Fine-Tuning Small Reasoning Models for Quantum Field Theory

Nathaniel S. Woodward[1,*], **Zhiqi Gao**[2], **Yurii Kvasiuk**[1], **Kendrick M. Smith**[3], **Frederic Sala**[2], **Moritz Münchmeyer**[1]

[1]Department of Physics, University of Wisconsin-Madison
[2]Department of Computer Science, University of Wisconsin-Madison
[3]Perimeter Institute for Theoretical Physics
[*] *Corresponding author: nwoodward2@wisc.edu*

https://huggingface.co/datasets/nswoodward/VerifiableQFT

https://github.com/nswood/VerifiableTPData

Abstract

Despite the growing application of LLMs to theoretical physics, there is little academic exploration of how domain-specific physics reasoning ability develops while training these models. To investigate this, we perform the first academic fine-tuning of small (7B-parameter) reasoning models dedicated specifically to problem solving in theoretical physics. Because open-source verifiable training data required to train such capabilities is scarce, we developed a robust data generation pipeline that can create synthetic problems and adapt existing human-authored problems suitable for model training. Selecting Quantum Field Theory (QFT) as our primary domain, we generated over 2,500 synthetic problems alongside a curated collection of human-adapted problems sourced from arXiv and standard pedagogical resources. We conduct both Reinforcement Learning (RL) and Supervised Fine-Tuning (SFT) experiments finding SFT excels on in-distribution QFT while RL generalizes better to other physics domains. Additionally, we analyze the underlying shifts in reasoning behavior, and examine the features driving challenging physics tasks. Finally, we publicly release our data pipeline, verifiable QFT training data, and $\sim$200M tokens of QFT reasoning traces.

1 Introduction

LLMs have recently become powerful at mathematical reasoning in theoretical physics [1–13]. It is interesting to study the training dynamics and internal reasoning processes that allow these capabilities. Currently most industry labs do not publish their training data, precise training setup, or intermediate results of their training processes. Therefore to study learning and reasoning dynamics of LLMs in academia (at a lower level of model capability) it is important to be able to run smaller experiments within our target domain. To our knowledge, the present work is the first public exploration of fine-tuning LLM reasoning models specifically for theoretical physics problem solving (here quantum field theory) in academia[1]. Quantum field theory (QFT) provides an ideal domain for this exploration due to a wealth of online pedagogical resources, pedagogical calculations being tedious yet analytically tractable without standard computational tools, and following a well-defined progression (from classical field theory and scalar fields, to fermions and non-abelian gauge theories) which provides a clear gradient of topic complexity.

[1]Prior papers have tuned LLMs with SFT on physics corpora, but not explored reasoning fine-tuning for problem-solving [14–19].

Conference on Physics and AI at Stanford University (PAI 2026).

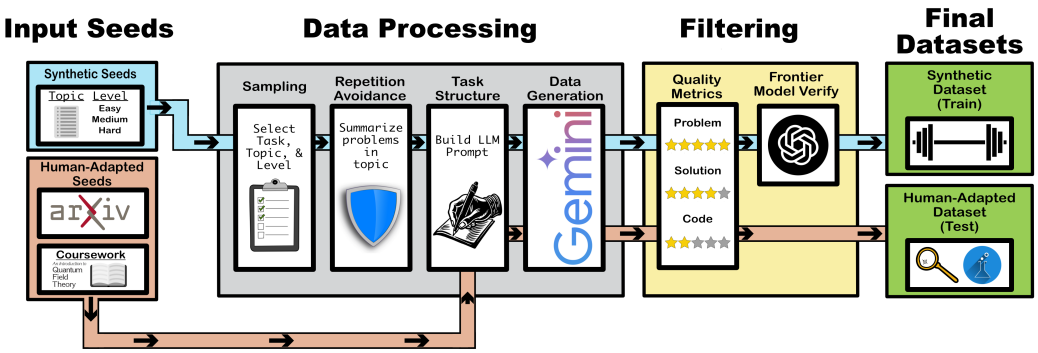

Figure 1: Schematic overview of the data curation pipeline. The workflow is divided into two tracks: a fully synthetic pipeline (blue arrows) utilized for generating the training dataset, and a human-adapted pipeline (orange arrows) derived from arXiv manuscripts and coursework for the testing dataset.

2 Data Pipeline and Datasets

Most of the progress thus far in LLM mathematical reasoning has been driven by training with outcome verification [20–23]. Therefore, we design our dataset and experiments around verifiable tasks in QFTs. In discussing these tasks, we distinguish between *domain* and *operational* difficulty. **Domain difficulty** reflects how conceptually challenging a topic is to understand and the depth of background knowledge required to approach it. Conversely, **operational difficulty** measures the mechanical and logical demands of problem solving strictly independent of domain knowledge. In the following sections, we detail the curation of three synthetic training datasets (Easy, Medium, and Hard QFT) alongside two evaluation sets. Drawn from a fixed topic list spanning advanced undergraduate to post-graduate levels, each synthetic dataset isolates a specific operational difficulty across a wide spectrum of domain difficulties.

Verifiable Physics Problems For all tasks, we require the model to implement its final analytic solution within a specified Python function and employ a suite of test cases for verification. We decompose a single problem into: **problem statement**: the setup and tasks, **answer requirements**: the Python skeleton that the model implements, **solution**: the author's necessary derivations and final result, and **code**: the fully functional script that numerically evaluates the derived analytic result. We categorize verifiable tasks into five types: **direct calculation** (computing numerical values or functions), **hidden-coefficient derivation** (extracting constants from expressions), **ratio and comparison** (outputting a ratio or boolean from related quantities), **categorical classification** (assigning discrete labels), and **logical consistency checks** (verifying physical conditions for binary outputs).

Problem Seeds Our data generation and processing pipeline is structured around *problem seeds*, which we categorize into **human-adapted** and **synthetic** classes. **Human-adapted seeds** are unprocessed problem-solution pairs which we collected from available theoretical physics resources. We manually curated 565 QFT pedagogical problem seeds [24–31] as well as scraped arXiv using an agentic framework to collect 397 seeds from over 100 theoretical physics arXiv manuscripts [32–162]. **Synthetic seeds** consist of physics topic and a target operational difficulty. To ensure a comprehensive curriculum and a varied domain difficulty, we organize our topics into four distinct pedagogical categories: Advanced Undergraduate (AU), Graduate (GR), Advanced Graduate (AG), and Post Graduate (PG). This hierarchical structure is crucial for two reasons: it ensures dataset breadth and minimizes data overlap by restricting generation to well-defined subgroups.

