

EXTREMAL DESCENDANT INTEGRALS ON MODULI SPACES OF CURVES: AN INEQUALITY DISCOVERED AND PROVED IN COLLABORATION WITH AI

JOHANNES SCHMITT

ABSTRACT.

For the pure ψ -class intersection numbers $\langle \tau_{e_1} \cdots \tau_{e_n} \rangle_g$ on the moduli space $\overline{\mathcal{M}}_{g,n}$ of stable curves, we determine for which choices of e_1, \dots, e_n their value becomes extremal. The intersection number is minimal for powers of a single ψ -class (i.e. all e_i but one vanish), whereas maximal values are obtained for balanced vectors ($|e_i - e_j| \leq 1$ for all i, j). The proof uses the nefness of the ψ -classes combined with Khovanskii–Teissier log-concavity.

AUTHOR'S NOTE.

The question of finding extremal values of the ψ -intersection numbers first occurred to the author when looking for a toy problem to explore using the software OpenEvolve [Sha25]. The conjecture that balanced exponents lead to the maximal values is a natural guess, and was indeed discovered quickly by the tested model. To the author's knowledge, this optimization-style problem was novel and not covered by existing literature: it is a simple and natural question, but somewhat orthogonal to the questions usually studied in enumerative geometry. After some experimental verification and presenting the conjecture to several colleagues (who confirmed its open status), it was submitted as a problem to the IMPProofBench project [SBD⁺25]. This project collects research level mathematics questions and tests them against a range of AI models. As part of this evaluation, the conjecture was independently proven by several such models, without human intervention (see Appendix A for further details).

The proof itself turns out to be unexpectedly simple, bypassing the rich structural results on descendant invariants on $\overline{\mathcal{M}}_{g,n}$, and instead using just a few basic properties of ψ -classes (nefness and symmetry) together with general results on intersection numbers of divisor classes. As such, while the obtained theorem is a neat little result and original contribution to the literature, it would arguably be on the borderline of notability for a mathematical publication.

We still opted to use this opportunity and write the present note, both to present the result but in particular also to document the process that went into its discovery and subsequent write-up. Apart from showcasing the current level of sophistication of AI for these tasks, we also propose some best practices for attribution of AI use in mathematical writing (see Appendix B). We apply these principles in the current paper, e.g. by marking human-generated text with a blue bar or underlines, whereas AI-generated text is marked in brown. After an initial draft of the paper was completed, we also added an alternative version of the proof of the main theorem (Section 3), partially verified in Lean [MU21]. The argument is split into a purely combinatorial optimization theorem, formalized by Claude Code and GPT 5.2 with minimal guidance by the author, and a geometric proposition proving the applicability of the optimization result to the intersection numbers above.

1. INTRODUCTION

AI Let $g, n \in \mathbb{Z}_{\geq 0}$ with $2g - 2 + n > 0$. The moduli space $\overline{\mathcal{M}}_{g,n}$ of stable curves of genus g with n marked points has dimension $d = 3g - 3 + n$. For each marked point i , the cotangent line bundle \mathbb{L}_i at that marking defines the ψ -class $\psi_i = c_1(\mathbb{L}_i) \in H^2(\overline{\mathcal{M}}_{g,n}, \mathbb{Q})$. The *descendant integrals* (or *intersection numbers of ψ -classes*)

$$\langle \tau_{e_1} \cdots \tau_{e_n} \rangle_g := \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{e_1} \cdots \psi_n^{e_n}$$

are rational numbers that vanish unless $\sum_{i=1}^n e_i = d$. These integrals play a central role in the Witten–Kontsevich theorem [Wit91, Kon92], which establishes that their generating function is a τ -function for the KdV hierarchy.

Given g and n , let

$$E(g, n) = \left\{ \mathbf{e} = (e_1, \dots, e_n) \in \mathbb{Z}_{\geq 0}^n : \sum_{j=1}^n e_j = 3g - 3 + n \right\}$$

be the set of valid exponent vectors. Define the function $D : E(g, n) \rightarrow \mathbb{Q}$ by $D(\mathbf{e}) = \langle \tau_{e_1} \cdots \tau_{e_n} \rangle_g$.

Definition 1.1. A vector $\mathbf{e} \in E(g, n)$ is *balanced* if $|e_i - e_j| \leq 1$ for all $1 \leq i, j \leq n$.

Equivalently, writing $3g - 3 + n = an + b$ with $0 \leq b < n$ via Euclidean division, the balanced vectors are precisely the permutations of $(a, \dots, a, a+1, \dots, a+1)$ with a appearing $n-b$ times and $a+1$ appearing b times.

Theorem 1.2 (Extremal descendants). *Let $g \in \mathbb{Z}_{\geq 0}$ and $n \in \mathbb{Z}_{\geq 0}$ with $2g - 2 + n > 0$.*

(a) (Minimum) *The function D achieves its minimum at the concentrated vector $(3g - 3 + n, 0, \dots, 0)$ (or any permutation thereof), with value*

$$\langle \tau_{3g-3+n} \cdot \underbrace{\tau_0^{n-1}}_{\text{in (1)}} \rangle_g = \frac{1}{24^g g!}. \quad (1)$$

(b) (Maximum) *The function D achieves its maximum on a balanced vector $\mathbf{e} \in E(g, n)$.*

The [formula for the minimal value](#) in part (a) is well-known ([see e.g. equation \(5.36\) in \[IZ92\]](#)); it also follows from the generating function identity

$$1 + \sum_{g \geq 1} t^g \int_{\overline{\mathcal{M}}_{g,1}} \psi_1^{3g-2} = \exp(t/24) \quad (2)$$

established in [FP00]. Part (b) appears to be new.

In genus zero, part (b) follows immediately from the closed formula

$$\langle \tau_{e_1} \cdots \tau_{e_n} \rangle_0 = \frac{(n-3)!}{e_1! \cdots e_n!}, \quad (3)$$

valid when $\sum e_i = n - 3$. The multinomial coefficient on the right is maximized precisely when the e_i are as equal as possible. For higher genus, no such closed formula exists, and the result requires the more sophisticated argument given in Section 2.

