素数の偏りに関する系統的研究

A Systematic Study on the Bias of Primes

青木 美穂(代表)/ 島根大学 総合理工学部　　横山 重俊(副代表)/ 国立情報学研究所

1. 研究背景: チェビシェフの偏りと深リーマン予想

ガロア拡大における素数や素イデアルの分解は,**イデアル類群・単数群・コホモロジー群**など重要な不変量の構造を決定する.アーベル拡大では, 高木–Artin の**類体論**とチェボタレフ密度定理が分解を完全に記述するが, **有限の範囲では**現在知られているこれらの理論的結果からは得られない**素数の「偏り」**が観測されている.

「チェビシェフの偏り」(1853): 4 で割って 3 余る素数が 1 余る素数より早く現れる現象 – チェボタレフの密度定理からは説明できない長年の問題. 代数的整数論の言葉では,有限範囲では**ガウス数体 $\mathbb{Q}(i)$ で分解しない素数が分解する素数より早く現れる**ことを意味する.

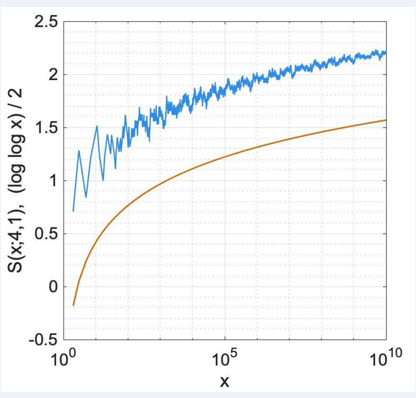
Aoki–Koyama 定理 (2023) [1]: 研究代表者らは,この現象が**深リーマン予想 (DRH)** と**同値**であることを証明した. DRH は Artin L 関数のオイラー積が中心値 $s=1/2$ で収束することを主張する予想で黒川信重氏により定式化され,リーマン予想・一般リーマン予想を含む強い予想である. 大域体の有限次ガロア拡大 L/K に対し $\sigma \in \mathrm{Gal}(L/K)$ について

$$\mathrm{bias}(x) \fallingdotseq (M(\sigma)+m(\sigma)) \log\log x,$$

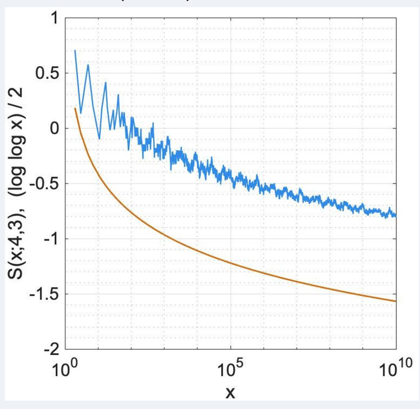
$M(\sigma)$ 代数的因子(ガロア群の構造で決まる), $m(\sigma)$ 解析的因子(Artin L 関数の中心値での零点の位数で決まる).

偏りの方向は $M+m$ の符号で決定: $>0 \Rightarrow$ 遅く現れる, $=0 \Rightarrow$ 偏りなし, $<0 \Rightarrow$ 早く現れる. **偏りの大きさ**は $|M+m|$ で決定. 素数の出現レースの**「ハンディキャップ」の正体は ガロア群の構造と L 関数中心値の零点の位数**に隠されている.本研究は**非可換ガロア拡大をはじめとする各種数論的対象**での系統的研究を目指す.

計算値(青)と理論値(茶)



$S(x;4,1)$: 1 余の素数



$S(x;4,3)$: 3 余の素数

2. 研究目的と学術的位置づけ

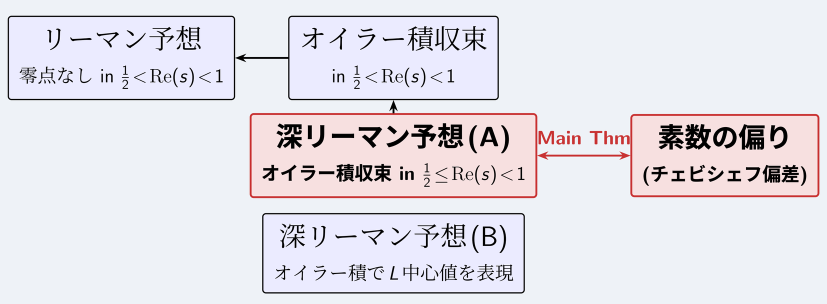
研究目的

代数体のガロア拡大などの数論的対象における素数・素イデアルの偏りと, 対応する Artin L 関数の中心値の振る舞いとの関係を, **AI エージェントと大規模計算による実験数学として系統的に研究し, AI for Math の有効な実践事例を生み出す**.

