Formalizing Physics: Automation, Presentation and Foundation Issues

Cezary Kaliszyk[1], Josef Urban[2](✉), Umair Siddique[3], Sanaz Khan-Afshar[3], Cvetan Dunchev[4], and Sofiène Tahar[3]

[1] University of Innsbruck, Innsbruck, Austria
[2] Radboud University, Nijmegen, The Netherlands
Josef.Urban@gmail.com
[3] Concordia University, Montreal, Canada
[4] University of Bologna, Bologna, Italy

Abstract. In this paper, we report our first experiments in using learning-assisted automated reasoning for the formal analysis of physical systems. In particular, we investigate the performance of automated proofs as compared to interactive ones done in HOL for the verification of ray and electromagnetic optics. Apart from automation, we also provide brief initial exploration of more general issues in formalization of physics, such as its presentation and foundations.

1 Introduction: Formalization, Automation and Physics

Twenty years after the QED Manifesto [1], there is an encouraging progress in building computer-understandable and formally verified mathematical corpora. Large projects in mathematics include the completed formal proofs of the Kepler conjecture (Flyspeck) [8], the Odd Order theorem [7], the Four Color theorem [6], and verification of more than a half of the Compendium of Continuous Lattices textbook [3]. Verification of the seL4 kernel [15] and the CompCert compiler [17] show comparable progress in full-scale verification of complicated software. Such projects are often linked to advances in verification technology, and in particular to strong automation [9,11,16] that allows less verbose formal proofs and increases the general understanding intelligence of the formal proof assistants.

This ongoing progress brings closer the possibility of eventually expressing in a computer-understandable form all of today's scientific knowledge, and in particular the vast knowledge accumulated by exact sciences such as physics. Such a *Formalization of Physics* (FOP) project raises a number of interesting issues, ranging from philosophical and theoretical to very practical ones, on a scale that may eventually dwarf the current applications of formal verification. Just optical components are today a basis of a growing multi-billion business, technologies involving quantum-level phenomena become more and more important, the safety of space/air flight and other means of transport (particularly self-driving) may greatly benefit from formal treatment, and perhaps even more some of the big and dangerous "prides" of modern physics such as nuclear power

M. Kerber et al. (Eds.): CICM 2015, LNAI 9150, pp. 288–295, 2015.
DOI: 10.1007/978-3-319-20615-8_19

plants, tokamaks, and large hadron colliders. An interesting multidisciplinary problem is the formal analysis of engineering systems which requires formalized theories of Physics, Probability and Information Theory.

One of the first practical hurdles in FOP is the unfamiliarity with theorem proving in the Physics community. An attractive step that may reduce this gap is to wrap the internal complexities of tactical theorem proving systems in powerful high-level automation, user-friendly interfaces, collaborative reasoning platforms and proof advice systems. The main concrete contribution of this paper is to describe the first experiments in deploying and using such strong automation – the HOLyHammer system [12] – over the first formal physics developments. Section 2 briefly describes such projects in the area of *Formal Optics* (Formalization of Physics) and Sect. 3 describes first steps and experiments in using HOLyHammer for these developments. This initial experience leads us to discuss in Sect. 4 some wider and more concrete issues related to the present and future FOP project(s).

2 Formal Optics

Optical systems are becoming increasingly important by resolving many bottlenecks in todays communication, aerospace and biomedical systems. However, given the continuous nature of optics, the inability to efficiently analyze optical system models using traditional paper-and-pencil and computer simulation approaches sets limits especially in safety-critical applications.

In 2009, a project[1] was started at the Hardware Verification Group (HVG) of Concordia University in order to build a comprehensive framework for the formal analysis of optical systems. The project can be divided into three sub-projects:

- Formalization of Ray Optics in which light is considered as a ray, i.e., a simple geometrical line.
- Formalization of Electromagnetic Optics in which light is characterized as electromagnetic waves.
- Formalization of Quantum Optics in which light is characterized as a stream of photons.