Remark 1.3. The [result above was also](#) verified computationally for all $0 \leq g \leq 11$ and $1 \leq n \leq 11$ using the SageMath [The24] software package `admcycles` [DSvZ21].

Remark 1.4. Uniqueness need not hold: due to symmetry, all permutations of a balanced vector achieve the same value, and in some cases plateaus may occur where strict inequality fails.

Given the simple formula (1) for the minimal descendant, a natural question is:

Question 1.5. For $3g - 3 + n = an + b$, does there exist a simple description (closed formula, generating function, recursion, ...) for the balanced descendant invariant

$$\langle \tau_a^{n-b} \tau_{a+1}^b \rangle_g = \max_{\mathbf{e} \in \underline{E}(g, n)} D(\mathbf{e}) ? \quad (4)$$

A particularly simple regime might be the case $n \geq 3g - 3$, where $a = 1$ and the dilaton equation allows to reduce all invariants to

$$\langle \tau_1^{n-(3g-3)} \tau_2^{3g-3} \rangle_g = \frac{(2g-3+n)!}{(5g-6)!} \langle \tau_2^{3g-3} \rangle_g. \quad (5)$$

Questions of extremizing enumerative invariants within a finite parameter space are somewhat unusual in enumerative geometry, where one more typically studies generating functions or asymptotic behavior. It would be interesting to investigate analogous extremal questions for other families of intersection numbers, such as Hodge integrals or double Hurwitz numbers.

Outline. In Section 2 we present the proof of Theorem 1.2, as it was found and stated by GPT-5 (for part (b)) and Gemini 3 Pro (for part (a)). We chose to leave the formulation largely untouched to allow the reader to see (for better or worse) the current level of proficiency in proof generation exhibited by these models.

Conversely, section 3 presents an alternative argument for Theorem 1.2, which can be read independently of Section 2. The treatment there splits the proof into:

- a purely combinatorial optimization result (Theorem 3.1) which has been formalized in Lean [MU21] based on the mathlib library [mC20],
- a geometric argument (Proposition 3.2), with proof entirely human-written, proving that the descendant function D from Theorem 1.2 satisfies the assumptions of this optimization result.

In Appendix A we describe the various forms of AI assistance that went into the preparation of the present article. In Appendix B we collect some proposals and tools for attributing AI contributions in mathematical papers.

Acknowledgments. We want to thank Javier Gómez-Serrano and Jaume De Dios Pont for inspiring and organizing, respectively, the OpenEvolve hackathon at which the present project was initiated. We are grateful to our colleagues who gave comments on Theorem 1.2, in particular for confirming that to their knowledge it is not already covered in the literature. The proposals in Appendix B were inspired by both a [public discussion thread started by Terry Tao](#) and conversations with colleagues, such as Jasper Dekoninck, Jeremy Feusi, Tim Gehrunger as well as the members of the writing group of the [Leiden declaration on the use of AI in mathematical research](#), in particular Jarod Alper, Michael Harris, David Holmes, Ursula Martin and Jim Portegies, whom we thank warmly!

Finally, we are grateful to Jarod Alper, Honglu Fan, Jeremy Feusi, David Holmes and Jim Portegies for comments on various drafts of the present article.

The author was supported by SwissMAP and the IMPProofBench project received financial support from Google/DeepMind as part of their Google Gift program.

2. PROOF OF THE MAIN THEOREM

AI We write $d = \dim \overline{\mathcal{M}}_{g,n} = 3g - 3 + n$ throughout.

2.1. Maximal descendants. The following proof was generated by OpenAI's GPT-5 model. It is reproduced with only minimal formatting changes; see Appendix A for context. Note: The original AI output contained an error in the genus-0 formula stated in the remarks, giving $(n - 3)!/\prod_i(2e_i - 1)!!$ instead of the correct formula (3). This has been corrected.

Proof of Theorem 1.2 (b). We prove that among all \mathbf{e} with fixed sum, $D(\mathbf{e})$ attains its maximum at a balanced vector, i.e., when the entries differ by at most 1.

Step 1 (Two-point slice and its basic properties). Fix distinct indices $i \neq j$ and set

$$M := \prod_{k \neq i,j} \psi_k^{e_k}, \quad q := e_i + e_j.$$

For $t = 0, 1, \dots, q$, consider the one-variable sequence

$$S_t := \int_{\overline{\mathcal{M}}_{g,n}} \psi_i^t \psi_j^{q-t} M.$$

By the natural action of the symmetric group permuting the markings on $\overline{\mathcal{M}}_{g,n}$, the classes ψ_i and ψ_j are exchanged by an automorphism that leaves M invariant; hence

$$S_t = S_{q-t} \quad (\text{palindromicity}).$$

Step 2 (Khovanskii–Teissier log-concavity). It is standard that each ψ_i is nef on $\overline{\mathcal{M}}_{g,n}$. For nef classes A, B and fixed nef classes H_3, \dots, H_d on a projective (orbifold) variety of dimension d , the Khovanskii–Teissier inequalities as described in Variant 1.6.2 of [Laz04] (equivalently, the mixed Hodge–Riemann bilinear relations) give the discrete log-concavity

$$\left(\int A^t B^{q-t} H_3 \cdots H_d \right)^2 \geq \left(\int A^{t-1} B^{q-t+1} H_3 \cdots H_d \right) \left(\int A^{t+1} B^{q-t-1} H_3 \cdots H_d \right)$$

for all $1 \leq t \leq q - 1$. Applying this with $A = \psi_i$, $B = \psi_j$, $H_3 \cdots H_d = M$, we obtain

$$S_t^2 \geq S_{t-1} S_{t+1} \quad (1 \leq t \leq q - 1) \quad (\text{log-concavity}).$$

Step 3 (Monotonicity towards the middle). For a positive log-concave sequence, the ratios $R_t := S_{t+1}/S_t$ are weakly decreasing in t . Using palindromicity,

$$R_{q-t-1} = \frac{S_{q-t}}{S_{q-t-1}} = \frac{S_t}{S_{t+1}} = \frac{1}{R_t}.$$

Hence for $t \leq \lfloor q/2 \rfloor - 1$ we have $R_t \geq R_{q-t-1} = 1/R_t$, so $R_t \geq 1$, i.e.,

$$S_{t+1} \geq S_t \quad \text{for } t < \frac{q}{2}, \quad S_{t-1} \geq S_t \quad \text{for } t > \frac{q}{2}.$$

In words: along the two-point slice $\{(t, q-t) : 0 \leq t \leq q\}$, the values increase up to the middle and then decrease.