- **非可換ガロア拡大**での偏りを **LMFDB 規模で系統的検証**.
- L 中心値の**零点の位数** $m(\sigma)$ **• Root number** $W(\rho)$ と偏りの相関を解明.
- 代数的構造と零点の位数を結ぶ理論を予想し,証明の手がかりを得る.
- **保型 L 関数等**への展開でラングランズ対応に新知見を加える.

学術的位置づけ

- チェビシェフの素数偏りは **深リーマン予想と同値**. 深リーマン予想からリーマン予想・一般リーマン予想が従う – その検証の一翼を担う.
- **BSD 予想**(1960 年代,Hasse–Weil L 関数の数値実験から定式化)の精神を継承する **Artin L 関数版**. アーベル拡大の他 [1] から**非可換ガロア拡大** (Q_8, D_4, Q_{16} 他)への系統的拡張は**世界に先駆ける試み**.
- Artin L 関数と保型 L 関数の偏りの比較は**ラングランズ対応の数値的検証**への新たな道筋.



深リーマン予想と素数の偏りの同値関係 (Aoki–Koyama [1] Main Thm 2.2)

3. 予備実験: Q_8 拡大体での検証

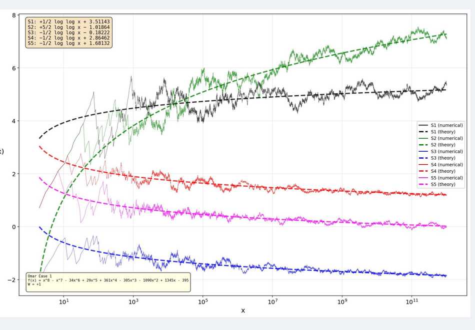
概要: 有理数体 $\mathbb{Q}$ 上のガロア群が四元数群 Q_8 と同型なガロア拡大体 $L/\mathbb{Q}$ に対し,**唯一の既約 2 次元表現** ρ_0 に付随する Artin L 関数 $L(s,\rho_0)$ の中心値 $s=1/2$ における**零点の位数** $m(\rho_0)=\mathrm{ord}_{s=1/2}L(s,\rho_0)$ **と素数の偏り (Chebyshev bias) の関係**を数値的に検証. フロベニウス共役類 σ に対する偏り関数 $S_\sigma(x)=\pi_{1/2}(x)-\frac{|G|}{|c_\sigma|}\pi_{1/2}(x;\sigma)$ は青木–小山予想 [1] により $S_\sigma(x)=(M(\sigma)+m(\sigma))\log\log x+c+o(1)$ と漸近的に等しくなる. $M(\sigma)$ は**代数的因子**(ガロア群の構造), $m(\sigma)$ は**解析的因子**(L 関数の零点の位数).

データ: Omar [2] の Q_8 拡大体 **23 例**, $x \le 10^{12}$ で偏り関数 S_1–S_5 を計算,[1] の理論値とプロット:

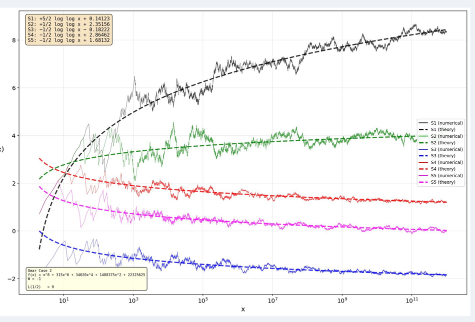
- $W(\chi_0)=+1$, $m(\rho_0)=0$ **(9 ケース)**: S_1 係数 $+1/2$, S_2 係数 $+5/2$
- $W(\chi_0)=-1$, $m(\rho_0)=1$ **(14 ケース)**: S_1 係数 $+5/2$, S_2 係数 $+1/2$
  S_1 と S_2 が入れ替わる

実線: 数値計算値,破線: 理論値.

