A peer-reviewed version of this preprint was published in PeerJ on 2 January 2017.

<u>View the peer-reviewed version</u> (peerj.com/articles/cs-103), which is the preferred citable publication unless you specifically need to cite this preprint.

Meurer A, Smith CP, Paprocki M, Čertík O, Kirpichev SB, Rocklin M, Kumar A, Ivanov S, Moore JK, Singh S, Rathnayake T, Vig S, Granger BE, Muller RP, Bonazzi F, Gupta H, Vats S, Johansson F, Pedregosa F, Curry MJ, Terrel AR, Roučka Š, Saboo A, Fernando I, Kulal S, Cimrman R, Scopatz A. 2017. SymPy: symbolic computing in Python. PeerJ Computer Science 3:e103 https://doi.org/10.7717/peerj-cs.103

SymPy: Symbolic Computing in Python

- ² Aaron Meurer¹, Christopher P. Smith², Mateusz Paprocki³, Ondřej Čertík⁴,
- ³ Sergey B. Kirpichev⁵, Matthew Rocklin⁶, AMiT Kumar⁷, Sergiu Ivanov⁸,
- Jason K. Moore⁹, Sartaj Singh¹⁰, Thilina Rathnayake¹¹, Sean Vig¹², Brian
- 5 E. Granger¹³, Richard P. Muller¹⁴, Francesco Bonazzi¹⁵, Harsh Gupta¹⁶,
- ⁶ Shivam Vats¹⁷, Fredrik Johansson¹⁸, Fabian Pedregosa¹⁹, Matthew J.
- 7 Curry²⁰, Andy R. Terrel²¹, Štěpán Roučka²², Ashutosh Saboo²³, Isuru
- ⁸ Fernando²⁴, Sumith Kulal²⁵, Robert Cimrman²⁶, and Anthony Scopatz²⁷
- ⁹ ¹University of South Carolina, Columbia, SC 29201 (asmeurer@gmail.com).
- ¹⁰ ²Polar Semiconductor, Inc., Bloomington, MN 55425 (smichr@gmail.com).
- ¹¹ ³Continuum Analytics, Inc., Austin, TX 78701 (mattpap@gmail.com).
- ¹² ⁴Los Alamos National Laboratory, Los Alamos, NM 87545 (certik@lanl.gov).
- ¹³ ⁵Moscow State University, Faculty of Physics, Leninskie Gory, Moscow, 119991, Russia
- 14 (skirpichev@gmail.com).
- ¹⁵ ⁶Continuum Analytics, Inc., Austin, TX 78701 (mrocklin@gmail.com).
- ¹⁶ ⁷Delhi Technological University, Shahbad Daulatpur, Bawana Road, New Delhi 110042,
- 17 India (dtu.amit@gmail.com).
- ¹⁸ ⁸Université Paris Est Créteil, 61 av. Général de Gaulle, 94010 Créteil, France
- ¹⁹ (sergiu.ivanov@u-pec.fr).
- ²⁰ ⁹University of California, Davis, Davis, CA 95616 (jkm@ucdavis.edu).
- ²¹ ¹⁰Indian Institute of Technology (BHU), Varanasi, Uttar Pradesh 221005, India
- ²² (singhsartaj94@gmail.com).
- ²³ ¹¹University of Moratuwa, Bandaranayake Mawatha, Katubedda, Moratuwa 10400, Sri
- Lanka (thilinarmtb.10@cse.mrt.ac.lk).
- ²⁵ ¹²University of Illinois at Urbana-Champaign, Urbana, IL 61801 (sean.v.775@gmail.com).
- ²⁶ ¹³California Polytechnic State University, San Luis Obispo, CA 93407
- 27 (ellisonbg@gmail.com).
- ²⁸ ¹⁴Center for Computing Research, Sandia National Laboratories, Albuquerque, NM
- 29 87185 (rmuller@sandia.gov).
- ³⁰ ¹⁵Max Planck Institute of Colloids and Interfaces, Department of Theory and
- ³¹ Bio-Systems, Science Park Golm, 14424 Potsdam, Germany
- 32 (francesco.bonazzi@mpikg.mpg.de).
- ¹⁶Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India
- ³⁴ (hargup@protonmail.com).
- ¹⁷Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India
- 36 (shivamvats.iitkgp@gmail.com).
- ³⁷ ¹⁸INRIA Bordeaux-Sud-Ouest LFANT project-team, 200 Avenue de la Vieille Tour,
- ³⁸ 33405 Talence, France (fredrik.johansson@gmail.com).
- ³⁹ ¹⁹INRIA SIERRA project-team, 2 Rue Simone IFF, 75012 Paris, France (f@bianp.net).
- ⁴⁰ ²⁰Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM
- ⁴¹ 87131 (mattjcurry@gmail.com).
- ⁴² ²¹Fashion Metric, Inc, Austin, TX 78681 (andy.terrel@gmail.com).
- ⁴³ ²²Faculty of Mathematics and Physics, Charles University in Prague, V Holešovičkách 2,
- ⁴⁴ 180 00 Praha, Czech Republic (stepan.roucka@mff.cuni.cz).
- ⁴⁵ ²³Birla Institute of Technology and Science, Pilani, K.K. Birla Goa Campus, NH 17B
- ⁴⁶ Bypass Road, Zuarinagar, Sancoale, Goa 403726, India
- 47 (ashutosh.saboo96@gmail.com).
- ⁴⁸ ²⁴University of Moratuwa, Bandaranayake Mawatha, Katubedda, Moratuwa 10400, Sri
- ⁴⁹ Lanka (isuru.11@cse.mrt.ac.lk).
- ⁵⁰ ²⁵Indian Institute of Technology Bombay, Powai, Mumbai, Maharashtra 400076, India
- ⁵¹ (sumith@cse.iitb.ac.in).
- ⁵² ²⁶New Technologies Research Centre, University of West Bohemia, Univerzitní 8, 306

