under four readings of the citations	0.320 to 0.372	failed
V4	optima certified and bounds respected	all blocks	passed
Informative null	Equation (22) with $c = 95\%$	62.5% of 64 specifications	no
Robustness analyses without a decision rule			
Block removal	15 removals	0.315 to 0.400	
Deleted links	200 deletions of 10%	0.303 to 0.381 (central 95%)	
Multiverse, all links	32 specifications	0.307 to 0.372	
Multiverse, inside blocks	32 specifications	0.492 to 0.535; 24 above 0.5	

All rules were fixed in the frozen plan before any cost of the text's order was computed. V2 was required only for a positive result and was run after

the null as a control. Every optimum in every replicate and specification was certified.

Table 6: Change Of the Index of the Elements with Each Choice of the Multiverse: Mean Difference In Ω (Range) Between Specifications That Differ Only in That Choice.

Choice	All dependencies	Inside blocks
Heath's notes included	0.000 (−0.004 to 0.003)	−0.001 (−0.006 to 0.003)
Porisms and lemmas not attributed	−0.041 (−0.045 to −0.037)	−0.026 (−0.030 to −0.022)

Repeated citations weighted	0.010 (0.005 to 0.014)	0.012 (0.008 to 0.016)
Fixed elements added	-0.010 (-0.012 to -0.008)	0.000 (0.000 to 0.000)
Citation of X.10 repaired	-0.003 (-0.005 to -0.000)	-0.001 (-0.005 to 0.003)
Only dependencies inside blocks	0.175 (0.161 to 0.192)	

Columns: objective of the specifications compared, with 16 pairs per cell. Last row: the change of objective, over the 32 pairs that share the other five choices.

Under the objective inside blocks, citations of fixed elements only enter through dependencies with other blocks, so they cannot change the index.

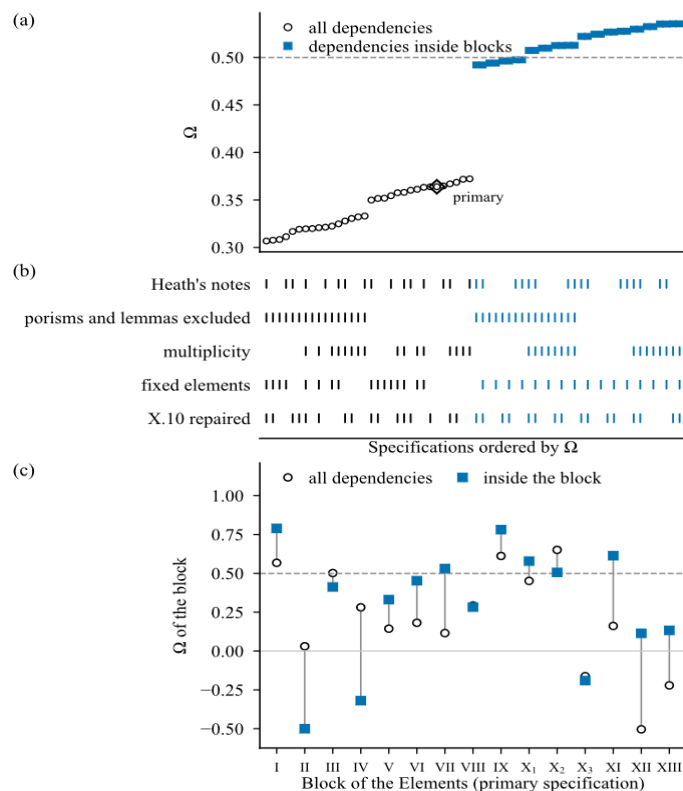


Figure 7: Multiverse Of the Elements. (A) Ω Of Every Specification in Increasing Order; Open Circles: Objective with All Dependencies; Filled Squares: Only the Dependencies Inside the Blocks; Diamond: Primary Specification; Dashed: Threshold. (B) Choices That Depart from the Primary Specification, In the Same Order and with the Same Coding. (C) Ω Of Each Block in the Primary Specification Under the Two Objectives.