Step 4 (The balancing step). Suppose $e_i \geq e_j + 2$. Then $e_i > q/2$. By the monotonicity just proved,

$$D(e_1, \dots, e_i, \dots, e_j, \dots, e_n) = S_{e_i} \leq S_{e_i-1} = D(e_1, \dots, e_i-1, \dots, e_j+1, \dots, e_n).$$

Thus, whenever two entries differ by at least 2, transferring one unit from the larger to the smaller weakly increases the value of the integral.

Step 5 (Conclusion by iteration). Starting from any $\mathbf{e} \in E(g, n)$, repeatedly apply the balancing step to any pair (i, j) with $|e_i - e_j| \geq 2$. This process terminates at a balanced vector \mathbf{e}^* (all entries differ by at most 1), and along the way the value of D never decreases. Hence

$$D(\mathbf{e}) \leq D(\mathbf{e}^*) \quad \text{for some balanced } \mathbf{e}^* \in E(g, n).$$

In particular, the maximum of D on $E(g, n)$ is achieved at a balanced vector. \square

Remark 2.1. The log-concavity in Step 2 is a special case of the Khovanskii–Teissier (or Alexandrov–Fenchel) inequalities for nef classes; it can be proved by restricting to a general complete intersection surface and applying the Hodge index theorem, or via the mixed Hodge–Riemann bilinear relations. The reduction from the Deligne–Mumford stack to a smooth projective variety can be made using a finite level-structure cover; intersection numbers scale by the degree of the cover, so inequalities are preserved.

2.2. **Minimal descendants.** The minimum claim follows from the observation that iteratively concentrating exponents on a single marking yields a non-increasing sequence of descendant integrals, combined with the known formula for one-point integrals.

Proof. We employ the logic of Part (b) in reverse. Let $\mathbf{e} = (e_1, \dots, e_n) \in E(g, n)$ be an arbitrary exponent vector. We wish to show that $D(\mathbf{e})$ is bounded below by the value of the concentrated vector.

Let k be an index such that $e_k = \max_{1 \leq i \leq n} e_i$. If \mathbf{e} is not fully concentrated on the k -th marking, there exists an index $j \neq k$ such that $e_j > 0$. Consider the pair of indices (k, j) and let $q = e_k + e_j$. Since e_k is maximal in \mathbf{e} , we have $e_k \geq e_j$, which implies $e_k \geq q/2$.

Recall from Step 3 of the proof of part (b) that the sequence S_t (representing the integral as a function of the exponent at the k -th marking) is non-increasing for $t \geq q/2$. Consequently, shifting the entire weight from the j -th marking to the k -th marking does not increase the value of the invariant:

$$D(\dots, e_k, \dots, e_j, \dots) \geq D(\dots, e_k + e_j, \dots, 0, \dots).$$

By iteratively applying this procedure for all $j \neq k$ with $e_j > 0$, we transform \mathbf{e} into the concentrated vector

$$\mathbf{e}_{\min} = (\underbrace{0, \dots, 0, d, 0, \dots, 0}_{\text{at index } k},$$

(where the weight $d = 3g - 3 + n$ is at index k) while ensuring the value of the integral never increases. Thus $D(\mathbf{e}) \geq D(\mathbf{e}_{\min})$.

It remains to calculate the value of this minimum. The integral is given by $\langle \tau_d \tau_0^{n-1} \rangle_g$, where τ_0 denotes the fundamental class. We apply the string equation

$$\langle \tau_0 \prod_{i=1}^m \tau_{a_i} \rangle_g = \sum_{j=1}^m \langle \tau_{a_j-1} \prod_{i \neq j} \tau_{a_i} \rangle_g$$

repeatedly to remove the $n - 1$ insertions of τ_0 . This works for $g \geq 1$; for $g = 0$ we only remove $n - 3$ such insertions, to avoid falling into the unstable range $n \leq 2$. In each step, the term on the right-hand side vanishes unless we act on the single non-zero exponent (since τ_{-1} is zero). We therefore obtain the reduction

$$\langle \tau_{3g-3+n} \tau_0^{n-1} \rangle_g = \langle \tau_{3g-4+n} \tau_0^{n-2} \rangle_g = \cdots = \langle \tau_{3g-2} \rangle_g.$$

The value of the one-point integral $\langle \tau_{3g-2} \rangle_g = 1/(24^g g!)$ is a classical result following from the Witten–Kontsevich theorem (specifically, the coefficient of t^g in the partition function restricted to one marked point). See in particular the discussion and references around equation (2). \square

3. FORMALIZED PROOF VIA AN ABSTRACT OPTIMIZATION THEOREM

AI This section presents an alternative approach to proving Theorem 1.2, based on a Lean 4 formalization available at [the project blueprint](#). The key insight is to separate the *abstract optimization argument*—which applies to any function satisfying certain axioms—from the *geometric input* specific to ψ -class intersection numbers. This modular structure clarifies the logical dependencies and yields a fully machine-verified proof of the combinatorial core. Throughout this section, we work with *weak compositions*: for integers $n \geq 1$ and $d \geq 0$, define

$$E(n, d) = \left\{ \mathbf{e} = (e_1, \dots, e_n) \in \mathbb{Z}_{\geq 0}^n : \sum_{j=1}^n e_j = d \right\}.$$

For $\mathbf{e} \in E(n, d)$ and distinct indices i, j , we write $\mathbf{e} - \delta_i + \delta_j$ for the vector obtained by decreasing e_i by 1 and increasing e_j by 1 (defined when $e_i \geq 1$).