Data Generation Pipeline Our generation pipeline processes both synthetic and human-adapted seeds. To generate diverse synthetic datasets, we sample uniformly across domain levels while using a registry of previous problem summaries to prevent repetition. Frontier models (`Gemini-2.5-pro` for Easy/Medium; `Gemini-3-pro` for Hard/human-adapted) are prompted with these seeds within our system prompt to generate complete problem-solution pairs and numerical test cases. This approach automates the construction of code-based verification used in TPBench [4]. We then enforce a rigorous quality filter, retaining only problems that achieve "Excellent" ratings from LLM judges

across a five-point rubric assessing problem definition, solution completeness, explanatory quality, test case validity, and source-seed correspondence (if applicable). Finally for fully synthetic problems, we keep only those problems that `GPT-5` could successfully solve within a strict attempt threshold.

Curated Datasets Our **Synthetic** datasets consist of Easy, Medium, and Hard QFT splits. Combined, we generate 4,602 synthetic problems and successive quality filtering yields approximately 1,000 training problems for both Easy and Medium, and 550 for Hard, with each split paired with an 80-problem validation set balanced across domain levels. Second, our **Human-Adapted** datasets leverage human-authored seeds to generate the arXiv (333) and QFT Pedagogy (480) validation sets. Comprising over 800 quality-checked problems, these sets provide insight into reasoning capabilities both within QFT and across broader theoretical physics domains. In total, all data generation costs approximately $2500.

3 QFT Fine-Tuning Experiments

We fine-tune the reasoning model `DeepSeek-R1-Distill-Qwen-7B` [21], referred to as `DeepSeek-7B` for brevity, on the Easy portion of our synthetic dataset using both reinforcement learning (RL) and supervised fine-tuning (SFT). Refer to the full pre-print [163] for further details.

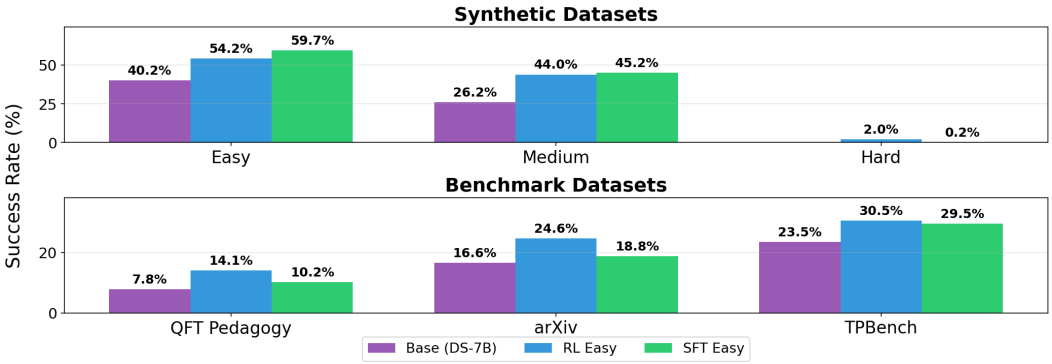

Figure 2: Validation accuracy (mean of 5 rollouts per problem) of the base DeepSeek-7B model compared to RL and SFT fine-tuned variants trained on Easy QFT.

SFT Data Generation To develop a robust SFT dataset for physics reasoning, we require high-quality chain-of-thought (CoT) solutions from capable "teacher" models. We test three teacher models within our compute budget: `oss-120b` (medium reasoning effort) [164], `Qwen3-30B-A3B` [22], and `Qwen3.5-122B-A10B` [165]. We generate five solution attempts ($T = 1.0$, 32,768 token responses) for every problem across the training and validation splits of our synthetic datasets. We apply rejection sampling [166] to filter all generated CoTs, retaining only the traces that yield the correct final result. In testing, we find `Qwen3-30B-A3B` to be best suited for `DeepSeek-7B` SFT.

Training Parameters All experiments were conducted on 4xH200 GPUs at PI. For RL (with Verl [167], parameters inspired by [168]), we use GRPO [20] with an accuracy-only reward function (all test cases passing), 32 rollouts per prompt, temperature of $T = 1.0$, a constant learning rate of 1×10^{-6}, a global batch size of 4, and a maximum response length of 20,000 tokens. We run SFT using Axolotl [169] with sample packing of 35,000 tokens, a micro-batch of 1 per GPU, 2 gradient accumulation steps, fused AdamW optimizer with a peak learning rate of 2×10^{-5}, a cosine decay after a 10% warmup period, and a weight decay of 0.01.

Results We fine-tune `DeepSeek-7B` with RL for 2,500 steps (160 hrs) on Easy QFT and with SFT using `Qwen3-30B` rejection sampled CoT responses to Easy QFT for 5 epochs ($\sim$1 hr). In Figure 2 we compare mean accuracies (5 attempts) over the Easy, Medium, and Hard QFT validation sets, the human-adapted datasets (arXiv, QFT pedagogy), and human-authored TPBench [4]. On the synthetic in-distribution datasets, SFT achieves higher performance compared to RL with meaningful gains on Easy QFT (+5.5) and slight improvements on Medium QFT (+1.2). However, we find RL generalizes better to our out-of-distribution benchmarks. We find gains on the human-adapted benchmarks (+3.9 QFT, +5.8 arXiv) with slight improvements on TPBench (+1.0). To better characterize performance

on QFT Hard, we calculate pass@100 attempts yielding 5.0% base, 12.5% RL, and 8.8 % SFT where both fine-tuned models demonstrate gains over the base.

CoT Error Analysis To further understand the change in validation performance after fine-tuning we generated 100 attempts for the 80 problems in the Easy QFT validation set. Next, to understand *how* fine-tuning improves model reasoning, we develop an automated error analysis pipeline that classifies the types of errors present in incorrect rollouts. Our *distill-then-classify* approach first distills the reasoning models' verbose Chain-of-Thought (CoT) rollouts into sequences of reasoning steps and then references the solution to identify errors. We categorize these errors into a global taxonomy of four types: **Factual** (physics or parsing errors), **Mathematical** (calculation mistakes), **Logical** (invalid deductions), and **Executional** (code translation bugs).