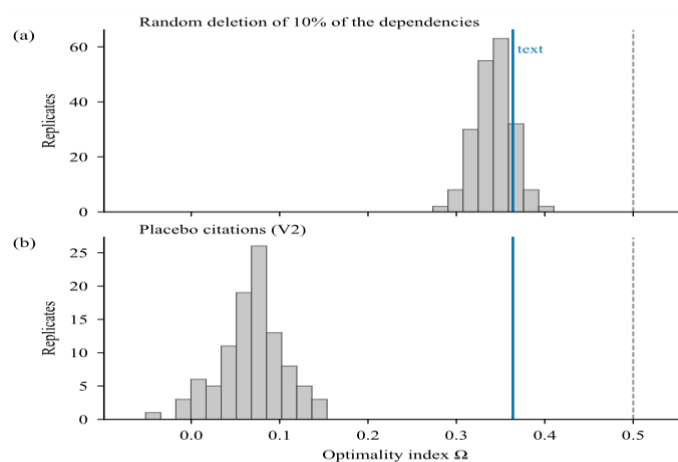


Figure 8: Distribution Of the Index When a Random 10% Of the Dependencies Is Deleted (Top) And of the Index of the Text Under Placebo Citations (Bottom). Solid Line: Index Of the Primary Specification; Dashed Line: Threshold.

Replication in Spinoza's Ethics. In the Ethics, with its five parts as blocks, the text's total dependency length is 31,611, against a random expectation of 34,171.8 and a minimum of 26,924. The text removes 2,560.8 of the 7,247.8 units that separate the two, so $\Omega = 0.353$ (Monte Carlo standard error 1.6×10^{-4}), with randomization $p = 1.0 \times 10^{-4}$ over 10,000 orders. As in the Elements, the order is significantly shorter than random admissible orders but below the threshold, so the secondary hypothesis is not supported either.

Across its parts the index ranges from 0.32 to 0.68 (Table 7; Figure 4b); only the first part exceeds the threshold. In its multiverse, the index ranges from 0.319 to 0.353 in the specifications that count every dependency, and is at most 0.534 when only the dependencies inside each part count: an upper bound, because there the integer programme did not certify the optimum of every part within its time limit.

Table 7: Reduction In Total Dependency Length Achieved by the Text and Achievable by the Optimum in Each Part of the Ethics, Relative to the Expectation Under Uniformly Random Admissible Orders.

Part	$\square$	Links	Achieved	Achievable	$\square$
I	36	69	372.4	547.4	1.0×10^{-4}
II	49	114	621.3	1,950.3	0.0032
III	59	140	508.1	1,540.1	0.0015
IV	73	98	759.5	2,402.5	6.0×10^{-4}
V	42	43	299.6	807.6	0.0029
All parts	259	464	2,560.8	7,247.8	1.0×10^{-4}

Columns as in Table 3; p : one-sided randomization test of each part, unadjusted (the plan adjusts only the blocks of the Elements), from 10,000 random orders. In Part IV, E0 is a Monte Carlo estimate for one component.

Computation. The decomposition with dynamic programming certified the optimum of every block of the Elements, whereas the integer programme, stopped after 900 s, certified it in 14 of the 15 blocks, always with the same value (Table 8). Solved piece by piece, as the decomposition allows, the integer programme certified the same optimum in all 15 blocks.

The integer programme needed 1,096.3 s for all the blocks, and the decomposition with dynamic programming 3.0 s. The finest decomposition splits the blocks into 170 pieces, the largest with 42 propositions.

A simple order that respects it, taking inside each piece the available proposition of largest weight, lies

within a factor of 1.028 of the optimum in every block, inside the guarantee of two, yet using it instead of the optimum would change the index from 0.364 to 0.418.

In the 5 parts of the Ethics, the decomposition reproduced the optima of the replication; the integer programme was not run on them.

On the benchmark (Table 9), the null algorithm took a median of 1.99 s for all the blocks, ten thousand exact samples of the whole text 5.15 s and the exact optimum 3.28 s.