Theorem 3.1 (Optimization Theorem -). *Let $D : E(n, d) \rightarrow \mathbb{Q}$ be a function satisfying:*

- (S) **Symmetry:** $D(\mathbf{e} \circ \sigma) = D(\mathbf{e})$ for all permutations $\sigma \in S_n$.
- (LC) **Log-concavity:** For all $\mathbf{e} \in E(n, d)$ and distinct i, j with $e_i, e_j \geq 1$,

$$D(\mathbf{e})^2 \geq D(\mathbf{e} - \delta_i + \delta_j) \cdot D(\mathbf{e} + \delta_i - \delta_j).$$

- (P) **Strict positivity:** $D(\mathbf{e}) > 0$ for all $\mathbf{e} \in E(n, d)$.

Then:

- (a) D achieves its maximum on a balanced vector (where $|e_i - e_j| \leq 1$ for all i, j).
- (b) D achieves its minimum on a concentrated vector (where $\mathbf{e} = d \cdot \delta_k$ for some k).

The proof is presented in [Sections 3.1 to 3.4](#). Assuming the theorem for now, we verify that the descendant integral satisfies the hypotheses [above](#).

Proposition 3.2. *Let $g, n \in \mathbb{Z}_{\geq 0}$ with $2g - 2 + n > 0$ and $n \geq 1$. The function*

$$D : E(n, 3g - 3 + n) \rightarrow \mathbb{Q}, \quad D(\mathbf{e}) = \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{e_1} \cdots \psi_n^{e_n}$$

satisfies conditions (S), (LC), and (P) of Theorem 3.1.

Proof. For the symmetry (S), we observe that for any $\sigma \in S_n$ with inverse $\tau = \sigma^{-1}$ there exists a natural isomorphism

$$\phi_\sigma : \overline{\mathcal{M}}_{g,n} \rightarrow \overline{\mathcal{M}}_{g,n}, (C, p_1, \dots, p_n) \mapsto (C, p_{\sigma(1)}, \dots, p_{\sigma(n)}), \quad (6)$$

which satisfies $\phi_\sigma^* \psi_i = \psi_{\tau(i)}$. Using the projection formula together with the pushforward relation $(\phi_\sigma)_* [\overline{\mathcal{M}}_{g,n}] = [\overline{\mathcal{M}}_{g,n}]$ of the fundamental class, we obtain:

$$\begin{aligned} D(\mathbf{e} \circ \sigma) &= \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{e_{\sigma(1)}} \cdots \psi_n^{e_{\sigma(n)}} = \int_{\overline{\mathcal{M}}_{g,n}} \psi_{\tau(1)}^{e_1} \cdots \psi_{\tau(n)}^{e_n} \\ &= \deg \left(\psi_{\tau(1)}^{e_1} \cdots \psi_{\tau(n)}^{e_n} \frown [\overline{\mathcal{M}}_{g,n}] \right) \\ &= \deg \left((\phi_\sigma)^* (\psi_1^{e_1} \cdots \psi_n^{e_n}) \frown (\phi_\sigma)_* [\overline{\mathcal{M}}_{g,n}] \right) \\ &= \deg \left((\phi_\sigma)_* (\psi_1^{e_1} \cdots \psi_n^{e_n} \frown [\overline{\mathcal{M}}_{g,n}]) \right) \\ &= \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{e_1} \cdots \psi_n^{e_n} = D(\mathbf{e}). \end{aligned}$$

To prove property (LC) we first recall from Variant 1.6.2 of [Laz04] that for X an irreducible complete scheme of dimension $d \geq 2$ and nef divisor classes

$$\alpha_1, \alpha_2, \beta_1, \dots, \beta_{d-2} \in N_1(X)_{\mathbb{R}}$$

we have

$$(\alpha_1 \cdot \alpha_2 \cdot \beta_1 \cdots \beta_{d-2})^2 \geq (\alpha_1^2 \cdot \beta_1 \cdots \beta_{d-2}) \cdot (\alpha_2^2 \cdot \beta_1 \cdots \beta_{d-2}). \quad (7)$$

Now fix $d = 3g - 3 + n$ and $\mathbf{e} \in E(n, d)$ and distinct i, j with $e_i, e_j \geq 1$, and choose $\tilde{\alpha}_1 = \psi_j$, $\tilde{\alpha}_2 = \psi_i$ and $\tilde{\beta}_1, \dots, \tilde{\beta}_{d-2} \in \{\psi_1, \dots, \psi_n\}$ such that

$$\psi_1^{e_1} \cdots \psi_n^{e_n} = \tilde{\alpha}_1 \cdot \tilde{\alpha}_2 \cdot \tilde{\beta}_1 \cdots \tilde{\beta}_{d-2}.$$

Note that this step is just a reordering of factors, using that $e_i, e_j \geq 1$ to isolate the two factors $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$. Let $\varphi : X \rightarrow \overline{\mathcal{M}}_{g,n}$ be a finite flat cover of degree M by an irreducible complete scheme X (see e.g. [ACV03, Theorem 7.4.2]). Write $\alpha_i = \varphi^* \tilde{\alpha}_i$ and $\beta_i = \varphi^* \tilde{\beta}_i$. Since all ψ -classes are nef ([ACG11, Chapter XIV, Corollary 5.14]) and this property is preserved under pullback by φ , we can apply (7) together with the projection formula and obtain:

$$\begin{aligned} D(\mathbf{e})^2 &= \left(\int_{\overline{\mathcal{M}}_{g,n}} \tilde{\alpha}_1 \cdot \tilde{\alpha}_2 \cdot \tilde{\beta}_1 \cdots \tilde{\beta}_{d-2} \right)^2 = \frac{1}{M^2} \left(\int_X \alpha_1 \cdot \alpha_2 \cdot \beta_1 \cdots \beta_{d-2} \right)^2 \\ &\geq \left(\frac{1}{M} \int_X \alpha_1^2 \cdot \beta_1 \cdots \beta_{d-2} \right) \cdot \left(\frac{1}{M} \int_X \alpha_2^2 \cdot \beta_1 \cdots \beta_{d-2} \right) \\ &= \left(\int_{\overline{\mathcal{M}}_{g,n}} \tilde{\alpha}_1^2 \cdot \tilde{\beta}_1 \cdots \tilde{\beta}_{d-2} \right) \cdot \left(\int_{\overline{\mathcal{M}}_{g,n}} \tilde{\alpha}_2^2 \cdot \tilde{\beta}_1 \cdots \tilde{\beta}_{d-2} \right) \\ &= D(\mathbf{e} - \delta_i + \delta_j) \cdot D(\mathbf{e} + \delta_i - \delta_j). \end{aligned}$$