This analysis reveals three core findings: **1. Error shifts:** We select the 20 most improved Easy QFT problems (overlap of 10) for the SFT and RL-finetuned models. Figure 3 illustrates error analysis results showcasing that all error categories decrease in absolute count. Per incorrect rollout, both finetuned models dramatically reduce factual error frequency (0.55 $\rightarrow$ 0.28 for RL, 0.32 for SFT), while mathematical errors constitute a larger share of the remaining incorrect attempts (0.50 $\rightarrow$ 0.64 for RL, 0.60 for SFT). **2. Backtracking:** To characterize reasoning shifts, we analyze backtracking frequency, defined as the number of explicit self-corrections per trace. We find that while base and finetuned models share nearly identical overall rates, fine-tuning redistributes this behavior among correct/incorrect traces (Figure 4). Specifically, the finetuned model increases productive self-correction on ultimately successful attempts while reducing futile backtracking loops on incorrect ones. **3. Difficulty Drivers:** Finally, expanding our evaluation to include the Medium and Hard QFT validation sets (Figure 5) reveals that LLM solve rates decline as operational difficulty increases, but we find no consistent correlation with domain difficulty.

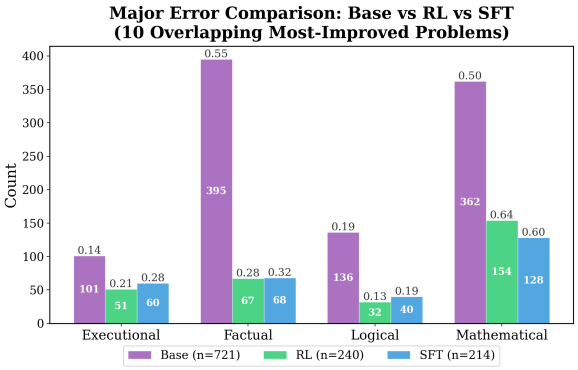

Figure 3: Overlap (10) between the RL and SFT top-20 most improved sets. Bars show absolute counts; values above bars show rates per incorrect rollout.

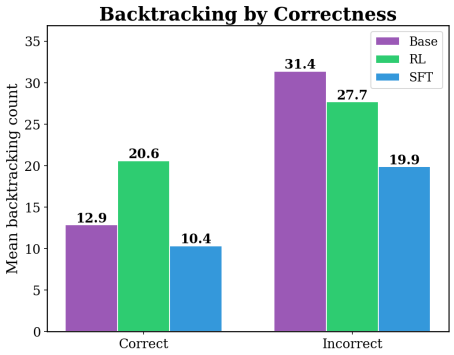

Figure 4: Number of backtracking occurrences per correct/incorrect response for base, RL, and SFT models.

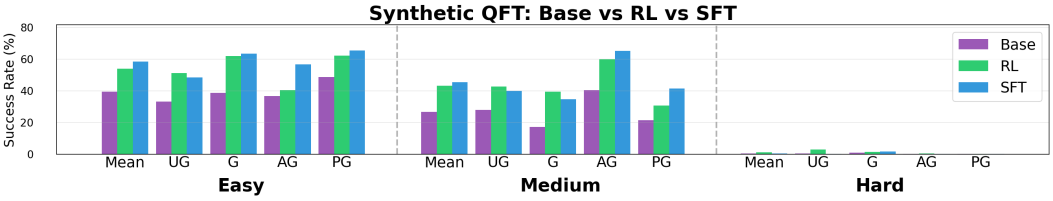

Figure 5: Validation performance trends (mean of 100 rollouts per problem) on base and fine-tuned models on UG, G, AG, PG subsets of the three synthetic QFT datasets.

4 Conclusion

In this work, we perform the first academic exploration of LLM learning dynamics during domain-specific fine-tuning in for theoretical physics problem-solving. Using our synthetic QFT data, we compared the performance gains of RL and SFT as well as explore the shifts in CoT errors, backtracking frequency, and find that operational difficulty is the primary driver of LLM solve rate on our tasks. We release our verifiable data pipeline, our QFT training datasets, and rejection sampled CoT datasets to aid further development in the academic community.

References

[1] Aitor Lewkowycz et al. *Solving Quantitative Reasoning Problems with Language Models*. 2022. arXiv: 2206.14858 [cs.CL]. URL: https://arxiv.org/abs/2206.14858.

[2] Ross Taylor et al. *Galactica: A Large Language Model for Science*. 2022. arXiv: 2211.09085 [cs.CL]. URL: https://arxiv.org/abs/2211.09085.

[3] Elliot Glazer et al. *FrontierMath: A Benchmark for Evaluating Advanced Mathematical Reasoning in AI*. 2024. arXiv: 2411.04872 [cs.AI]. URL: https://arxiv.org/abs/2411.04872.

[4] Daniel J. H. Chung et al. *Theoretical Physics Benchmark (TPBench) – a Dataset and Study of AI Reasoning Capabilities in Theoretical Physics*. 2025. arXiv: 2502.15815 [cs.LG]. URL: https://arxiv.org/abs/2502.15815.

[5] Kaiyue Feng et al. *PHYSICS: Benchmarking Foundation Models on University-Level Physics Problem Solving*. 2025. arXiv: 2503.21821 [physics.ed-ph]. URL: https://arxiv.org/abs/2503.21821.

[6] Xinyu Zhang et al. *PhysReason: A Comprehensive Benchmark towards Physics-Based Reasoning*. 2025. arXiv: 2502.12054 [cs.AI]. URL: https://arxiv.org/abs/2502.12054.

[7] Xinyu Pang et al. *Physics Reasoner: Knowledge-Augmented Reasoning for Solving Physics Problems with Large Language Models*. 2024. arXiv: 2412.13791 [cs.CL]. URL: https://arxiv.org/abs/2412.13791.

[8] Xin Xu et al. *UGPhysics: A Comprehensive Benchmark for Undergraduate Physics Reasoning with Large Language Models*. 2025. arXiv: 2502.00334 [cs.CL]. URL: https://arxiv.org/abs/2502.00334.

[9] Paul Tschisgale et al. "Evaluating GPT- and reasoning-based large language models on Physics Olympiad problems: Surpassing human performance and implications for educational assessment". In: *Physical Review Physics Education Research* 21.2 (Aug. 2025). ISSN: 2469-9896. DOI: 10.1103/6fmx-bsnl. URL: http://dx.doi.org/10.1103/6fmx-bsnl.