On Book I, the decomposition with dynamic programming was 103.3 times faster than the integer programme (95% bootstrap interval 102.0 to 104.9), and the floating-point null was 8.6 times faster than exact rational arithmetic (8.3 to 9.1). Splitting the blocks into components made the null algorithm 3.2 times faster than enumerating the ideals of each whole block, on the 11 blocks where that is feasible (3.1 to 3.3).

Table 8: Exact Optimum of Each Block by the Integer Programme and by the Sidney Decomposition with Dynamic Programming, And Quality of a Sidney-Compatible Order.

Block	Integer programme (s)	Decomposition (s)	Pieces	Heuristic ratio
I	55.34	0.41	5	1.002
II	0.01	0.03	10	1.000
III	10.28	0.31	16	1.000
IV	0.09	0.03	11	1.000
V	1.07	0.09	8	1.000
VI	15.38	0.33	4	1.007
VII	30.06	0.32	11	1.005
VIII	0.74	0.09	9	1.002
IX	8.22	0.22	15	1.003
X (first group)	900.08†	0.44	5	1.020
X (second group)	3.22	0.28	24	1.000
X (third group)	0.04	0.16	30	1.000
XI	71.68	0.23	9	1.013

XII	0.03	0.04	7	1.028
XIII	0.04	0.04	6	1.009
Ethics I	—	0.43	8	1.003
Ethics II	—	1.27	10	1.015
Ethics III	—	49.85	7	1.008
Ethics IV	—	2.45	22	1.005
Ethics V	—	0.44	16	1.001

Wall-clock times of single runs on the virtual machine described in the Methods.

Decomposition: the Sidney decomposition with dynamic programming inside each piece. Pieces: pieces of the finest Sidney decomposition. Heuristic

ratio: cost of the order that takes, inside each Sidney piece, the available proposition of largest weight, divided by the optimum, both measured above the smallest possible cost of the block. †: stopped at the time limit before certifying the optimum. —: not run.

Table 9: Wall-Clock Time, Its Spread, CPU Time And Peak Memory of Each Stage of the Pipeline.

Stage	Time (s)	Spread (s)	CPU (s)	Memory (MiB)
Reading the TEI file	0.060	0.002	0.061	183
Building the blocks	0.097	0.006	0.099	183
Null, all blocks	1.993	0.079	2.030	223
Ten thousand samples	5.149	0.009	5.241	222
Optimum, all blocks	3.281	0.093	3.308	204
Null in rationals, I	3.251	0.164	3.272	236
Null in floating point, I	0.380	0.016	0.387	197
Optimum, I	0.527	0.013	0.533	194
Integer programme, I	54.389	0.222	55.618	341
Whole-block lattices	4.327	0.139	4.406	281
Component lattices	1.352	0.029	1.377	207

Medians over at least 5 runs after a discarded warm-up, each in a fresh process with the numerical libraries restricted to one thread, on the virtual machine described in the Methods; spread: interquartile range; memory: peak resident memory of the process, interpreter included. I: Book I. Lattices: Algorithm 1 on the 11 blocks whose whole lattice of ideals can be enumerated, without and with the split into components.

Bounds. Bounds are not a substitute for the

$$\gamma_b(B) = \frac{T_{\min} - B}{E_0 - T_{\min}} \quad (26)$$

Verification. An independent re-implementation of the primary analysis, written from the plan without access to the pipeline's code, agreed with it in 100 of 100 comparisons at a relative tolerance of 10^{-9} , and its exact rational expectations were identical to the pipeline's in all 15 blocks. It computed the optimum by a different method, an integer programme over linear orderings, and it was itself checked against exhaustive enumeration on 40 random instances. A second run of the pipeline reproduced every value of the primary analysis.

4. DISCUSSION

The order of the Elements is economical in a strong statistical sense and far from optimal in an optimisation sense. No random admissible order

optimum. For a lower bound B on $T_{\min}$, the gap Equation (26) is the fraction by which B would enlarge the improvable range, so that replacing the optimum by B divides the index by $1 + \gamma_b(B)$. The relaxation of the integer programme leaves gaps of up to 0.58, and the combinatorial bound gaps between 0.56 and 2.83. For the whole text, using the relaxation instead of the optimum would lower the index from 0.364 to 0.299, and using the combinatorial bound would lower it to 0.145.

among the 100,000 sampled was as short as the text, yet the text captures just over a third of the reduction that reordering within the blocks could achieve.