Currently, fundamentals of ray optics, electromagnetic optics and quantum optics have been formalized [14] in HOL Light. This allowed the formal verification of some interesting and safety-critical optical systems such as optical resonators [19], laser resonator [13] and optical quantum flip gate [18]. In the sequel, we explore automation and presentation issues of these projects.

3 HOLyHammer and Formal Optics

HOLyHammer [12] is a recently developed online AI/ATP system for assisting formal (computer-understandable) verification done in HOL Light. The service

[1] http://hvg.ece.concordia.ca/projects/optics/.

allows its users to upload and automatically process an arbitrary formal development (project) based on HOL Light, and to attack arbitrary conjectures that use the concepts defined in some of the uploaded projects. The service uses several automated theorem provers (ATPs) combined with several premise selection methods trained on all the project proofs. The ITP (interactive theorem prover) and ATP proof data and theorems from different (possibly incompatible) projects and their versions are pooled together using a recursive content-based (MD5) naming of symbols and theorems, providing a large base of proofs to learn from. Authorized users can upload a new project against an arbitrary existing project (saved as standard and proof-recording checkpointed images), allowing fast processing of HOL Light projects that import large libraries such as the Multivariate Analysis. The system also provides version control and heuristic HTML-ization (cross-linking) of the uploaded projects. Users can ask parallel asynchronous queries to the service either from its web interface or directly from the HOL Light mode for Emacs. Below we describe the steps to deploy and test HOLyHammer for Formal Optics.

3.1 Deployment

We have streamlined the HOLyHammer installation and deployed it on a faster dedicated machine with 12 hyperthreading 2.6 GHz Xeons in Canada (U. of Alberta), which was serving so far the users of the similar online service for Mizar [9]. The HVG members were given access rights to upload their developments there, to update them, and their Emacs mode was configured to ask queries to this server. Such a dedicated/local HOLyHammer installation is now quite easy and we hope that more users will use this option and we will eventually build a network of such online "hammer" installations that will further synchronize between them their proof data, projects, CPU-load, etc., in the spirit of large distributed formal wikis [2].

3.2 Experiments with Complete Automation

We have measured the strength of the HOLyHammer automation on the Ray (Ray Optics) and EMF (Electromagnetic Optics) formalizations. These two projects are both based on HOL Light's Complex Multivariate Analysis, and they together contain 482 proved toplevel theorems and 125 definitions.[2] Table 1 shows the performance of 11 ATPs in proving the 482 theorems from their recorded HOL Light dependencies, and Table 2 shows the performance of various strategies that combine the three best ATPs with premise selection using learning from previous proofs[3]. The learning method used in all cases was distance-weighted k-nearest neighbor with IDF-weighted normalized term-based features [10]. The results are encouraging: the combined strength of the methods reaches nearly 50 % (239

[2] Many definitions are just abbreviations introducing proper physics terminology.

[3] The complete set of ATP inputs generated by HOLyHammer and the corresponding ATP outputs are available at http://cl-informatik.uibk.ac.at/~cek/cicm15/data.tgz.

problems solved) in the first scenario when the premises are chosen by the user. 236 of these problems are already solved by one of the best three ATPs (Epar, Vampire 3.0, and Z3 4.0). The performance is 45 % (217 problems solved) in the fully automated mode when the relevant premises are chosen automatically by machine learning, and seven different combinations of premise selection and ATPs are needed for this. Note that there are 105 problems that Paradox found counter-satisfiable. This means that the incompleteness of the currently used HOL-to-FOL translation shows quite considerably on these problems, making more complete encodings an interesting problem to address in this context.