[10] Haining Pan et al. "Quantum many-body physics calculations with large language models". In: *Communications Physics* 8.1 (Jan. 2025). ISSN: 2399-3650. DOI: 10.1038/s42005-025-01956-y. URL: http://dx.doi.org/10.1038/s42005-025-01956-y.

[11] Xiansheng Cai et al. "Learning-at-Criticality in Large Language Models for Quantum Field Theory and Beyond". In: *Chinese Physics Letters* 42.12 (Nov. 2025), p. 120002. ISSN: 1741-3540. DOI: 10.1088/0256-307x/42/12/120002. URL: http://dx.doi.org/10.1088/0256-307X/42/12/120002.

[12] Kristian G. Barman et al. "Large physics models: towards a collaborative approach with large language models and foundation models". In: *The European Physical Journal C* 85.9 (Sept. 2025). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-025-14707-8. URL: http://dx.doi.org/10.1140/epjc/s10052-025-14707-8.

[13] Alfredo Guevara et al. *Single-minus gluon tree amplitudes are nonzero*. 2026. arXiv: 2602.12176 [hep-th]. URL: https://arxiv.org/abs/2602.12176.

[14] Paul Richmond et al. "FeynTune: large language models for high-energy theory". In: *Machine Learning: Science and Technology* 7.2 (Feb. 2026), p. 025012. ISSN: 2632-2153. DOI: 10.1088/2632-2153/ae47bb. URL: http://dx.doi.org/10.1088/2632-2153/ae47bb.

[15] Felix Grezes et al. *Building astroBERT, a language model for Astronomy Astrophysics*. 2021. arXiv: 2112.00590 [cs.CL]. URL: https://arxiv.org/abs/2112.00590.

[16] Tuan Dung Nguyen et al. *AstroLLaMA: Towards Specialized Foundation Models in Astronomy*. 2023. arXiv: 2309.06126 [astro-ph.IM]. URL: https://arxiv.org/abs/2309.06126.

[17] Ernest Perkowski et al. *AstroLLaMA-Chat: Scaling AstroLLaMA with Conversational and Diverse Datasets*. 2024. arXiv: 2401.01916 [astro-ph.IM]. URL: https://arxiv.org/abs/2401.01916.

[18] Yuan-Sen Ting et al. *AstroMLab 1: Who Wins Astronomy Jeopardy!?* 2024. arXiv: 2407.11194 [astro-ph.IM]. URL: https://arxiv.org/abs/2407.11194.

[19] Tijmen de Haan. "cosmosage: A natural-language assistant for cosmology". In: *Astronomy and Computing* 51 (Apr. 2025), p. 100934. ISSN: 2213-1337. DOI: 10.1016/j.ascom.2025.100934. URL: http://dx.doi.org/10.1016/j.ascom.2025.100934.

[20] Zhihong Shao et al. *DeepSeekMath: Pushing the Limits of Mathematical Reasoning in Open Language Models*. 2024. arXiv: 2402.03300 [cs.CL]. URL: https://arxiv.org/abs/2402.03300.

[21] DeepSeek-AI. *DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning*. 2025. arXiv: 2501.12948 [cs.CL]. URL: https://arxiv.org/abs/2501.12948.

[22] Qwen Team. *Qwen3 Technical Report*. 2025. arXiv: 2505.09388 [cs.CL]. URL: https://arxiv.org/abs/2505.09388.

[23] Thomas Hubert et al. "Olympiad-Level Formal Mathematical Reasoning with Reinforcement Learning". In: *Nature* (2025). URL: https://www.nature.com/articles/s41586-025-09833-y.

[24] A. Zee. *Quantum field theory in a nutshell*. 2003. ISBN: 978-0-691-14034-6.

[25] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5, 978-0-429-50355-9, 978-0-429-49417-8. DOI: 10.1201/9780429503559.

[26] Steven Weinberg. *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, June 2005. ISBN: 978-0-521-67053-1, 978-0-511-25204-4. DOI: 10.1017/CBO9781139644167.

[27] Voja Radovanovic. *Problem book in quantum field theory*. 2008.

[28] Ta-Pei [0000-0002-1137-0969] Cheng and Ling-Fong [0000-0002-8035-3329] Li. *Gauge Theory of Elementary Particle Physics*. Oxford, UK: Oxford University Press, 1984. ISBN: 978-0-19-851961-4, 978-0-19-851961-4.

[29] Hong Liu. *8.323 Relativistic Quantum Field Theory I*. Massachusetts Institute of Technology: MIT OpenCourseWare. Spring 2023. License: Creative Commons BY-NC-SA. 2023. URL: https://ocw.mit.edu/courses/8-323-relativistic-quantum-field-theory-i-spring-2023/.

[30] Hong Liu. *8.324 Relativistic Quantum Field Theory II*. Massachusetts Institute of Technology: MIT OpenCourseWare. Fall 2010. License: Creative Commons BY-NC-SA. 2010. URL: https://ocw.mit.edu/courses/8-324-relativistic-quantum-field-theory-ii-fall-2010/.

[31] Frank Wilczek. *8.325 Relativistic Quantum Field Theory III*. Massachusetts Institute of Technology: MIT OpenCourseWare. Spring 2003. License: Creative Commons BY-NC-SA. 2003. URL: https://ocw.mit.edu/courses/8-325-relativistic-quantum-field-theory-iii-spring-2003/.

[32] P. Blasiak et al. "Combinatorics and Boson normal ordering: A gentle introduction". In: (2007). DOI: 10.1119/1.2723799. eprint: arXiv:0704.3116.

[33] Jean-Yves Ollitrault. "Relativistic hydrodynamics for heavy-ion collisions". In: (2007). DOI: 10.1088/0143-0807/29/2/010. eprint: arXiv:0708.2433.

[34] Raphael Bousso. "TASI Lectures on the Cosmological Constant". In: (2007). DOI: 10.1007/s10714-007-0557-5. eprint: arXiv:0708.4231.

[35] Roberto Emparan and Harvey S. Reall. "Black Holes in Higher Dimensions". In: (2008). DOI: 10.12942/lrr-2008-6. eprint: arXiv:0801.3471.

[36] G. H. E. Duchamp et al. *Hopf Algebras in General and in Combinatorial Physics: a practical introduction*. 2008. eprint: arXiv:0802.0249.