Both statements are needed. The first shows that, with respect to citation distance, the order is not an arbitrary choice among the orders that the logic permits; the second shows that minimising that distance is not the principle that fixes it.

Because the index is exact and the threshold was fixed in advance, the conclusion does not rest on a choice made after seeing the data, and the robustness analyses place it well away from the threshold under every reading of the citations examined, as long as the objective counts every dependency. When it counts only the dependencies inside each block, the index rises to about one half, so the pre-registered

null is specific to the total dependency length and, by the criterion of the plan, not informative.

The economy that does exist is specific to the citations of the text. When each citation is redirected to a random earlier proposition of its block, the same order scores 0.069 on average instead of 0.364.

The order therefore does not look economical merely because proofs cite backwards; it places propositions near the particular results they use, only not as near as the dependencies would allow.

The exploratory analysis shows where the remaining economy lies and why the optimum looks different from an exposition.

The objective rewards stating a widely used result at the end of its book, next to the later books that use it. The text instead keeps such results with the propositions they belong to by subject: I.6 immediately follows I.5, of which it is the converse, and I.32, whose proof rests on I.29 and I.31, follows the propositions on parallel lines, although no later proposition of Book I cites either of them.

The gap between the text and the optimum is therefore consistent with an arrangement by subject, in which a result is stated where it belongs rather than where it will be used, and the index measures one consideration among several that an order of exposition balances. That the gains are unrelated to Heiberg's brackets and to Heath's notes on interpolation suggests that they reflect the arrangement itself rather than later alterations of the text. The two objectives of the multiverse differ precisely in this pull of the later books: counting only the dependencies inside each block removes it, and the index of the text then rises to about one half.

The economy is concentrated in a few blocks. Book I, in which the deductive organization of I.1-25 is very tight (Schiefsky, 2007), scores above the threshold, as do Book IX and the second group of Book X, whereas the third group of Book X and the solid geometry of Books XII and XIII score below chance.

The three groups of Book X span both extremes, and Knorr pointed to detectable seams in Book X (Knorr, 1975); the index does not identify seams, but it shows that the three groups are arranged very differently with respect to citation distance. This pattern is consistent with Dodgson's remark that many propositions have a partly arbitrary place (Dodgson, 1879): where the logic leaves room, the text does not always use it to shorten citations.

The replication on the Ethics gives a lower index, 0.353, than the 0.364 of the Elements. The two texts share a method of exposition, and the layout of the Ethics imitates Euclid (Steenbakkers, 2009), but the

Ethics is shorter, its parts are fewer and larger, and one of its components is large enough to need the Monte Carlo null, so the comparison is between two measurements of the same kind rather than a test of a common principle. An earlier computer analysis of the deductive structure of the Ethics (Wischik, 2000) examined its graph but not its order against an optimum.

The comparison with syntactic dependencies is suggestive but indirect. On a scale of the same form, half of the languages are optimized to 70% or more (Ferrer-i-Cancho et al., 2022), but that scale compares each sentence with all permutations of its words, whereas the present baseline must respect the logical dependencies, so the two values are not directly comparable. The index should also be read against the scale of the problem. Most of the total length is the constant K, fixed by the order of the books, and the text exceeds the minimum by only 5.4%, against 8.4% for a random admissible order. The index is a statement about how the freedom left by the logic is used, not about the length of the citations as such.

The analysis has limitations. The dependencies are Heath's: Heath's identifications of proposition use are far from reliable (Schiefsky, 2007), and different editors make different dependencies explicit. The multiverse and the random deletions vary how the citations are read, and every variant that counts all dependencies stays well below the threshold, but they cannot replace a new census of the proofs. The text is Heiberg's, as translated by Heath, and the number of propositions differs across textual traditions (Heath, 1908), although the alterations respected the global logical structure (Vitrac, 2004). The blocks are kept fixed, so the analysis measures the order within books and says nothing about the order of the books. The uniform distribution over admissible orders is one baseline among several possible ones, and weighting every dependency equally is a modelling choice.