Finally, the positivity of all descendant invariants $D(\mathbf{e})$ is well known, following e.g. from the recursive formulas for computing these invariants derived from the Witten–Kontsevich theorem [Wit91, Kon92]. \square

Remark 3.3. Theorem 1.2(a) and (b) follow immediately from Theorem 3.1 and Proposition 3.2. The formula (1) for the minimal value is then obtained by the string equation argument given in Section 2.

3.1. Slice sequences. The proof of Theorem 3.1 proceeds by analyzing how D varies along “slices” where only two coordinates change.

Definition 3.4 (Slice sequence - **LEAN**). Let $D : E(n, d) \rightarrow \mathbb{Q}_{>0}$ satisfy (S), (LC), (P). Given $\mathbf{e} \in E(n, d)$ and distinct indices $i \neq j$, define the *slice sequence* $S : \{0, 1, \dots, q\} \rightarrow \mathbb{Q}_{>0}$ by

$$S_t = D(\mathbf{e}^{(t)}),$$

where $q = e_i + e_j$ and $\mathbf{e}^{(t)}$ is the composition agreeing with \mathbf{e} at all indices $k \notin \{i, j\}$, with $e_i^{(t)} = t$ and $e_j^{(t)} = q - t$.

The key observation is that S inherits strong structural properties from D .

Lemma 3.5 (Palindromicity - **W**, **LEAN**). *The slice sequence satisfies $S_t = S_{q-t}$ for all $0 \leq t \leq q$.*

Proof. The transposition $\sigma = (i \ j) \in S_n$ exchanges the i -th and j -th coordinates. Since $\mathbf{e}^{(t)} \circ \sigma = \mathbf{e}^{(q-t)}$, symmetry (S) gives $S_t = D(\mathbf{e}^{(t)}) = D(\mathbf{e}^{(q-t)}) = S_{q-t}$. \square

Lemma 3.6 (Log-concavity - **W**, **LEAN**). *The slice sequence satisfies $S_t^2 \geq S_{t-1} \cdot S_{t+1}$ for all $1 \leq t \leq q-1$.*

Proof. For $1 \leq t \leq q-1$, the composition $\mathbf{e}^{(t)}$ has $e_i^{(t)} = t \geq 1$ and $e_j^{(t)} = q - t \geq 1$. Applying condition (LC) to $\mathbf{e}^{(t)}$ with indices i, j yields

$$D(\mathbf{e}^{(t)})^2 \geq D(\mathbf{e}^{(t)} - \delta_i + \delta_j) \cdot D(\mathbf{e}^{(t)} + \delta_i - \delta_j) = D(\mathbf{e}^{(t-1)}) \cdot D(\mathbf{e}^{(t+1)}),$$

which is exactly $S_t^2 \geq S_{t-1} \cdot S_{t+1}$. \square

3.2. Unimodality of log-concave palindromic sequences. The following general lemma is the heart of the argument.

Lemma 3.7 (Unimodality - **W**, **LEAN**). *Let $S : \{0, 1, \dots, q\} \rightarrow \mathbb{Q}_{>0}$ be a positive sequence satisfying:*

- *Palindromicity: $S_t = S_{q-t}$ for all $0 \leq t \leq q$.*
- *Log-concavity: $S_t^2 \geq S_{t-1} \cdot S_{t+1}$ for all $1 \leq t \leq q-1$.*

Then S is unimodal with maximum at the center:

- (i) $S_t \leq S_{t+1}$ whenever $2t < q$.
- (ii) $S_t \leq S_{t-1}$ whenever $2t > q$.

Proof. Define the ratio $R_t = S_{t+1}/S_t$ for $0 \leq t \leq q-1$. Log-concavity implies that the ratios are weakly decreasing: from $S_t^2 \geq S_{t-1}S_{t+1}$ we obtain $S_t/S_{t-1} \geq S_{t+1}/S_t$, i.e., $R_{t-1} \geq R_t$.

Palindromicity gives a reflection identity for ratios:

$$R_{q-1-t} = \frac{S_{q-t}}{S_{q-1-t}} = \frac{S_t}{S_{t+1}} = \frac{1}{R_t}.$$

For t with $2t < q - 1$, monotonicity of ratios gives $R_t \geq R_{q-1-t} = 1/R_t$, hence $R_t^2 \geq 1$. Since $R_t > 0$, we conclude $R_t \geq 1$, i.e., $S_{t+1} \geq S_t$. This proves (i).

Part (ii) follows from (i) by palindromicity: if $2t > q$, then $2(q-t) < q$, so $S_{q-t} \leq S_{q-t+1}$, which by palindromicity becomes $S_t \leq S_{t-1}$. \square

3.3. The balancing step.

Lemma 3.8 (Balancing increases D - **✓, LEAN**). *Let D satisfy (S), (LC), (P). If $\mathbf{e} \in E(n, d)$ has $e_i \geq e_j + 2$ for some $i \neq j$, then*

$$D(\mathbf{e}) \leq D(\mathbf{e} - \delta_i + \delta_j).$$

Proof. Consider the slice sequence S for \mathbf{e} with indices i, j . The original composition \mathbf{e} corresponds to S_{e_i} , and the modified composition $\mathbf{e} - \delta_i + \delta_j$ corresponds to S_{e_i-1} .