Table 1. ATP re-proving with 300 s time limit on the 482 Emf and Ray top-level problems

Prover	Theorem (%)	CounterSat (%)
Epar	219 (45.436)	0
Vampire 3.0	210 (43.568)	0
Z3 4.0	210 (43.568)	0
CVC4 1.3	201 (41.701)	0
Vampire 2.6	198 (41.079)	0
E 1.8	189 (39.212)	0
SPASS 3.5	154 (31.950)	0
Metis 2.3	152 (31.535)	0
iProver 1	116 (24.066)	0
Prover9 09.11a	114 (23.651)	0
Paradox 4.0	0 (0.000)	105 (21.784)
any	239 (49.585)	105 (21.784)

A brief review of the fully automatically solved problems shows that HOLyHammer is particularly useful in automating proofs about complex vectors (used in the representation of planar waves) in Electromagnetic Optics, for example the following relation[4] between collinearity and orthogonality of complex vectors is proved by Epar using 17 other previous theorems:

```
∀x y:complex^N.
  collinear_cvectors x y ∧ ¬(x=cvector_zero) ∧ ¬(y=cvector_zero)
    ⟹ ¬(corthogonal x y)
```

An example of a fully automatically proved lemma in Ray Optics is a statement[5] about the stability of an optical resonator (represented by its ray transfer matrix) under certain conditions. In this case the AI/ATP found a relevant special lemma where most of the hard proving work was done, and which together with six auxiliary lemmas can be used to automatically prove the more general statement:

[4] http://mizar.cs.ualberta.ca/hh/ses/Emf202/cvectors.html#CORTHOGONAL_COLLINEAR_CVECTORS.

[5] http://mizar.cs.ualberta.ca/hh/ses/Ray203/resonator.html#STABILITY_LEMMA_GENERAL_SYM.

Table 2. ATP proving with k-NN premise selection and 300 s time limit on the 482 Emf and Ray top-level problems

Prover	Premises	Theorem (%)
Epar	1024	170 (35.565)
Epar	128	155 (32.158)
Vampire 3.0	128	121 (25.104)
Vampire 3.0	1024	119 (24.895)
E 1.8	128	104 (21.577)
Z3 4.0	128	103 (21.369)
Epar	32	102 (21.162)
Vampire 3.0	32	92 (19.087)
E 1.8	32	91 (18.880)
Z3 4.0	32	89 (18.465)
E 1.8	1024	68 (14.226)
Z3 4.0	1024	64 (13.389)
any		217 (45.021)

```
∀ (M:real^2^2) xi thetai.
 (det (M) = &1) ∧ ( −&1 < (M$1$1 + M$2$2) / &2) ∧ (M$1$1 + M$2$2)/&2 < &1
    ⟹ ∃(Y:real^2). ∀n.
        abs (((M pow 2) pow n ** vector [xi; thetai])$1) ≤ Y$1
      ∧ abs (((M pow 2) pow n ** vector [xi; thetai])$2) ≤ Y$2
```

3.3 Linking to Informal Physics Explanations

Formal mathematics as a science enjoys a remarkable property: it is in some sense fully "understood" by machines. Computers can correctly parse the formal definitions and statements, verify the proofs, and sometimes even find proofs independently of humans, regardless of any possible motivation and underlying intuition ivolved in proposing the definitions, theorems, proofs and theories. In this sense, formal mathematics is completely self-explanatory. While (physical) intuition may play varied part in formulation of various theories, such theories as formal mathematical objects are independent and decoupled from their (possible) underlying intuition. It is not unusual that for some abstract theory a new application is found, which has very little in common with the original intuition. Similarly, the popular term "abstract nonsense" refers to abstract arguments (e.g., in category theory) which are hard to link to any particular intuition. While some physicists (notably Feynman) criticized such decoupling from physical intuition as harmful, it is a fact that many mathematicians (to say nothing about computers) do mathematics without such links.

We believe that here is a real difference between (formal/abstract) mathematics and physics, and this difference really needs to be addressed by appropriate tools assisting formalization of physics. In physics, there is always first some underlying intuition about (part of) the real world, and this intuition is more or less perfectly captured by various abstract mathematical models. An important part of physics is the *informal* understanding of the (intended) correspondence between the physical phenomena and their formal models. This understanding however is not (yet) part of the actual formal code. In particular, those of us who are not experts in optics have found it significantly harder to understand some of the formal definitions modelling the physical systems and phenomena. While abstract concepts like sets, quasigroups, categories and topological spaces are acceptable to mathematicians as just such abstract concepts described by their formal definitions, taking an "optical resonator" to be just its formal definition does not seem to be right, because it forgets the "real" physical phenomenon that is linked to (and motivating) the particular choice of the formal model.