The deviations from the plan, including the correction of the decomposition, are listed in the Methods.

From an engineering point of view, the exact computation is feasible although both problems are hard in general. The blocks split into weakly connected components with few order ideals, and the Sidney decomposition splits the optimum into small pieces, so the whole primary analysis, with exact rational arithmetic, runs in less than a minute on a desktop computer. The finest decomposition matters most when the objective counts only the dependencies inside each block: every weakly connected component then has ratio zero, all pieces

tie, and only the finest decomposition, whose pieces are the components themselves, keeps the dynamic programme small; the implementation that merged tied pieces had sent these specifications to integer programming. The same pipeline applies to any text whose units cite earlier units, from other deductive treatises to textbooks, where the admissible orders are the learning paths of the prerequisite structure.

5. CONCLUSIONS

Measured exactly and against a threshold fixed in advance, the order of Euclid's Elements is economical but not close to optimal. It is shorter than every one of 100,000 random orders that respect the logical dependencies, yet it recovers just over a third (0.364) of the reduction in total citation distance that the dependencies allow, below the pre-registered threshold of 0.5, and its economy is concentrated in a few books. Counting only the citations inside each

book, or each group of Book X, it recovers about half, between 0.49 and 0.54, so the pre-registered null holds for the total distance but not for the distances within those blocks. The largest departures from the optimum come from fundamental, widely cited results that the text states early, whereas minimising citation distance would move them to the end of their books. Minimising the distance between a proposition and the results it uses thus describes part of the arrangement of the Elements, not the whole of it. The economy that exists is specific to what the proofs cite: citations redirected at random to earlier propositions make the same order score far lower. The Ethics of Spinoza, measured in the same way, is also far more economical than chance and also below the threshold, with an index of 0.353. The formulation applies to any text whose units cite earlier units, and the pipeline computes both reference points exactly for texts of this size.

Declarations

Author Contributions: Luis Eduardo Muñoz Guerrero: conceptualization, methodology, software, validation, formal analysis, data curation, writing (original draft, review and editing) and visualization.

Ethics: Not applicable: the study analyses a published text and contains no data on living persons. The text is used under the licence stated in the data availability statement.

Data Availability: The data are public: the TEI file of Heath's translation in the repository canonical-greekLit of the Perseus Digital Library (licensed CC BY-SA 4.0) and Tredwell's dependency table of the Ethics in the repository Autio/Ethics (licensed GPL-2.0), each used at a fixed commit; a script downloads them and verifies their checksums.

Code Availability: The code and the frozen analysis plan are available from the author upon reasonable request.

Reporting guideline. No specific reporting guideline applies to this design. The report follows the analysis plan fixed before the analysis and states every deviation from it.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflicts of interest.

REFERENCES

- Alama, Jesse, Mamane, Lionel, and Urban, Josef (2012). Dependencies in Formal Mathematics: Applications and Extraction for Coq and Mizar. *Lecture Notes in Computer Science*, 1-16. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31374-5_1
- Atkinson, Jesse (2016). The Pythagorean Theorem as a Rooted In-tree Dependency Graph. *Proceedings of Bridges 2016: Mathematics, Music, Art, Architecture, Education, Culture*, 503-506. <https://archive.bridgesmathart.org/2016/bridges2016-503.html>
- Autio, Petri (2015). Ethics: Yet Another Graph Visualisation of the logical dependencies in Spinoza's Ethics (based on data compiled by R. F. Tredwell), commit 1028c55. GitHub repository Autio/Ethics. <https://github.com/Autio/Ethics>
- Avigad, Jeremy, Dean, Edward, and Mumma, John (2009). A formal system for Euclid's Elements. *The Review of Symbolic Logic* 2(4), 700-768. <https://doi.org/10.1017/S1755020309990098>