[37] Stefan Vandoren and Peter van Nieuwenhuizen. *Lectures on instantons*. 2008. eprint: arXiv:0802.1862.

[38] Ruth Gregory. "Braneworld black holes". In: (2008). DOI: 10.1007/978-3-540-88460-6_7. eprint: arXiv:0804.2595.

[39] Christos Charmousis. "Higher order gravity theories and their black hole solutions". In: (2008). DOI: 10.1007/978-3-540-88460-6_8. eprint: arXiv:0805.0568.

[40] Torsten A. Ensslin, Mona Frommert, and Francisco S. Kitaura. "Information field theory for cosmological perturbation reconstruction and non-linear signal analysis". In: (2008). DOI: 10.1103/PhysRevD.80.105005. eprint: arXiv:0806.3474.

[41] Y. N. Srivastava, A. Widom, and L. Larsen. *A Primer for Electro-Weak Induced Low Energy Nuclear Reactions*. 2008. eprint: arXiv:0810.0159.

[42] Mihalis Dafermos and Igor Rodnianski. “Lectures on black holes and linear waves”. In: (2008). eprint: arXiv:0811.0354.

[43] Matthew Headrick. *A solution manual for Polchinski’s "String Theory"*. 2008. eprint: arXiv:0812.4408.

[44] M. Sami. “A primer on problems and prospects of dark energy”. In: (2009). eprint: arXiv:0904.3445.

[45] Daniel Baumann. *TASI Lectures on Inflation*. 2009. eprint: arXiv:0907.5424.

[46] C. R. Laumann et al. *Statistical mechanics of classical and quantum computational complexity*. 2010. eprint: arXiv:1009.1635.

[47] James M Yearsley. “A review of the decoherent histories approach to the arrival time problem in quantum theory”. In: (2010). DOI: 10.1088/1742-6596/306/1/012056. eprint: arXiv:1012.2575.

[48] Matthias Staudacher. “Review of AdS/CFT Integrability, Chapter III.1: Bethe Ansätze and the R-Matrix Formalism”. In: (2010). DOI: 10.1007/s11005-011-0530-9. eprint: arXiv:1012.3990.

[49] Eric Poisson, Adam Pound, and Ian Vega. “The motion of point particles in curved spacetime”. In: (2011). DOI: 10.12942/lrr-2011-7. eprint: arXiv:1102.0529.

[50] Frederik Denef. “TASI lectures on complex structures”. In: (2011). DOI: 10.1142/9789814350525_0007. eprint: arXiv:1104.0254.

[51] Jean Claude Dutailly. *Yang Mills model of interacting particles in the classical field theory*. 2011. eprint: arXiv:1109.3897.

[52] Mathieu Beau. “Feynman Path Integral approach to electron diffraction for one and two slits, analytical results”. In: (2011). DOI: 10.1088/0143-0807/33/5/1023. eprint: arXiv:1110.2346.

[53] G. S. Beloglazov et al. “Mathematical Physics : Problems and Solutions of The Students Training Contest Olympiad in Mathematical and Theoretical Physics (May 21st - 24th, 2010)”. In: (2011). eprint: arXiv:1110.4864.

[54] Edward Anderson. *The Problem of Time and Quantum Cosmology in the Relational Particle Mechanics Arena*. 2011. eprint: arXiv:1111.1472.

[55] Xiaoping Xu. *Algebraic Approaches to Partial Differential Equations*. 2012. eprint: arXiv:1205.6535.

[56] John R. Klauder. *Enhanced Quantum Procedures that Resolve Difficult Problems*. 2012. eprint: arXiv:1206.4017.

[57] C. P. Burgess. *The Cosmological Constant Problem: Why it’s hard to get Dark Energy from Micro-physics*. 2013. eprint: arXiv:1309.4133.

[58] M. Sami and R. Myrzakulov. “Late time cosmic acceleration: ABCD of dark energy and modified theories of gravity”. In: (2013). eprint: arXiv:1309.4188.

[59] Yu. L. Bolotin and I. V. Tanatarov. *A Thousand Problems in Cosmology: Horizons*. 2013. eprint: arXiv:1310.6329.

[60] Sevag Gharibian et al. “Quantum Hamiltonian Complexity”. In: (2014). DOI: 10.1561/0400000066. eprint: arXiv:1401.3916.

[61] Igor Mol. *Revisiting the Schwarzschild and the Hilbert-Droste Solutions of Einstein Equation and the Maximal Extension of the Latter*. 2014. eprint: arXiv:1403.2371.

[62] Ken Roberts and S. R. Valluri. “Solution of the quantum finite square well problem using the Lambert W function”. In: (2014). DOI: 10.1139/cjp-2016-0602. eprint: arXiv:1403.6685.

[63] Stefan M. Wild, Jason Sarich, and Nicolas Schunck. “Derivative-free optimization for parameter estimation in computational nuclear physics”. In: (2014). DOI: 10.1088/0954-3899/42/3/034031. eprint: arXiv:1406.5464.

[64] Alex S. Miranda et al. “Separable wave equations for gravitoelectromagnetic perturbations of rotating charged black strings”. In: (2014). DOI: 10.1088/0264-9381/32/23/235002. eprint: arXiv:1412.6312.

[65] Richard Brito, Vitor Cardoso, and Paolo Pani. “Superradiance – the 2020 Edition”. In: (2015). DOI: 10.1007/978-3-030-46622-0. eprint: arXiv:1501.06570.

[66] John D. Barrow. *Some Generalities About Generality*. 2015. eprint: arXiv:1503.05723.

[67] Atsushi Inoue. *Lectures on Super Analysis – Why necessary and What's that?* 2015. eprint: arXiv:1504.03049.

[68] Arnaud Baguet, Olaf Hohm, and Henning Samtleben. *$E_{6(6)}$ Exceptional Field Theory: Review and Embedding of Type IIB*. 2015. eprint: arXiv:1506.01065.

[69] Oscar J. C. Dias, Jorge E. Santos, and Benson Way. "Numerical Methods for Finding Stationary Gravitational Solutions". In: (2015). DOI: 10.1088/0264-9381/33/13/133001. eprint: arXiv:1510.02804.

[70] Alexander Pavlov. *Intrinsic time in Geometrodynamics: introduction and application to Friedmann cosmology*. 2016. eprint: arXiv:1606.09460.