Since $e_i \geq e_j + 2$, we have $e_i > (e_i + e_j)/2 = q/2$, so $2e_i > q$. By the unimodality lemma (part (ii)), $S_{e_i} \leq S_{e_i-1}$, which gives $D(\mathbf{e}) \leq D(\mathbf{e} - \delta_i + \delta_j)$. \square

Proof of Theorem 3.1(a). Starting from any $\mathbf{e} \in E(n, d)$, repeatedly apply the balancing step: whenever there exist i, j with $e_i \geq e_j + 2$, replace \mathbf{e} by $\mathbf{e} - \delta_i + \delta_j$. By Lemma 3.8, each step weakly increases D .

This process terminates in finitely many steps because the *imbalance* $\sum_k e_k^2$ strictly decreases with each modification (one verifies $(a-1)^2 + (b+1)^2 < a^2 + b^2$ when $a \geq b+2$). The process terminates precisely when no pair (i, j) satisfies $|e_i - e_j| \geq 2$, i.e., when \mathbf{e} is balanced.

Thus every \mathbf{e} can be transformed to a balanced \mathbf{e}^* with $D(\mathbf{e}) \leq D(\mathbf{e}^*)$. In particular, the maximum of D over the finite set $E(n, d)$ is achieved at a balanced vector. \square

3.4. The concentrating step.

Lemma 3.9 (Concentrating decreases D - **✓, LEAN**). *Let D satisfy (S), (LC), (P). If $\mathbf{e} \in E(n, d)$ has $e_j \geq 1$ and $e_j \leq e_i$ for some $i \neq j$, then*

$$D(\mathbf{e} - \delta_j + \delta_i) \leq D(\mathbf{e}).$$

Proof. Consider the slice sequence S for \mathbf{e} with indices j, i (note the order). The original composition corresponds to S_{e_j} , and the modified composition $\mathbf{e} - \delta_j + \delta_i$ corresponds to S_{e_j-1} .

Since $e_j \leq e_i$, we have $2e_j \leq e_j + e_i = q$. If $2e_j < q$, then by unimodality (part (i)), $S_{e_j-1} \leq S_{e_j}$, giving the result. If $2e_j = q$, then $e_j = e_i$ and palindromicity gives $S_{e_j-1} = S_{q-(e_j-1)} = S_{e_j+1}$; unimodality then gives $S_{e_j-1} = S_{e_j+1} \leq S_{e_j}$. \square

Proof of Theorem 3.1(b). Starting from any $\mathbf{e} \in E(n, d)$, repeatedly apply the concentrating step: let i be an index achieving the maximum $\max_k e_k$, and if there exists $j \neq i$ with $e_j \geq 1$, replace \mathbf{e} by $\mathbf{e} - \delta_j + \delta_i$. By Lemma 3.9, each step weakly decreases D .

This process terminates because the maximum entry $\max_k e_k$ strictly increases with each step (and is bounded by d). The process terminates when all mass is concentrated at a single index, i.e., when $\mathbf{e} = d \cdot \delta_k$ for some k .

Thus every \mathbf{e} can be transformed to a concentrated \mathbf{e}_* with $D(\mathbf{e}_*) \leq D(\mathbf{e})$. In particular, the minimum of D is achieved at a concentrated vector. \square

This completes the proof of Theorem 3.1.

APPENDIX A. METHODOLOGY: AI-ASSISTED DISCOVERY AND PROOF

AI This appendix documents the role of AI systems in discovering the conjecture and producing the proof of Theorem 1.2(b).

A.1. Discovery of the conjecture. The initial observation arose during a hackathon on AI-assisted mathematics organized by Jaume de Dios Pont at ETH Zürich, focused on OpenEvolve [Sha25] — an open-source implementation of the AlphaEvolve system [NVE⁺25] for evolutionary search over functions. Inspired by recent applications of such systems for mathematical exploration [GGSTW25], the author chose as a test problem the optimization of descendant integrals on $\overline{\mathcal{M}}_{g,n}$ for fixed g and n .

While setting up the project using Claude Code (running Anthropic’s Claude Opus 4.5), the AI assistant observed from initial computations:

“The interesting pattern is that for $g = 0$, spreading the exponents seems to give larger values (e.g., $[1, 1, 1, 0, 0, 0] = 6$ vs. $[3, 0, 0, 0, 0, 0] = 1$.)”

In the configuration file for OpenEvolve, the model included:

Mathematical insight for genus 0:

- The integral equals $(n-3)! / (e_1! * e_2! * \dots * e_n!)$
- This is maximized when the exponents are as evenly distributed as possible
- For example, with $n=6$ and $\text{dim}=3$, $[1, 1, 1, 0, 0, 0]$ gives 6, while $[3, 0, 0, 0, 0, 0]$ gives 1

For higher genus:

- The formula is more complex (Witten-Kontsevich theorem)
- General pattern: spreading exponents tends to give larger values
- But the optimal distribution may differ from genus 0

The evolutionary search then quickly converged to balanced vectors as optima.

A.2. Submission to IMPproofBench. Based on these observations, the author formulated the conjecture (for the maximum; the minimum was added later) and submitted it to IMPproofBench [SBD⁺25], a benchmark for evaluating AI systems on research-level mathematical proof generation.

The problem was evaluated against several large language models in an agentic framework with access to SageMath and other computational tools. Results varied significantly:

- Several models produced inconclusive responses: proofs for genus 0 only, or vague suggestions for higher genus.
- Some models claimed false counterexamples.
- At least one model cited a reference that appears to be entirely hallucinated:

“Zhou, Y. (2018). A proof of the concavity conjecture for intersection numbers of ψ -classes. Journal of Differential Geometry, 110(2), 361–389.”

This paper does not exist; the above volume of JDG instead features the paper [Son18] on pages 345–377, being the last paper in this issue.

- Some models used the provided SageMath tools to verify the conjecture computationally for small values of g and n .