公開: `github.com/jxta/quaternion-bias-project`



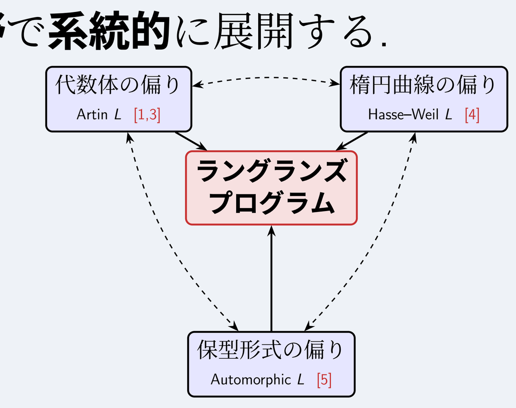
Case 1: $W(\rho_\sigma)=+1$, $m(\rho_\sigma)=0$



Case 2: $W(\rho_\sigma)=-1$, $m(\rho_\sigma)=1$

4. 本研究: ラングランズ視野での「系統的」展開

予備実験(Q_8 拡大体)を出発点とし,**素数の偏り**を**ラングランズプログラムの視野**で**系統的**に展開する.



方針: **(i)** 個別研究 → **(ii)** 対象間の対応関係(ラングランズ的) → **(iii)** 同一の理論枠組みで**ラングランズ対応の数値的検証**を系統的に実施.

既知の定理 (Aoki–Koyama 2023 [1] §2–4): 偏りは全て $\boxed{C\cdot\log\log x+O(1)}$ の形, 係数 C が対象により決定.

対象	係数 C	出典
Artin L 大域体	$M(\sigma)+m(\sigma)$	[1] Thm 2.2, [3]
Hasse–Weil L 楕円曲線	$\frac{1}{2}-\mathrm{rank}(K)$	[1] Thm 4.2, [4]
Automorphic L $\tau(p)/p^6$	$\frac{1}{2}$	[1] Thm 4.1, [5]

共通構造を *LMFDB*+α 規模で計測, 数論対象間の相関と「またがる構造」を解明.

5. アプローチ: AI 駆動型実験数学

LLM × 数式処理 × LMFDB × mdx 大規模並列 + **meta-AI による自己改善**

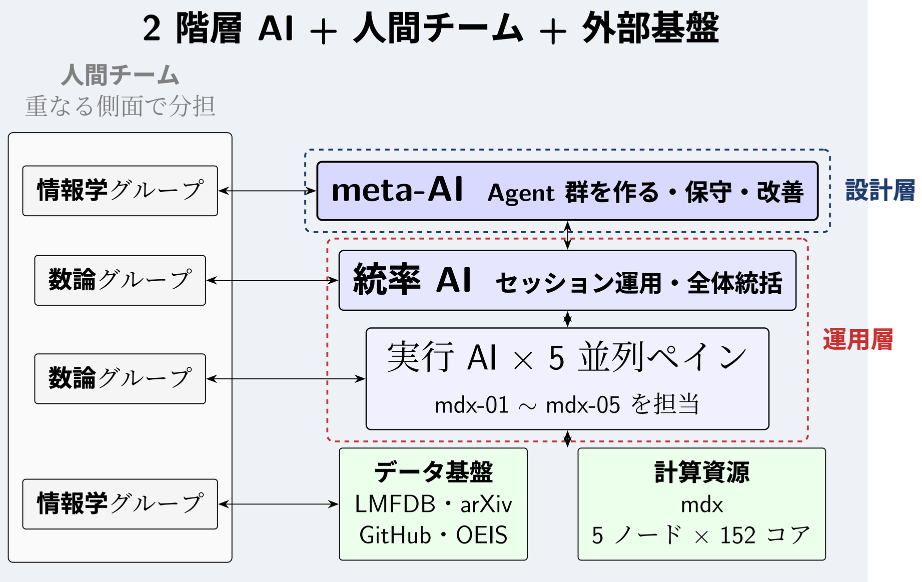
「系統的」の定義: 個別事例の確認ではなく, **LMFDB に登録された大規模な数論的対象**をスタートラインに対して**偏りを計測し各種特徴量との相関を網羅的に解明**.