[71] Ido Schaefer, Hillel Tal-Ezer, and Ronnie Kosloff. "Semi-global approach for propagation of the time-dependent Schrödinger equation for time-dependent and nonlinear problems". In: (2016). DOI: 10.1016/j.jcp.2017.04.017. eprint: arXiv:1611.06707.

[72] Vincent Desjacques, Donghui Jeong, and Fabian Schmidt. "Large-Scale Galaxy Bias". In: (2016). DOI: 10.1016/j.physrep.2017.12.002. eprint: arXiv:1611.09787.

[73] A. F. Reyes-Lega. "Some Aspects of Operator Algebras in Quantum Physics". In: (2016). DOI: 10.1142/9789814730884_0001. eprint: arXiv:1612.07718.

[74] L. Sebastiani, S. Vagnozzi, and R. Myrzakulov. "Mimetic gravity: a review of recent developments and applications to cosmology and astrophysics". In: (2016). DOI: 10.1155/2017/3156915. eprint: arXiv:1612.08661.

[75] John R. Klauder. *Enhanced Quantization: The Right way to Quantize Everything*. 2017. eprint: arXiv:1702.04713.

[76] Luca Fabbri. "Fundamental Theory of Torsion Gravity". In: (2017). DOI: 10.3390/universe7080305. eprint: arXiv:1703.02287.

[77] Shi-Ju Ran et al. "Lecture Notes of Tensor Network Contractions". In: (2017). DOI: 10.1007/978-3-030-34489-4. eprint: arXiv:1708.09213.

[78] Gábor Sárosi. "AdS_2 holography and the SYK model". In: (2017). DOI: 10.22323/1.323.0001. eprint: arXiv:1711.08482.

[79] Alexander Krikun. *Numerical Solution of the Boundary Value Problems for Partial Differential Equations. Crash course for holographer*. 2018. eprint: arXiv:1801.01483.

[80] Marius Oltean, Carlos F. Sopuerta, and Alessandro D. A. M. Spallicci. "Particle-without-Particle: a practical pseudospectral collocation method for linear partial differential equations with distributional sources". In: (2018). DOI: 10.1007/s10915-018-0873-9. eprint: arXiv:1802.03405.

[81] Dieter Lust and Ward Vleeshouwers. *Black Hole Information and Thermodynamics*. 2018. eprint: arXiv:1809.01403.

[82] Brendan Marsh. *Depictions of Quantum Reality in Kent's Interpretation of Quantum Theory*. 2018. eprint: arXiv:1811.08950.

[83] Sebastian Paeckel et al. "Time-evolution methods for matrix-product states". In: (2019). DOI: 10.1016/j.aop.2019.167998. eprint: arXiv:1901.05824.

[84] John Collins. *A new approach to the LSZ reduction formula*. 2019. eprint: arXiv:1904.10923.

[85] Ronald de Wolf. *Quantum Computing: Lecture Notes*. 2019. eprint: arXiv:1907.09415.

[86] Subenoy Chakraborty. *Treatise on Differential Geometry and its role in Relativity Theory*. 2019. eprint: arXiv:1908.10681.

[87] Thitipat Sainapha. *Gribov Ambiguity*. 2019. eprint: arXiv:1910.11659.

[88] José A. Pons and Daniele Viganò. "Magnetic, thermal and rotational evolution of isolated neutron stars". In: (2019). DOI: 10.1007/s41115-019-0006-7. eprint: arXiv:1911.03095.

[89] David Barral et al. "Quantum state engineering in arrays of nonlinear waveguides". In: (2020). DOI: 10.1103/PhysRevA.102.043706. eprint: arXiv:2005.07240.

[90] Marcel Balsiger et al. *Solutions to Problems at Les Houches Summer School on EFT*. 2020. eprint: arXiv:2005.08573.

[91] James M. Cline. *Advanced Concepts in Quantum Field Theory*. 2020. eprint: arXiv:2005.10241.

[92] Bijay Kumar Agrawal and Bhoomika Maheshwari. "Pairing, quasi-spin and seniority". In: (2020). DOI: 10.1140/epjst/e2020-000047-2. eprint: arXiv:2006.03909.

[93] Emmanuel Grenier and Toan T. Nguyen. *Stability of shear flows near a boundary*. 2020. eprint: arXiv:2008.00887.

[94] Arthur Hebecker. *Lectures on Naturalness, String Landscape and Multiverse*. 2020. eprint: arXiv:2008.10625.

[95] U. Boscain, M. Sigalotti, and D. Sugny. "Introduction to the Pontryagin Maximum Principle for Quantum Optimal Control". In: (2020). DOI: 10.1103/PRXQuantum.2.030203. eprint: arXiv:2010.09368.

[96] Y. T. Albert Law. "A Compendium of Sphere Path Integrals". In: (2020). DOI: 10.1007/JHEP12(2021)213. eprint: arXiv:2012.06345.

[97] David E. Bernal, Sridhar Tayur, and Davide Venturelli. *Quantum Integer Programming (QuIP) 47-779: Lecture Notes*. 2020. eprint: arXiv:2012.11382.

[98] Christian Brand et al. "Single-, double-, and triple-slit diffraction of molecular matter-waves". In: (2021). DOI: 10.1119/5.0058805. eprint: arXiv:2108.06565.

[99] Hector Jose Morrell, Anika Zaman, and Hiu Yung Wong. "Step-by-Step HHL Algorithm Walkthrough to Enhance the Understanding of Critical Quantum Computing Concepts". In: (2021). DOI: 10.1109/ACCESS.2023.3297658. eprint: arXiv:2108.09004.

[100] Arnold Neumaier. *Quantum tomography explains quantum mechanics*. 2021. eprint: arXiv:2110.05294.

[101] Raghav G. Jha. "Introduction to Monte Carlo for Matrix Models". In: (2021). DOI: 10.21468/SciPostPhysLectNotes.46. eprint: arXiv:2111.02410.

[102] Andrea Bevilacqua and Jerzy Kowalski-Glikman. *Doubly Special Relativity and Relative Locality*. 2022. eprint: arXiv:2203.04091.

[103] R. Arouca, Andrea Cappelli, and T. H. Hansson. "Quantum Field Theory Anomalies in Condensed Matter Physics". In: (2022). DOI: 10.21468/SciPostPhysLectNotes.62. eprint: arXiv:2204.02158.