⁵³ 14 Plzeň, Czech Republic (cimrman3@ntc.zcu.cz).

⁵⁴ ²⁷University of South Carolina, Columbia, SC 29201 (scopatz@cec.sc.edu).

55 ABSTRACT

- 55 SymPy is an open source computer algebra system written in pure Python. It is built with a focus on
- 57 extensibility and ease of use, through both interactive and programmatic applications. These characteristics
- have led SymPy to become a popular symbolic library for the scientific Python ecosystem. This paper
- ⁵⁹ presents the architecture of SymPy, a description of its features, and a discussion of select domain specific
- ⁶⁰ submodules. The supplementary materials provide additional examples and further outline details of the
- ⁶¹ architecture and features of SymPy.
- ⁶² Keywords: symbolic, Python, computer algebra system

63 1 INTRODUCTION

SymPy is a full featured computer algebra system (CAS) written in the Python programming
language [25]. It is free and open source software, licensed under the 3-clause BSD license [37].
The SymPy project was started by Ondřej Čertík in 2005, and it has since grown to over 500
contributors. Currently, SymPy is developed on GitHub using a bazaar community model [33].
The accessibility of the codebase and the open community model allow SymPy to rapidly respond
to the needs of users and developers.

Python is a dynamically typed programming language that has a focus on ease of use and readability. Due in part to this focus, it has become a popular language for scientific computing and data science, with a broad ecosystem of libraries [28]. SymPy is itself used by many libraries and tools to support research within a variety of domains, such as Sage [40] (pure mathematics), yt [45] (astronomy and astrophysics), PyDy [15] (multibody dynamics), and SfePy [10] (finite elements).

⁷⁶ Unlike many CASs, SymPy does not invent its own programming language. Python itself ⁷⁷ is used both for the internal implementation and end user interaction. By using the operator ⁷⁸ overloading functionality of Python, SymPy follows the embedded domain specific language ⁷⁹ paradigm proposed by Hudak [20]. The exclusive usage of a single programming language makes ⁸⁰ it easier for people already familiar with that language to use or develop SymPy. Simultaneously, ⁸¹ it enables developers to focus on mathematics, rather than language design.

SymPy is designed with a strong focus on usability as a library. Extensibility is important in its application program interface (API) design. Thus, SymPy makes no attempt to extend the Python language itself. The goal is for users of SymPy to be able to include SymPy alongside other Python libraries in their workflow, whether that be in an interactive environment or as a programmatic part in a larger system.