A solution that does not require much work from the formalizers (and which can even be done later by others) is to allow special comments in the formal text, that are during the HTML-ization turned into cross-links to informal explanations, in our case to Wikipedia. Such cross-links can be also harvested from the formalizations, thus providing an informal overview (and in some sense also high-level semantic anchors) of the physics topics dealt with in the formal code. About 20 such Wikipedia annotations have been inserted into the Ray Optics formalization,[6] making the resulting HTML presentation considerably easier to understand for some of us. Another very interesting informal resource that could provide such semantic anchors are the three volumes of Feynman's lectures that have been recently published online in a form that makes use of state-of-the-art informal presentation technologies such as MathJax.[7]

4 Some Issues and Considerations in Formal Physics

The tighter link between the formal mathematical theory and its underlying (physical) intuition is likely just one of several interesting differences between formalization of physics and formalization of mathematics. Clearly, the most obvious theoretical issue is whether it is possible to consistently formalize the whole of physics at all, and what should be the ultimate foundational framework for such formalization. For example, Beeson in [4] briefly derives (what he calls) a contradiction between quantum theory and general relativity that is apparently well-known to physicists, and which can perhaps be understood as quantum physics breaking some of the assumptions of general relativity about all possible worlds being regular solutions to Einstein's equations. There are probably several answers to this famous problem by current theoretical physics, the best-known involving various string and superstring theories for which we still lack enough experimental evidence.

This however just brings up the main issue with physics: it is about modelling the real world "well enough" which we do not fully know and probably

[6] See, e.g., http://mizar.cs.ualberta.ca/hh/ses/Ray203/resonator.html.

[7] http://www.feynmanlectures.caltech.edu/.

never will. As already the several approaches to the formalization of optics show, there are typically several models of the same phenomena. These models will often be "almost compatible" in terms of their predictions when used on their intended domain, e.g., the more complicated electromagnetic optics model will largely agree with the simpler ray optics model on an important class of optics problems. As one goes farther away from this class of problems, the predictions of these two models will disagree more and more. Some models designed for very different phenomena, such as the quantum-theoretical and relativistic, might quickly yield hard contradictions as soon as one tries to use both of them at once. A proper foundational framework should make such relations between the models as explicit as possible (e.g., by theorems exhibiting the asymptotic relations between the models and/or their incompatibilities and scope, perhaps enhancing by such explicit relations formalization frameworks such as Little Theories and Realms [5]), so that one can consistently and automatically combine the knowledge contained in them in the same way as the current large-theory AI/ATP methods do over large mathematical corpora.

An interesting related issue is to what extent such careful "theory engineering" could assist, emulate, or even replace "proper" mathematical solutions to inconsistencies in physics, such as the Dirac delta "function" (made consistent later by Schwartz's distributions), the physics way of treating the infinitesimals (made consistent by Robinson's ultraproduct models) or various approaches to counting with infinities (regularization, renormalization) in Feynman's diagrams.

There are also many practical issues and tasks that are already visible in our experiments. Physics is a heavy user of computation, and the pragmatic approach used sometimes by the HVG group is to just trust the results of computer algebra systems (e.g., using Mathematica to compute the numerical eigenvalues of the waveguide when there is no closed form solution [14]), temporarily adding them as axioms [14]. This is going to be a rich source of research problems for Calculemus-style projects, SMT solving, systems like MetiTarski, etc. In short, we suggest FOP as a rather exciting and very large and rewarding research topic whose automation, foundations and presentation issues will keep the formalization community busy in the next years, hopefully greatly expanding its current borders and methods.

Acknowledgements. Kaliszyk was supported by the Austrian Science Fund (FWF) grant P26201.