[104] Sashwat Tanay. *Integrability and action-angle-based solution of the post-Newtonian BBH system (lecture notes)*. 2022. eprint: arXiv:2206.05799.

[105] M. Le Delliou. *Advanced General Relativity Notes*. 2022. eprint: arXiv:2208.02506.

[106] Ilya L. Shapiro. "The background information about perturbative quantum gravity". In: (2022). DOI: 10.1007/978-981-19-3079-9_8-1. eprint: arXiv:2210.12319.

[107] Arnau Bas i Beneito, Gianluca Calcagni, and Lesław Rachwał. *Classical and quantum nonlocal gravity*. 2022. eprint: arXiv:2211.05606.

[108] Domenico Giordano, Pierluigi Amodio, and Felice Iavernaro. *A didactically motivated reexamination of a particle's quantum mechanics with square-well potentials*. 2022. eprint: arXiv:2211.06269.

[109] Ievgen Dubovyk, Janusz Gluza, and Gabor Somogyi. *Mellin-Barnes Integrals: A Primer on Particle Physics Applications*. 2022. eprint: arXiv:2211.13733.

[110] Daniel Green. *TASI Lectures on Cosmic Signals of Fundamental Physics*. 2022. eprint: arXiv:2212.08685.

[111] Luca Ciambelli. *From Asymptotic Symmetries to the Corner Proposal*. 2022. eprint: arXiv:2212.13644.

[112] Deep Ray, Orazio Pinti, and Assad A. Oberai. *Deep Learning and Computational Physics (Lecture Notes)*. 2023. eprint: arXiv:2301.00942.

[113] Raghav G. Jha. *Notes on Quantum Computation and Information*. 2023. eprint: arXiv:2301.09679.

[114] Alexei A. Deriglazov. "Lagrangian and Hamiltonian formulations of asymmetric rigid body, considered as a constrained system". In: (2023). DOI: 10.1088/1361-6404/ace80d. eprint: arXiv:2301.10741.

[115] Matthew Reece. *TASI Lectures: (No) Global Symmetries to Axion Physics*. 2023. eprint: arXiv:2304.08512.

[116] Pietropaolo Frisoni. *Introduction to Loop Quantum Gravity: Rovelli's lectures on LQG*. 2023. eprint: arXiv:2305.12215.

[117] Yasunori Lee and Keita Kanno. *Modal analysis on quantum computers via qubitization*. 2023. eprint: arXiv:2307.07478.

[118] John L. Friedman. *Notes on Gravitational Physics*. 2023. eprint: arXiv:2308.09826.

[119] Marcelo M. Disconzi. "Recent developments in mathematical aspects of relativistic fluids". In: (2023). DOI: 10.1007/s41114-024-00052-x. eprint: arXiv:2308.09844.

[120] Pietro G. Fré. *Lectures on resolutions à la Kronheimer of orbifold singularities, McKay quivers for Gauge Theories on D3 branes, and the issue of Ricci flat metrics on the resolved three-folds*. 2023. eprint: arXiv:2308.14022.

[121] Sari Ghanem. *The global stability of the Minkowski space-time solution to the Einstein-Yang-Mills equations in higher dimensions*. 2023. eprint: arXiv:2310.07954.

[122] Sari Ghanem. *Exterior stability of the* $(1+3)$*-dimensional Minkowski space-time solution to the Einstein-Yang-Mills equations*. 2023. eprint: arXiv:2310.08196.

[123] Vasileios Evangelakos, Emmanuel Paspalakis, and Dionisis Stefanatos. "Optimal STIRAP shortcuts using the spin to spring mapping". In: (2023). DOI: 10.1103/PhysRevA.107.052606. eprint: arXiv:2312.16643.

[124] Roman Berens, Trevor Gravely, and Alexandru Lupsasca. "Gravitational Waves on Kerr Black Holes I: Reconstruction of Linearized Metric Perturbations". In: (2024). DOI: 10.1088/1361-6382/ad6c9c. eprint: arXiv:2403.20311.

[125] Mohamed Sabba. *A quantum Pascal pyramid and an extended de Moivre-Laplace theorem*. 2024. eprint: arXiv:2404.03560.

[126] Peixiang Ji and Lijing Shao. "Scalar Dark Energy Models and Scalar-Tensor Gravity: Theoretical Explanations for the Accelerated Expansion of Present Universe". In: (2024). DOI: 10.1088/1572-9494/ad5aeb. eprint: arXiv:2406.04954.

[127] Darryl D. Holm. *31 Lectures on Geometric Mechanics*. 2024. eprint: arXiv:2408.09564.

[128] Qiushi Liu et al. "Fully-Optimized Quantum Metrology: Framework, Tools, and Applications". In: (2024). DOI: 10.1002/qute.202400094. eprint: arXiv:2409.07068.

[129] Fernando Quevedo and Andreas Schachner. *Cambridge Lectures on The Standard Model*. 2024. eprint: arXiv:2409.09211.

[130] Christian Northe. *Young Researchers School 2024 Maynooth: Lectures on CFT, BCFT and DCFT*. 2024. eprint: arXiv:2411.03381.

[131] Francisco Duque. *New Horizons for Psi: Extreme-Mass-Ratio Inspirals in Fundamental Fields*. 2024. eprint: arXiv:2412.08711.

[132] Miguel Correia et al. *Records from the S-Matrix Marathon: Gravitational Physics from Scattering Amplitudes*. 2024. eprint: arXiv:2412.11649.

[133] V. A. Babenko and A. V. Nesterov. "Study of the Quartic Anharmonic Oscillator Using the System's Wave Function Expansion in the Oscillator Basis". In: (2025). DOI: 10.1142/s0217751x25500733. eprint: arXiv:2505.06317.

[134] Nicolas Perez and Eyal Heifetz. *Landau levels in a time-dependent magnetic field: the Madelung fluid perspective*. 2025. eprint: arXiv:2505.08460.

[135] Muhammad Faryad. *Lecture Notes on Quantum Algorithms*. 2025. eprint: arXiv:2507.11565.

[136] Zohreh Davoudi. *TASI/CERN/KITP Lecture Notes on "Toward Quantum Computing Gauge Theories of Nature"*. 2025. eprint: arXiv:2507.15840.