As a library, SymPy does not have a built-in graphical user interface (GUI). However, SymPy exposes a rich interactive display system, and supports registering printers with Jupyter [30] frontends, including the Notebook and Qt Console, which will render SymPy expressions using MathJax [9] or IAT_EX.

The remainder of this paper discusses key components of the SymPy software. Section 2 discusses the architecture of SymPy. Section 3 enumerates the features of SymPy and takes a closer look at some of the important ones. The section 4 looks at the numerical features of SymPy and its dependency library, mpmath. Section 5 looks at the domain specific physics submodules for performing symbolic and numerical calculations in classical mechanics and quantum mechanics. Conclusions and future directions for SymPy are given in section 6.

97 2 ARCHITECTURE

⁹⁸ Software architecture is of central importance in any large software project because it establishes

predictable patterns of usage and development [39]. This section describes the essential structural

components of SymPy, provides justifications for the design decisions that have been made, and
 gives example user-facing code as appropriate.

102 2.1 Basic Usage

¹⁰³ The following statement imports all SymPy functions into the global Python namespace. From

¹⁰⁴ here on, all examples in this paper assume that this statement has been executed.¹

```
105 >>> from sympy import *
```

Symbolic variables, called symbols, must be defined and assigned to Python variables before they can be used. This is typically done through the symbols function, which may create multiple symbols in a single function call. For instance,

```
109 >>> x, y, z = symbols('x y z')
```

creates three symbols representing variables named x, y, and z. In this particular instance, these symbols are all assigned to Python variables of the same name. However, the user is free to assign them to different Python variables, while representing the same symbol, such as a, b, $c = \text{symbols}('x \ y \ z')$. In order to minimize potential confusion, though, all examples in this paper will assume that the symbols x, y, and z have been assigned to Python variables identical to their symbolic names.

Expressions are created from symbols using Python's mathematical syntax. For instance, the following Python code creates the expression $(x^2 - 2x + 3)/y$.

¹²⁰ Importantly, SymPy expressions are immutable. This simplifies the design of SymPy by ¹²¹ allowing expression interning. It also enables expressions to be hashed and stored in Python ¹²² dictionaries, thereby permitting features such as caching.

123 2.2 The Core

A computer algebra system (CAS) represents mathematical expressions as data structures. For example, the mathematical expression x + y is represented as a tree with three nodes, +, x, and y, where x and y are ordered children of +. As users manipulate mathematical expressions with traditional mathematical syntax, the CAS manipulates the underlying data structures. Automated optimizations and computations such as integration, simplification, etc. are all functions that consume and produce expression trees.

In SymPy every symbolic expression is an instance of a Python Basic class, a superclass
 of all SymPy types providing common methods to all SymPy tree-elements, such as traversals.
 The children of a node in the tree are held in the args attribute. A terminal or leaf node in the
 expression tree has empty args.

¹³⁴ For example, consider the expression xy + 2:

```
_{135} >>> expr = x*y + 2
```

¹³⁶ By order of operations, the parent of the expression tree for expr is an addition, so it is of type ¹³⁷ Add. The child nodes of expr are 2 and x*y.

- 138 >>> type(expr)
- 139 <class 'sympy.core.add.Add'>
- 140 >>> expr.args
- ¹⁴¹ (2, x*y)

Descending further down into the expression tree yields the full expression. For example, the next child node (given by expr.args[0]) is 2. Its class is Integer, and it has an empty args tuple, indicating that it is a leaf node.

 $^{^{1}}$ All examples in this paper use SymPy version 1.0.

```
145 >>> expr.args[0]
146 2
147 >>> type(expr.args[0])
148 <class 'sympy.core.numbers.Integer'>
149 >>> expr.args[0].args
150 ()
```

A useful way to view an expression tree is using the srepr function, which returns a string representation of an expression as valid Python code with all the nested class constructor calls to create the given expression.

```
154 >>> srepr(expr)
```

```
"Add(Mul(Symbol('x'), Symbol('y')), Integer(2))"
```

¹⁵⁶ Every SymPy expression satisfies a key identity invariant:

```
157 expr.func(*expr.args) == expr
```

This means that expressions are rebuildable from their args.² Note that in SymPy the == operator represents exact structural equality, not mathematical equality. This allows testing if any two expressions are equal to one another as expression trees. For example, even though $(x+1)^2$ and x^2+2x+1 are equal mathematically, SymPy gives

```
_{162} >>> (x + 1)**2 == x**2 + 2*x + 1
```

163 False

because they are different as expression trees (the former is a Pow object and the latter is an Add
 object).