- Bansal, Nikhil and Khot, Subhash (2009). Optimal Long Code Test with One Free Bit. 2009 50th Annual IEEE Symposium on Foundations of Computer Science, 453-462. IEEE. <https://doi.org/10.1109/FOCS.2009.23>
- Beeson, Michael, Narboux, Julien, and Wiedijk, Freek (2019). Proof-checking Euclid. *Annals of Mathematics and Artificial Intelligence* 85(2-4), 213-257. <https://doi.org/10.1007/s10472-018-9606-x>
- Boman, Eugene, Milbrand, Alexandra, Brown, Tyler, Dahiya, Siddharth, Roberge, Joseph, and Boman, Mary (2014). Euclid21: Euclid's Elements for the 21st Century. *Convergence*. <https://doi.org/10.4169/convergence20141202>
- Brightwell, Graham and Winkler, Peter (1991). Counting linear extensions. *Order* 8(3), 225-242. <https://doi.org/10.1007/BF00383444>
- Correa, José R. and Schulz, Andreas S. (2005). Single-Machine Scheduling with Precedence Constraints. *Mathematics of Operations Research* 30(4), 1005-1021. <https://doi.org/10.1287/moor.1050.0158>
- Dodgson, Charles L. (1879). *Euclid and his Modern Rivals*. Macmillan and Co. <https://archive.org/details/euclidandhismode000469mbp>
- Doignon, Jean-Paul and Falmagne, Jean-Claude (2015). Knowledge Spaces and Learning Spaces. *arXiv preprint arXiv:1511.06757*. <https://arxiv.org/abs/1511.06757>
- Eppstein, David (2008). Learning Sequences. *arXiv preprint arXiv:0803.4030*. <https://arxiv.org/abs/0803.4030>
- Falmagne, Jean-Claude and Doignon, Jean-Paul (2011). *Learning Spaces*. Springer. <https://doi.org/10.1007/978-3-642-01039-2>
- Ferrer i Cancho, Ramon (2004). Euclidean distance between syntactically linked words. *Physical Review E* 70(5). <https://doi.org/10.1103/PhysRevE.70.056135>
- Ferrer-i-Cancho, Ramon and Liu, Haitao (2014). The risks of mixing dependency lengths from sequences of different length. *Glottology* 5(2), 143-155. <https://doi.org/10.1515/glot-2014-0014>
- Ferrer-i-Cancho, Ramon, Gómez-Rodríguez, Carlos, Esteban, Juan Luis, and Alemany-Puig, Lluís (2022). Optimality of syntactic dependency distances. *Physical Review E* 105(1). <https://doi.org/10.1103/PhysRevE.105.014308>
- Futrell, Richard, Mahowald, Kyle, and Gibson, Edward (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences* 112(33), 10336-10341. <https://doi.org/10.1073/pnas.1502134112>
- Gildea, Daniel and Temperley, David (2010). Do Grammars Minimize Dependency Length? *Cognitive Science* 34(2), 286-310. <https://doi.org/10.1111/j.1551-6709.2009.01073.x>
- Graham, R.L., Lawler, E.L., Lenstra, J.K., and Rinnooy Kan, A.H.G. (1979). Optimization and Approximation in Deterministic Sequencing and Scheduling: a Survey. *Annals of Discrete Mathematics* 5, 287-326. Elsevier. [https://doi.org/10.1016/S0167-5060\(08\)70356-X](https://doi.org/10.1016/S0167-5060(08)70356-X)
- Grötschel, Martin, Jünger, Michael, and Reinelt, Gerhard (1984). A Cutting Plane Algorithm for the Linear Ordering Problem. *Operations Research* 32(6), 1195-1220. <https://doi.org/10.1287/opre.32.6.1195>
- Heath, Thomas L. (1908). *The Thirteen Books of Euclid's Elements*, translated from the text of Heiberg with introduction and commentary (3 vols.). Cambridge University Press. https://archive.org/details/bub_gb_UhgPAAAAIAAJ
- Heiberg, Johan Ludvig (1883). *Euclidis Opera omnia*, vols. I-IV: *Elementa*. B. G. Teubner. <https://archive.org/details/euclidisoperaomn01eucl>
- Huangfu, Qi and Hall, J. A. J. (2018). Parallelizing the dual revised simplex method. *Mathematical Programming Computation* 10(1), 119-142. <https://doi.org/10.1007/s12532-017-0130-5>
- Huber, Mark (2006). Fast perfect sampling from linear extensions. *Discrete Mathematics* 306(4), 420-428. <https://doi.org/10.1016/j.disc.2006.01.003>
- Karzanov, Alexander and Khachiyan, Leonid (1991). On the conductance of order Markov chains. *Order* 8(1), 7-15. <https://doi.org/10.1007/BF00385809>
- Knorr, Wilbur Richard (1975). *The Evolution of the Euclidean Elements*. D. Reidel (Springer Netherlands). <https://doi.org/10.1007/978-94-010-1754-1>
- Margot, François, Queyranne, Maurice, and Wang, Yaoguang (2003). Decompositions, Network Flows, and a Precedence Constrained Single-Machine Scheduling Problem. *Operations Research* 51(6), 981-992. <https://doi.org/10.1287/opre.51.6.981.24912>