2 階層 AI による自律的探索(右図参照):

- **meta-AI**(設計層): Agent 群を作る・保守・改善 – **自己改善ループ**が本研究の核
- **統率 AI**(運用層): セッション運用・ジョブ投入の意思決定
  - **実行 AI × 5 並列ペイン**: **mdx-01 ∼ mdx-05** を担当 / 並列実験・自己検証
- **SageMath / Julia / PARI / Nemo.jl** で厳密検算(*Generator – Verifier – Reviser*)

外部基盤: LMFDB 等**データ基盤**から対象取得・ **mdx I**(5 ノード × 152 コア)で並列実験.

2 階層 AI + 人間チーム + 外部基盤



6. 検証したい仮説

■ **数論研究**

- H1 **(主仮説)** 偏りの符号と大きさは Artin L 関数の**零点の位数** $m(\sigma)$ で決定される
- H2 偏りに関して**数論的な対象の中に共通の法則**が存在
- H3 **保型 L 関数**でも同様の偏りが観測され, **ラングランズ対応の数値的検証**

H1–H3 を LMFDB 規模で同時検証する点が「系統的」.

■ **AI 駆動研究**

- H4 **(主仮説)** 本研究の対象群が **AI for Math の試金石として機能**する

H4 を支持する 5 性質:

- P1 **厳密な検証可能性**: 数値計算で誤差なく確認可能
- P2 **仮説空間の階層構造**: Galois 群族 × 指標 × 素数範囲
- P3 **計算コストが手頃**: クラスタ + 数式処理系で到達可能
- P4 **形式化標的の明確性**: Lean 4 ステートメント化可能
- P5 **AI-ready なドメインデータベースの存在**: LMFDB

7. システム構成と分担例

AI 構成: Mac 上に **meta-AI**(設計層)・**統率 AI**(運用層)・**実行 AI × 5**(並列ペイン)が全て常駐. 実行 AI が **SSH 経由で mdx-01 ∼ mdx-05** 上で数値計算を実行.

データフロー:

- **LMFDB** から数体・L 関数を取得(ingest)
- **LLM API**(Claude / Gemini)で推論駆動
- **Generator – Verifier – Reviser** ループで自律検証