References

1. Boyer, R., et al.: The QED Manifesto. In: Bundy, Alan (ed.) CADE 1994. LNCS, vol. 814, pp. 238–251. Springer, Heidelberg (1994)
2. Alama, J., Brink, K., Mamane, L., Urban, J.: Large formal wikis: issues and solutions. In: Davenport, J.H., Farmer, W.M., Urban, J., Rabe, F. (eds.) MKM/Calculemus 2011. LNCS, vol. 6824, pp. 133–148. Springer, Heidelberg (2011)
3. Bancerek, G., Rudnicki, P.: A compendium of continuous lattices in MIZAR. J. Autom. Reasoning **29**(3–4), 189–224 (2002)

4. Beeson, M.: Constructivity, computability, and the continuum. In: Essays on the Foundations of Mathematics and Logic, Polimetrica, Milan, vol. 2 (2005)
5. Carette, J., Farmer, W.M., Kohlhase, M.: Realms: A structure for consolidating knowledge about mathematical theories. In: Watt, S.M., Davenport, J.H., Sexton, A.P., Sojka, P., Urban, J. (eds.) CICM 2014. LNCS, vol. 8543, pp. 252–266. Springer, Heidelberg (2014)
6. Gonthier, G.: The four colour theorem: engineering of a formal proof. In: Kapur, D. (ed.) ASCM 2007. LNCS (LNAI), vol. 5081, p. 333. Springer, Heidelberg (2008)
7. Gonthier, G.: Engineering mathematics: the odd order theorem proof. In: Giacobazzi, R., Cousot, R. (eds.) The 40th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2013, Rome, Italy, 23–25 January, pp. 1–2. ACM (2013)
8. Hales, T.: Dense Sphere Packings: A Blueprint for Formal Proofs. London Mathematical Society Lecture Note Series, vol. 400. Cambridge University Press, Cambridge (2012)
9. Kaliszyk, C., Urban, J.: MizAR 40 for Mizar 40. CoRR, abs/1310.2805 (2013)
10. Kaliszyk, C., Urban, J.: Stronger automation for Flyspeck by feature weighting and strategy evolution. In: Blanchette, J.C., Urban, J. (eds.) PxTP 2013. EPiC Series, vol. 14, pp. 87–95. EasyChair (2013)
11. Kaliszyk, C., Urban, J.: Learning-assisted automated reasoning with Flyspeck. J. Autom. Reasoning **53**(2), 173–213 (2014)
12. Kaliszyk, C., Urban, J.: HOL(y)Hammer: Online ATP service for HOL Light. Math. Comput. Sci. **9**(1), 5–22 (2015)
13. Khan-Afshar, S., Hasan, O., Tahar, S.: Formal analysis of electromagnetic optics. In: Proceedings of SPIE, vol. 9193, pp. 91930A–91930A-14 (2014)
14. Khan-Afshar, S., Siddique, U., Mahmoud, M.Y., Aravantinos, V., Seddiki, O., Hasan, O., Tahar, S.: Formal analysis of optical systems. Math. Comput. Sci. **8**(1), 39–70 (2014)
15. Klein, G., Huuck, R., Schlich, B.: Operating system verification. J. Autom. Reasoning **42**(2–4), 123–124 (2009)
16. Kühlwein, D., Blanchette, J.C., Kaliszyk, C., Urban, J.: MaSh: machine learning for sledgehammer. In: Blazy, S., Paulin-Mohring, C., Pichardie, D. (eds.) ITP 2013. LNCS, vol. 7998, pp. 35–50. Springer, Heidelberg (2013)
17. Leroy, X.: Formal verification of a realistic compiler. Commun. ACM **52**(7), 107–115 (2009)
18. Mahmoud, M.Y., Aravantinos, V., Tahar, S.: Formal verification of optical quantum flip gate. In: Klein, G., Gamboa, R. (eds.) ITP 2014. LNCS, vol. 8558, pp. 358–373. Springer, Heidelberg (2014)
19. Siddique, U., Aravantinos, V., Tahar, S.: Formal stability analysis of optical resonators. In: Brat, G., Rungta, N., Venet, A. (eds.) NFM 2013. LNCS, vol. 7871, pp. 368–382. Springer, Heidelberg (2013)