- Masucci, A.P. (2011). Formal versus self-organised knowledge systems: A network approach. *Physica A: Statistical Mechanics and its Applications* 390(23-24), 4652-4659. <https://doi.org/10.1016/j.physa.2011.06.074>
- Mondal, Debashis and Hinrichs, Nina (2016). Rank Tests from Partially Ordered Data Using Importance and MCMC Sampling Methods. *Statistical Science* 31(3). <https://doi.org/10.1214/16-STS549>
- Nadler, Steven (2006). *Spinoza's Ethics*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511815737>
- Neuenschwander, E. (1973). Die ersten vier Bücher der Elemente Euklids. *Archive for History of Exact Sciences* 9(4-5), 325-380. <https://doi.org/10.1007/BF00348367>
- Neuenschwander, E. (1974). Die stereometrischen Bücher der Elemente Euklids. *Archive for History of Exact Sciences* 14(2), 91-125. <https://doi.org/10.1007/BF00329825>
- Nicholls, Geoff K., Lee, Jeong Eun, Karn, Nicholas, Johnson, David, Huang, Rukuang, and Muir-Watt, Alexis (2022). Bayesian inference for partial orders from random linear extensions: power relations from 12th Century Royal Acta. arXiv preprint arXiv:2212.05524 (v1). <https://arxiv.org/abs/2212.05524v1>
- Perseus Digital Library (2023). canonical-greekLit: Euclid, Elementa (urn:cts:greekLit:tlg1799.tlg001), TEI of Heath's translation (perseus-eng1) and Heiberg's Greek text (perseus-grc2), commit 6451481. GitHub repository PerseusDL/canonical-greekLit. Tufts University. <https://github.com/PerseusDL/canonical-greekLit/tree/645148145584f04f4d2535e9c90e2f7373611908/data/tlg1799/tlg001>
- Phipson, Belinda and Smyth, Gordon K. (2010). Permutation P-values Should Never Be Zero: Calculating Exact P-values When Permutations Are Randomly Drawn. *Statistical Applications in Genetics and Molecular Biology* 9(1). <https://doi.org/10.2202/1544-6115.1585>
- Potts, C. N. (1980). An algorithm for the single machine sequencing problem with precedence constraints. *Mathematical Programming Studies* 13, 78-87. Springer Berlin Heidelberg. <https://doi.org/10.1007/BFb0120909>
- Propp, James Gary and Wilson, David Bruce (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms* 9(1-2), 223-252. [https://doi.org/10.1002/\(SICI\)1098-2418\(199608/09\)9:1/2<223::AID-RSA14>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1098-2418(199608/09)9:1/2<223::AID-RSA14>3.0.CO;2-O)
- Pał, Karol (2014). Improving legibility of natural deduction proofs is not trivial. *Logical Methods in Computer Science*. [https://doi.org/10.2168/LMCS-10\(3:23\)2014](https://doi.org/10.2168/LMCS-10(3:23)2014)
- Pał, Karol (2015). Improving Legibility of Formal Proofs Based on the Close Reference Principle is NP-Hard. *Journal of Automated Reasoning* 55(3), 295-306. <https://doi.org/10.1007/s10817-015-9337-1>
- Schiefsky, Mark J. (2007). New technologies for the study of Euclid's Elements. Archimedes Project, Harvard University (working paper). http://web.archive.org/web/20230307051713/http://archimedes.fas.harvard.edu/euclid/euclid_paper.pdf
- Shiloach, Yossi (1979). A Minimum Linear Arrangement Algorithm for Undirected Trees. *SIAM Journal on Computing* 8(1), 15-32. <https://doi.org/10.1137/0208002>
- Sidney, Jeffrey B. (1975). Decomposition Algorithms for Single-Machine Sequencing with Precedence Relations and Deferral Costs. *Operations Research* 23(2), 283-298. <https://doi.org/10.1287/opre.23.2.283>
- Sidney, Jeffrey B. and Steiner, George (1986). Optimal Sequencing by Modular Decomposition: Polynomial Algorithms. *Operations Research* 34(4), 606-612. <https://doi.org/10.1287/opre.34.4.606>
- Steenbakkers, Piet (2009). The Geometrical Order in the Ethics. *The Cambridge Companion to Spinoza's Ethics*, 42-55. Cambridge University Press. <https://doi.org/10.1017/CCOL9780521853392.003>
- Talvitie, Topi, Niinimäki, Teppo, and Koivisto, Mikko (2017). The Mixing of Markov Chains on Linear Extensions in Practice. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, 524-530. International Joint Conferences on Artificial Intelligence Organization. <https://doi.org/10.24963/ijcai.2017/74>
- Temperley, David and Gildea, Daniel (2018). Minimizing Syntactic Dependency Lengths: Typological/Cognitive Universal? *Annual Review of Linguistics* 4(1), 67-80. <https://doi.org/10.1146/annurev-linguistics-011817-045617>