[137] Simon Badger et al. "Scattering Amplitudes in Quantum Field Theory". In: (2023). DOI: 10.1007/978-3-031-46987-9. eprint: arXiv:2306.05976.

[138] Benjamin D. Wandelt, Eric Hivon, and Krzysztof M. Gorski. "The Pseudo-C_l method: Cosmic microwave background anisotropy power spectrum statistics for high precision cosmology". In: (2000). DOI: 10.1103/PhysRevD.64.083003. eprint: arXiv:astro-ph/0008111.

[139] Douglas C Heggie. *Escape in Hill's Problem*. 2000. eprint: arXiv:astro-ph/0011294.

[140] Andrew R. Zentner. "The Excursion Set Theory of Halo Mass Functions, Halo Clustering, and Halo Growth". In: (2006). DOI: 10.1142/S0218271807010511. eprint: arXiv:astro-ph/0611454.

[141] Massimo Giovannini. "Why CMB physics?" In: (2007). DOI: 10.1142/S0217751X07036919. eprint: arXiv:astro-ph/0703730.

[142] Johannes Voit. "One-Dimensional Fermi liquids". In: (1995). DOI: 10.1088/0034-4885/58/9/002. eprint: arXiv:cond-mat/9510014.

[143] Ian Affleck. "Conformal Field Theory Approach to the Kondo Effect". In: (1995). eprint: arXiv:cond-mat/9512099.

[144] H. Saleur. *Lectures on Non Perturbative Field Theory and Quantum Impurity Problems*. 1998. eprint: arXiv:cond-mat/9812110.

[145] Christian Klein. "Exact relativistic treatment of stationary counter-rotating dust disks I: Boundary value problems and solutions". In: (2001). DOI: 10.1103/PhysRevD.63.064033. eprint: arXiv:gr-qc/0102083.

[146] Sascha Husa. "Numerical relativity with the conformal field equations". In: (2002). DOI: 10.1007/3-540-36973-2_9. eprint: arXiv:gr-qc/0204057.

[147] Geoffrey Compere. *An introduction to the mechanics of black holes*. 2006. eprint: arXiv:gr-qc/0611129.

[148] Thomas Mohaupt. "Black Holes in Supergravity and String Theory". In: (2000). DOI: 10.1088/0264-9381/17/17/303. eprint: arXiv:hep-th/0004098.

[149] Chun-Ming Chiang and Choon-Lin Ho. "Charged particles in external fields as physical examples of quasi-exactly solvable models: a unified treatment". In: (2000). DOI: 10.1103/PhysRevA.63.062105. eprint: arXiv:hep-th/0011005.

[150] M. Chaves. "An introduction to generalized Yang-Mills theories". In: (2001). eprint: arXiv:hep-th/0102055.

[151] Justin R. David, Gautam Mandal, and Spenta R. Wadia. "Microscopic Formulation of Black Holes in String Theory". In: (2002). DOI: 10.1016/S0370-1573(02)00271-5. eprint: arXiv:hep-th/0203048.

[152] Erick J. Weinberg and Piljin Yi. "Magnetic Monopole Dynamics, Supersymmetry, and Duality". In: (2006). DOI: 10.1016/j.physrep.2006.11.002. eprint: arXiv:hep-th/0609055.

[153] James M. Cline. *String Cosmology*. 2006. eprint: arXiv:hep-th/0612129.

[154] Iosif Bena and Nicholas P. Warner. "Black Holes, Black Rings and their Microstates". In: (2007). DOI: 10.1007/978-3-540-79523-0. eprint: arXiv:hep-th/0701216.

[155] Alexander A. Migdal. "Loop Equation and Area Law in Turbulence". In: (1993). DOI: 10.1142/S0217751X94000558. eprint: arXiv:hep-th/9310088.

[156] Koichi Yamawaki. *Zero-Mode Problem on the Light Front*. 1998. eprint: arXiv:hep-th/9802037.

[157] Bertrand Eynard. *Large N asymptotics of orthogonal polynomials, from integrability to algebraic geometry*. 2005. eprint: arXiv:math-ph/0503052.

[158] Peter W. Shor. *Introduction to Quantum Algorithms*. 2000. eprint: arXiv:quant-ph/0005003.

[159] P. Deuar and P. D. Drummond. "Gauge P-representations for quantum-dynamical problems: Removal of boundary terms". In: (2002). DOI: 10.1103/PhysRevA.66.033812. eprint: arXiv:quant-ph/0203025.

[160] Y. S. Kim and Marilyn E. Noz. "Coupled oscillators, entangled oscillators, and Lorentz-covariant harmonic oscillators". In: (2005). DOI: 10.1088/1464-4266/7/12/005. eprint: arXiv:quant-ph/0502096.

[161] A. Deloff. "Semi-spectral Chebyshev method in Quantum Mechanics". In: (2006). DOI: 10.1016/j.aop.2006.07.004. eprint: arXiv:quant-ph/0606100.

[162] L. P. Gilbert et al. "Playing Quantum Physics Jeopardy with zero-energy eigenstates". In: (2006). DOI: 10.1119/1.2238884. eprint: arXiv:quant-ph/0606196.

[163] Nathaniel S. Woodward et al. *Fine-Tuning Small Reasoning Models for Quantum Field Theory*. 2026. arXiv: 2604.18936 [cs.LG]. URL: https://arxiv.org/abs/2604.18936.

[164] OpenAI. *gpt-oss-120b gpt-oss-20b Model Card*. 2025. arXiv: 2508.10925 [cs.CL]. URL: https://arxiv.org/abs/2508.10925.

[165] Qwen Team. *Qwen3.5: Towards Native Multimodal Agents*. 2026. URL: https://qwen.ai/blog?id=qwen3.5.

[166] Zheng Yuan et al. *Scaling Relationship on Learning Mathematical Reasoning with Large Language Models*. 2023. eprint: arXiv:2308.01825.

[167] Guangming Sheng et al. "HybridFlow: A Flexible and Efficient RLHF Framework". In: *arXiv preprint arXiv: 2409.19256* (2024).

[168] Bingxiang He et al. "JustRL: Scaling a 1.5 B LLM with a Simple RL Recipe". In: *arXiv preprint arXiv:2512.16649* (2025).

[169] Axolotl maintainers and contributors. *Axolotl: Open Source LLM Post-Training*. 2023. URL: https://github.com/axolotl-ai-cloud/axolotl.