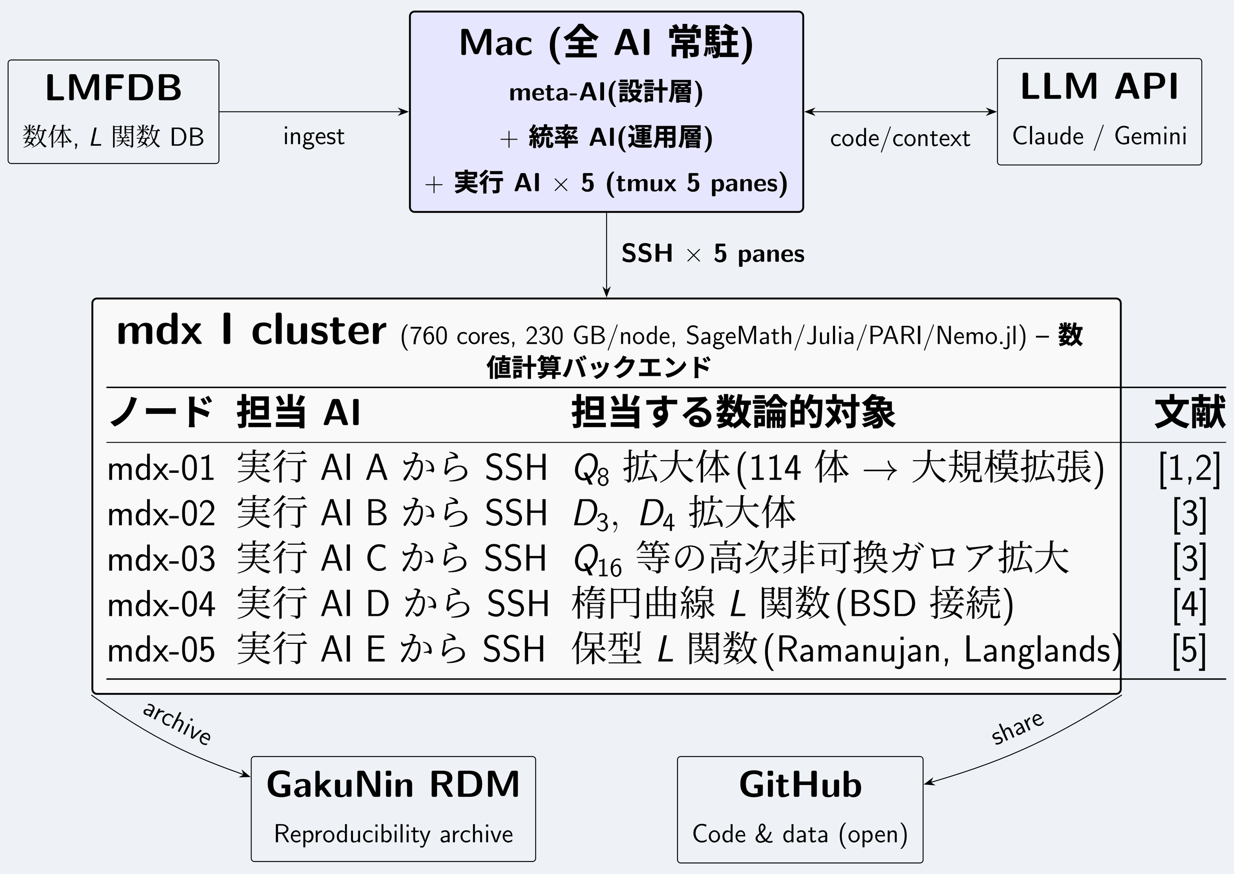
分担: 各**実行 AI** が **異なる数論的対象**を分担し, ブロック 4 の 3 スポーク(Artin L / Hasse–Weil L / Automorphic L)を同時走査.

出力:

- **GakuNin RDM** へアーカイブ(再現性確保)
- **GitHub** で公開(open access)

ブロック 4 との対応:

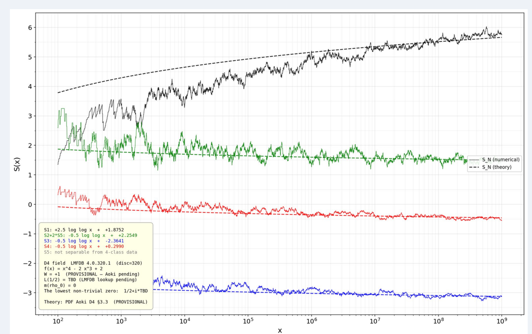
- Node 1, 2, 3 = Artin L(代数体ガロア拡大)
- Node 4 = Hasse–Weil L(楕円曲線)
- Node 5 = Automorphic L(保型形式)



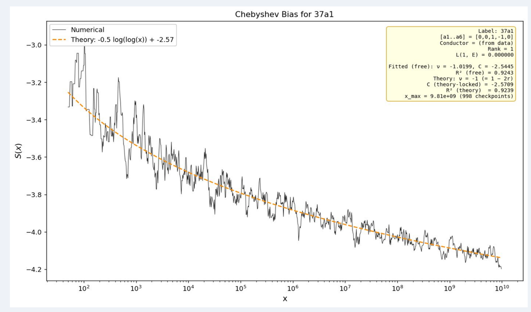
ノード	担当 AI	担当する数論的対象	文献
mdx-01	実行 AI A から SSH	Q_8 拡大体(114 体 → 大規模拡張)	[1,2]
mdx-02	実行 AI B から SSH	D_3, D_4 拡大体	[3]
mdx-03	実行 AI C から SSH	Q_{16} 等の高次非可換ガロア拡大	[3]
mdx-04	実行 AI D から SSH	楕円曲線 L 関数(BSD 接続)	[4]
mdx-05	実行 AI E から SSH	保型 L 関数(Ramanujan, Langlands)	[5]