- Multiple evaluations of [the AI models](#) o3, GPT-5 and GPT-5 Pro converged on the proof strategy using nefness and Khovanskii–Teissier log-concavity presented in Section 2.

The proof in Section 2 is taken from one of the GPT-5 evaluations. Interestingly, the single evaluation of the problem against GPT-5.1 did not solve the problem.

A.3. Partial formalization and Lean Blueprint.

Based on discussions with colleagues and inspired by [a proposal of Emily Riehl](#), we also decided to formalize at least parts of the proof found by GPT-5, as presented in Section 3. To our knowledge, the prerequisite results in algebraic geometry going into the proof (divisors on algebraic schemes, intersection numbers, nef classes, moduli spaces of curves, the Witten–Kontsevich theorem, etc) are not yet formalized in Lean. Thus we decided to split the proof in a purely combinatorial optimization result (Theorem 3.1), and a geometric part proved non-formally (Theorem 3.2).

As for the work of formalization itself: the author has not had any prior experience with Lean or the mathlib. For all manipulations of [the](#) respective `.lean` file we used Claude Code (Opus 4.5) which loaded the [Lean 4 Skills](#) collection developed by Cameron Freer. For particularly tricky problems or intermediate planning we also copy-pasted summaries of the current formalization state by Claude together with the latest version of the Lean file into ChatGPT 5.2, pasting its answer back to Claude Code.

Progress was slow, but steady, and after several hours of work the Lean file compiled without errors or missing proofs. During this time, the contribution of the author was confined to:

- providing an initial description of the result to be formalized, with existing AI-generated proof for context,
- giving very few pointers on possible mathematical issues,
- occasionally soliciting updates, general management of the order of work, and interventions via GPT 5.2 commentary when progress stagnated.

In particular, we did not edit a single line of `.lean` code. A transcript of the conversation can be found [here](#).

After the formalization was complete, the Lean code was made available via a GitHub repository using the [Lean blueprint](#) plugin developed by Patrick Massot and collaborators. Again, with exceptions of a few clicks on the GitHub website and debugging information provided by the author, all files were created and edited by Claude Code.

The resulting annotated file `balanced-vectors-blueprint.lean` was then converted back to the human-readable proofs in Section 3 by [Claude Opus 4.5](#).

The whole process described above took about a day to complete.

A.4. Paper preparation.

This paper was prepared with assistance from Claude (Anthropic), which helped structure the document, identify issues in the AI-generated proof (including the incorrect genus-0 formula in the original remarks), and draft portions of this methodology section.

Below we collect some links to shared example conversations with AI models, which include the prompts used to get the relevant outputs:

- [Paper drafting session](#) (Claude Opus 4.5)

- Summary of the IMPProofBench project and literature overview created to provide context for the paper drafting (Claude Opus 4.5 + Research Mode)
- Reference for (2) (Gemini 3 Pro), Reference for finite cover of $\overline{\mathcal{M}}_{g,n}$ by scheme (ChatGPT 5.2)
- Proof of Theorem 1.2(a) (Gemini 3 Pro)
- Writing code to verify equation (5) in examples (Claude Opus 4.5)
- Writing the LaTeX proofs of Section 3 based on the existing Lean Blueprint file (Claude Opus 4.5)
- Examples of Lean assistance provided by ChatGPT 5.2 (Conversation 1, Conversation 2, Conversation 3, Conversation 4)
- Creating LaTeX environments for labelling human and AI paragraphs : Version 1, Version 2, Version 3 (Claude Opus 4.5)

The first raw draft of the paper was also submitted to Claude Code, Gemini 3 Pro and ChatGPT 5.1 for feedback and proofreading. This resulted in several improvements, e.g.

- Restricting to $n > 0$ in Theorem 1.2 to avoid statements about mimima/maxima of the empty set, and adding the term τ_0^{n-1} in that theorem, for clarity.
- Pointing out that in the proof of Theorem 1.2(a), the case $g = 0$ cannot reduce all the way to $n = 1$ via the string equation (see the corresponding human correction).

The usage of the described AI models was covered by subscriptions in the form of Claude Max (\$100/month), ChatGPT Plus (\$20/month) and Google AI Pro (\$20/month).

APPENDIX B. BEST PRACTICES FOR ATTRIBUTION OF AI USE

HUMAN

After some first pioneering papers [GGL24, RSL⁺25], recent months have seen a steadily growing list of articles [AK25, JR25, AM25, IX25, DPS25, For25, Dob25, BCE⁺25] using Large Language Models both to study mathematical questions and writing up the results. See e.g. the [awesome-ai-for-math](#) repository maintained by Seewoo Lee for a list of examples. Given the potential impact of these tools, it is important for the mathematical community to develop some norms and best practices for attributing their use in our work. Below we propose some ideas for such measures, and apply these to the present article.

Methodology section

Mathematical papers that benefited from using LLMs, Deep-Learning, proof assistants, or other software tools should include a section clearly attributing this usage. This section should include or reference any sample prompts, data repositories or code files that could help others to examine and reproduce the employed methods, and to adapt them for their own research.

Transparency about writing tools

When using AI tools for writing first drafts of parts of the paper (e.g. literature overviews, abstract or content summary, proof sketches, etc), all paragraphs with substantial AI contributions could be typeset in a visually distinctive way.

AI

This paragraph was generated by Claude Opus 4.5 (Anthropic) to demonstrate the visual formatting conventions proposed in this appendix. The orange margin bar and label allow readers to immediately identify AI-authored content, complementing the blue markers used for human-written passages.

A first LaTeX template for such sidebars (created jointly with Claude Opus 4.5) will be made available in the auxiliary files published together with the present paper.

Contextualization for non-expert readers

Given the current public interest in AI applications for academic research, results obtained with the help of AI tend to gain increased levels of attention. In particular, those papers are exposed to an audience outside of the specialized area in which the result would usually be received. This creates some incentive, both for the author and the developers of the relevant AI model, to overstate the importance of the presented result. Thus we believe it is the responsibility of the author to contextualize the significance of the presented work for non-experts, ideally in a somewhat prominent place (as with the Author’s Note on the first page of this article).