Python allows classes to override mathematical operators. The Python interpreter translates the above x*y + 2 to, roughly, (x.__mul__(y)).__add__(2). Both x and y, returned from the symbols function, are Symbol instances. The 2 in the expression is processed by Python as a literal, and is stored as Python's built in int type. When 2 is passed to the __add__ method of Symbol, it is converted to the SymPy type Integer(2) before being stored in the resulting expression tree. In this way, SymPy expressions can be built in the natural way using Python operators and numeric literals.

173 2.3 Assumptions

SymPy performs logical inference through its assumptions system. The assumptions system allows users to specify that symbols have certain common mathematical properties, such as being positive, imaginary, or integral. SymPy is careful to never perform simplifications on an expression unless the assumptions allow them. For instance, the identity $\sqrt{t^2} = t$ holds if t is nonnegative ($t \ge 0$). However, for general complex t, no such identity holds.

By default, SymPy performs all calculations assuming that symbols are complex valued. This assumption makes it easier to treat mathematical problems in full generality.

```
181 >>> t = Symbol('t')
```

```
182 >>> sqrt(t**2)
```

```
183 sqrt(t**2)
```

¹⁸⁴ By assuming the most general case, that symbols are complex by default, SymPy avoids ¹⁸⁵ performing mathematically invalid operations. However, in many cases users will wish to simplify ¹⁸⁶ expressions containing terms like $\sqrt{t^2}$.

Assumptions are set on Symbol objects when they are created. For instance Symbol('t', positive=True) will create a symbol named t that is assumed to be positive.

```
189 >>> t = Symbol('t', positive=True)
190 >>> sqrt(t**2)
191 t
```

 $^{^{2}}$ expr.func is used instead of type(expr) to allow the function of an expression to be distinct from its actual Python class. In most cases the two are the same.

Some of the common assumptions that SymPy allows are positive, negative, real, nonpositive, integer, prime and commutative.³ Assumptions on any object can be checked with the is_assumption attributes, like t.is_positive.

Assumptions are only needed to restrict a domain so that certain simplifications can be performed. They are not required to make the domain match the input of a function. For instance, one can create the object $\sum_{n=0}^{m} f(n)$ as Sum(f(n), (n, 0, m)) without setting integer=True when creating the Symbol object n.

The assumptions system additionally has deductive capabilities. The assumptions use a three-valued logic using the Python built in objects True, False, and None. Note that False is returned if the SymPy object doesn't or can't have the assumption. For example, both I.is_real and I.is_prime return False for the imaginary unit I.

None represents the "unknown" case. This could mean that given assumptions do not unambiguously specify the truth of an attribute. For instance, Symbol('x', real=True).is_positive will give None because a real symbol might be positive or negative. The None could also mean that not enough is known or implemented to compute the given fact. For instance, (pi + E).is_irrational gives None, because determining whether $\pi + e$ is rational or irrational is an open problem in mathematics [24].

Basic implications between the facts are used to deduce assumptions. For instance, the assumptions system knows that being an integer implies being rational, so Symbol('x', integer=True).is_rational returns True. Furthermore, expressions compute the assumptions on themselves based on the assumptions of their arguments. For instance, if x and y are both created with positive=True, then (x + y).is_positive will be True whereas (x - y).is_positive will be None.

214 2.4 Extensibility

While the core of SymPy is relatively small, it has been extended to a wide variety of domains by a broad range of contributors. This is due in part because the same language, Python, is used both for the internal implementation and the external usage by users. All of the extensibility capabilities available to users are also utilized by SymPy itself. This eases the transition pathway from SymPy user to SymPy developer.

The typical way to create a custom SymPy object is to subclass an existing SymPy class, usually Basic, Expr, or Function. All SymPy classes used for expression trees⁴ should be subclasses of the base class Basic, which defines some basic methods for symbolic expression trees. Expr is the subclass for mathematical expressions that can be added and multiplied together. Instances of Expr typically represent complex numbers, but may also include other "rings" like matrix expressions. Not all SymPy classes are subclasses of Expr. For instance, logic expressions such as And(x, y) are subclasses of Basic but not of Expr.