8. これまでの成果

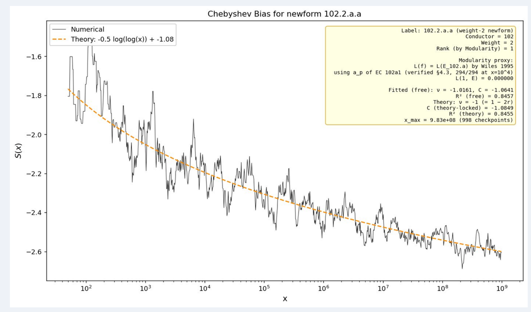
- **代数体: D_3, D_4, Q_{16} 拡張への展開**
  - Q_8 114 体での予備実験を基盤に **LMFDB 5 万件規模**へのスケール拡張を開始
  - D_3 / D_4 パイプライン構築完了; **Bailleul [3] 予測の照合**を開始 (右図 D_4 例)
- **楕円曲線への適用 [4]**
  - $\Sigma\, a_v/q_v$ パイプライン稼働; LMFDB rank $=0,1,2,3,4$ を網羅
  - スロープ $\alpha\approx(1-2\,\mathrm{rank})/2$ の予備的支持; 37a1 **(Rank 1)** で $\nu\approx-1$ を確認(右図)
  - Kaneko–Koyama 公式 [4] の数値的検証フェーズ
- **保型形式: Ramanujan τ [5] と weight-2 newform**
  - $\Sigma\,\tau(p)/p^6$ の $x\le10^9$ 数値検証フレームワーク稼働
  - $\frac{1}{2}\log\log x$ **漸近**を予備的に確認 [5] Thm 4.1 に整合
  - `102.2.a.a` **(weight-2 newform, Rank 1)** で $\nu\approx-1$ を確認(右図, modularity proxy 経由)
  - $\mathrm{sym}^2\Delta$ $(-\frac{1}{2}\log\log x)$ も計算スタック構築済み
- **ラングランズ対応の数値的検証**
  - **3 領域**(代数体・楕円曲線・保型形式)で共通の $C\cdot\log\log x$ 構造を確認
  - 共通理論枠 (M,m,W) の系統的計測体制が完成
  - AK formula $B(x)=\frac{1}{2}\log\log x+C_0+2\Sigma\,\mathrm{Ci}(\gamma_n\log x)$ の数値支持



D_4 拡大体 *(disc=320):* $x\le10^9$



EC 37a1 (Rank 1): $\nu=-1.02$, $x\le10^{10}$



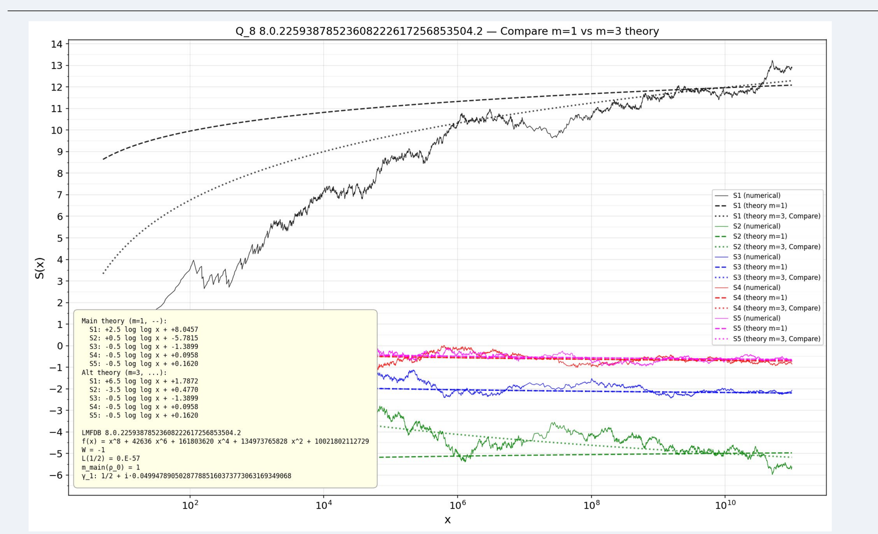
MF 102.2.a.a (Rank 1): $\nu=-1.02$, $x\le10^9$