- Tredwell, R. F. (2009). Acyclic Directed Graph of Spinoza's Ethics. *Studia Spinoziana Publication Services* (discussion paper). <https://capone.mtsu.edu/rbombard/RB/Spinoza/tredwell.html>
- Viteri, Scott and DeDeo, Simon (2022). Epistemic phase transitions in mathematical proofs. *Cognition* 225, 105120. <https://doi.org/10.1016/j.cognition.2022.105120>
- Vitrac, Bernard (2004). À propos des démonstrations alternatives et autres substitutions de preuves dans les *Éléments* d'Euclide. *Archive for History of Exact Sciences* 59(1), 1-44. <https://doi.org/10.1007/s00407-004-0087-4>
- Wilson, David Bruce (2004). Mixing times of lozenge tiling and card shuffling Markov chains. *The Annals of Applied Probability* 14(1). <https://doi.org/10.1214/aoap/1075828054>
- Wischik, Lucian (2000). The shape of deduction in Spinoza's "Ethics". Unpublished essay (PDF on the website of J. Martin, University of Cincinnati). https://homepages.uc.edu/~martinj/Spinoza_&_Hobbes/Secondary_Literature/Deduction%20in%20Spinoza's%20Ethics.pdf
- Woeginger, Gerhard J (2003). On the approximability of average completion time scheduling under precedence constraints. *Discrete Applied Mathematics* 131(1), 237-252. [https://doi.org/10.1016/S0166-218X\(02\)00427-4](https://doi.org/10.1016/S0166-218X(02)00427-4)
- Wolfram, Stephen (2002). *A New Kind of Science*. 1176 pp. Wolfram Media. <https://www.wolframscience.com/nks/notes-12-9--empirical-metamathematics/>
- Wolfram, Stephen (2021). The Empirical Metamathematics of Euclid and Beyond. arXiv preprint arXiv:2107.07337. <https://arxiv.org/abs/2107.07337>