REFERENCES

- [ACG11] Enrico Arbarello, Maurizio Cornalba, and Phillip A. Griffiths. *Geometry of algebraic curves. Volume II*, volume 268 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2011. With a contribution by Joseph Daniel Harris.
- [ACV03] Dan Abramovich, Alessio Corti, and Angelo Vistoli. Twisted bundles and admissible covers. volume 31, pages 3547–3618. 2003. Special issue in honor of Steven L. Kleiman.
- [AK25] Sergey Avvakumov and Roman Karasev. Tensor rank of the determinant and periodic triangulations of \mathbb{R}^n , 2025.
- [AM25] Boris Alexeev and Dustin G. Mixon. Forbidden sidon subsets of perfect difference sets, featuring a human-assisted proof, 2025.
- [BCE⁺25] Sébastien Bubeck, Christian Coester, Ronen Eldan, Timothy Gowers, Yin Tat Lee, Alexandru Lupsasca, Mehtaab Sawhney, Robert Scherrer, Mark Sellke, Brian K. Spears, Derya Unutmaz, Kevin Weil, Steven Yin, and Nikita Zhivotovskiy. Early science acceleration experiments with gpt-5, 2025.
- [Dob25] Edgar Dobriban. Solving a research problem in mathematical statistics with ai assistance, 2025.
- [DPS25] Ved Datar, Vamsi Pritham Pingali, and Harish Seshadri. The complex monge-ampere equation and an application to uniformisation of surfaces, 2025.
- [DSvZ21] Vincent Delecroix, Johannes Schmitt, and Jason van Zelm. admcycles – a Sage package for calculations in the tautological ring of the moduli space of stable curves. *Journal of Software for Algebra and Geometry*, 11(1):89–112, 2021.
- [For25] Lance Fortnow. Search versus decision for S_2^P , 2025.
- [FP00] Carel Faber and Rahul Pandharipande. Hodge integrals and Gromov-Witten theory. *Inventiones Mathematicae*, 139(1):173–199, 2000.
- [GGL24] Robert Ghrist, Julian Gould, and Miguel Lopez. Lattice-valued bottleneck duality, 2024.
- [GGSTW25] Bogdan Georgiev, Javier Gómez-Serrano, Terence Tao, and Adam Zsolt Wagner. Mathematical exploration and discovery at scale, 2025.
- [IX25] Paata Ivanisvili and Xinyuan Xie. Counterexample to majority optimality in ncd with erasures, 2025.
- [IZ92] C. Itzykson and J.-B. Zuber. Combinatorics of the modular group. II. The Kontsevich integrals. *Internat. J. Modern Phys. A*, 7(23):5661–5705, 1992.
- [JR25] Uijeong Jang and Ernest K. Ryu. Point convergence of nesterov’s accelerated gradient method: An ai-assisted proof, 2025.
- [Kon92] Maxim Kontsevich. Intersection theory on the moduli space of curves and the matrix Airy function. *Communications in Mathematical Physics*, 147(1):1–23, 1992.
- [Laz04] Robert Lazarsfeld. *Positivity in algebraic geometry. I*, volume 48 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results*

in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]. Springer-Verlag, Berlin, 2004. Classical setting: line bundles and linear series.

[mC20] The mathlib Community. The lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs*, POPL '20, page 367–381. ACM, January 2020.

[MU21] Leonardo de Moura and Sebastian Ullrich. The lean 4 theorem prover and programming language. In *Automated Deduction – CADE 28: 28th International Conference on Automated Deduction, Virtual Event, July 12–15, 2021, Proceedings*, page 625–635, Berlin, Heidelberg, 2021. Springer-Verlag.

[NVE⁺25] Alexander Novikov, Ngan Vu, Marvin Eisenberger, Emilien Dupont, Po-Sen Huang, Adam Zsolt Wagner, Sergey Shirobokov, Borislav Kozlovskii, Francisco J. R. Ruiz, Abbas Mehrabian, M. Pawan Kumar, Abigail See, Swarat Chaudhuri, George Holland, Alex Davies, Sebastian Nowozin, Pushmeet Kohli, and Matej Balog. Alphaevolve: A coding agent for scientific and algorithmic discovery, 2025.

[RSL⁺25] Tomer Raz, Michael Shalyt, Elyashev Leibtag, Rotem Kalisch, Shachar Weinbaum, Yaron Hadad, and Ido Kaminer. From euler to ai: Unifying formulas for mathematical constants, 2025.

[SBD⁺25] Johannes Schmitt, Gergely Berczi, Jasper Dekoninck, Jeremy Feusi, Tim Gehrunger, Raphael Appenzeller, Jim Bryan, Niklas Canova, Timo de Wolff, Filippo Gaia, Michel van Garrel, Baran Hashemi, David Holmes, Aitor Iribar Lopez, Victor Jaeck, Martina Jorgensen, Steven Kelk, Stefan Kuhlmann, Adam Kurpisz, Chiara Meroni, Ingmar Metzler, Martin Moller, Samuel Muoz-Echaniz, Robert Nowak, Georg Oberdieck, Daniel Platt, Dylan Possama, Gabriel Ribeiro, Ral Sanchez Galn, Zheming Sun, Josef Teichmann, Richard P. Thomas, and Charles Vial. IMPProofBench: Benchmarking AI on research-level mathematical proof generation, 2025.

[Sha25] Asankhaya Sharma. Openevolve: an open-source evolutionary coding agent, 2025.

[Son18] Antoine Song. Embeddedness of least area minimal hypersurfaces. *J. Differential Geom.*, 110(2):345–377, 2018.

[The24] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 10.5)*, 2024. <https://www.sagemath.org>.

[Wit91] Edward Witten. Two-dimensional gravity and intersection theory on moduli space. In *Surveys in Differential Geometry*, volume 1, pages 243–310. Lehigh University, Bethlehem, PA, 1991.

DEPARTMENT OF MATHEMATICS, ETH ZURICH, SWITZERLAND

Email address: johannes.schmitt@math.ethz.ch