- **ハードケース現象**

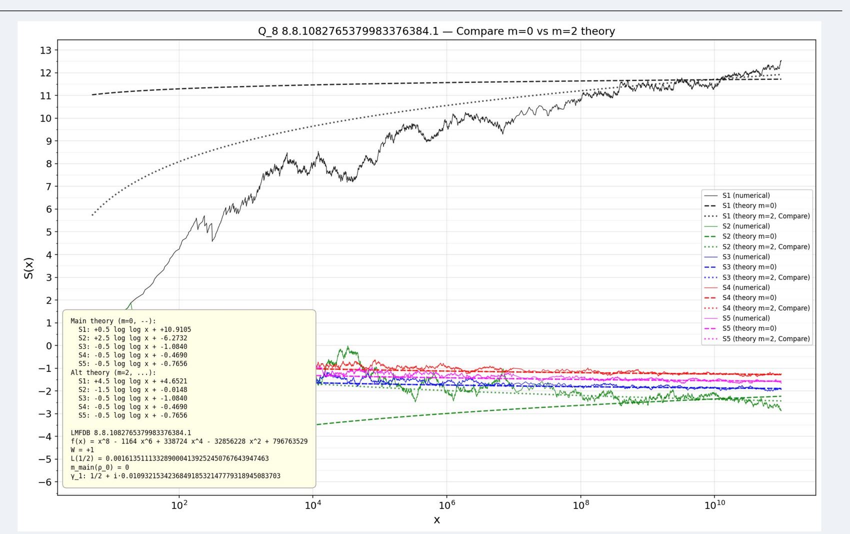
零点虚部 γ_1 が小さい体では,理論値 $m=m_{\mathrm{main}}$ なのに見かけ係数が $m_{\mathrm{main}}+2$ 寄りに観測される現象を発見(下図 2 枚: Q_8 で $m=1\to3$ および $m=0\to2$ の例).

原因仮説: Explicit formula 由来の Ci 振動が, $\gamma_1\log x\lesssim2\pi$ の領域では $\log\log x$ の主項を凌駕.

他対象での確認: **Q_8 拡大体で発見したハードケース現象は, D_3/D_4 拡大体・楕円曲線・保型形式など他の数論的対象でも同様に確認されつつある**.



Q_8 ハードケース *(*$m=1$ *vs* $m=3$*):* $\gamma_1\approx0.0499$, $x\le10^{10}$



Q_8 ハードケース *(*$m=0$ *vs* $m=2$*):* $\gamma_1\approx0.0109$, $x\le10^{10}$

9. 今後の研究計画

■ **数論研究**

- **LMFDB** による偏り事例の**蓄積と整理**
- **LMFDB 以外**の事例への拡張
- 偏りによる**指紋定義**と活用

■ **AI 駆動研究**

- **AI 駆動型実験数学環境**の活用と改善
- **再現性の高い研究成果発表**の実践と論文執筆

10. 期待される成果

■ **数論研究**

- **素数偏り**と数論的対象の特徴量の網羅的検証
- **ラングランズ対応**に関する新たな数値的・理論的知見

■ **AI 駆動研究**

- **「AI を活用した再現可能な研究」**の方法論実証
- **AI for Science 全体**への展開条件の整理

参考文献

[1] M. Aoki, S. Koyama. *J. Number Theory* 245:233–262, 2023.
[2] S. Omar. *Experiment. Math.* 10(2):237–245, 2001.
[3] A. Bailleul. *Algebra Number Theory* 15(4):999–1041, 2021.
[4] I. Kaneko, S. Koyama. arXiv:2206.05445, 2022.
[5] S. Koyama, N. Kurokawa. *Proc. Jpn. Acad.* 98A:35–39, 2022.
[6] 横山重俊. 京大数理研講究録 No. 2333, 2026.

関連リソース

GitHub: `github.com/jxta/quaternion-bias-project`
LMFDB: `www.lmfdb.org`　**mdx**: `mdx.jp`

謝辞

本研究は JHPCN 共同研究課題 **jh261018** として実施. **mdx I** の利用に感謝申し上げます.