The Function class is a subclass of Expr which makes it easier to define mathematical functions called with arguments. This includes named functions like $\sin(x)$ and $\log(x)$ as well as undefined functions like f(x). Subclasses of Function should define a class method eval, which returns a canonical form of the function application (usually an instance of some other class, i.e. a Number) or None, if for given arguments that function should not be automatically evaluated.

Many SymPy functions perform various evaluations down the expression tree. Classes define their behavior in such functions by defining a relevant _eval_* method. For instance, an object can indicate to the diff function how to take the derivative of itself by defining the _eval_derivative(self, x) method, which may in turn call diff on its args. (Subclasses of Function should implement fdiff method instead, it returns the derivative of the function without considering the chain rule.) The most common _eval_* methods relate to the assumptions: _eval_is_assumption is used to deduce assumption on the object.

As an example of the notions presented in this section, Listing 1 presents a minimal version of the gamma function $\Gamma(x)$ from SymPy, which evaluates itself on positive integer arguments, has the positive and real assumptions defined, can be rewritten in terms of factorial with gamma(x).rewrite(factorial), and can be differentiated. self.func is used throughout instead of referencing gamma explicitly so that potential subclasses of gamma can reuse the methods.

³If A and B are Symbols created with commutative=False then SymPy will keep $A \cdot B$ and $B \cdot A$ distinct.

 $^{^4\}mathrm{Some}$ internal classes, such as those used in the polynomial module, do not follow this rule for efficiency reasons.

```
Listing 1. A minimal implementation of sympy.gamma.
    from sympy import Integer, Function, floor, factorial, polygamma
244
245
    class gamma(Function)
246
        @classmethod
247
        def eval(cls, arg):
248
             if isinstance(arg, Integer) and arg.is positive:
249
                 return factorial(arg - 1)
250
251
        def eval is positive(self):
252
             x = self.args[0]
253
             if x.is positive:
254
                 return True
255
             elif x.is noninteger:
256
                 return floor(x).is_even
257
258
        def eval is real(self):
259
             x = self.args[0]
260
            # noninteger means real and not integer
261
             if x.is positive or x.is noninteger:
262
                 return True
263
264
        def eval rewrite as factorial(self, z):
265
             return factorial(z - 1)
266
267
        def fdiff(self, argindex=1):
268
             from sympy.core.function import ArgumentIndexError
269
             if argindex == 1:
270
                 return self.func(self.args[0])*polygamma(0, self.args[0])
271
             else:
272
                 raise ArgumentIndexError(self, argindex)
273
```

The gamma function implemented in SymPy has many more capabilities than the above listing, such as evaluation at rational points and series expansion.

276 **3 FEATURES**

Although SymPy's extensive feature set cannot be covered in-depth in this paper, calculus and
other bedrock areas are discussed in their own subsections. Additionally, Table 1 gives a compact
listing of all major capabilities present in the SymPy codebase. This grants a sampling from the
breadth of topics and application domains that SymPy services. Unless stated otherwise, all
features noted in Table 1 are symbolic in nature. Numeric features are discussed in Section 4.

 Table 1. SymPy Features and Descriptions

Feature	Description
Calculus	Algorithms for computing derivatives, integrals, and limits.
Category Theory	Representation of objects, morphisms, and diagrams. Tools for drawing
	diagrams with Xy-pic.
Code Generation	Generation of compilable and executable code in a variety of different
	programming languages from expressions directly. Target languages
	include C, Fortran, Julia, JavaScript, Mathematica, MATLAB and
	Octave, Python, and Theano.
Combinatorics &	Permutations, combinations, partitions, subsets, various permutation
Group Theory	groups (such as polyhedral, Rubik, symmetric, and others), Gray
	codes [27], and Prufer sequences [4].

Concrete Math	Summation, products, tools for determining whether summation and
	product expressions are convergent, absolutely convergent, hypergeo- metric, and for determining other properties; computation of Gosper's
Cryptography	normal form [32] for two univariate polynomials. Block and stream ciphers, including shift, Affine, substitution, Vi- genère's, Hill's, bifid, RSA, Kid RSA, linear-feedback shift registers,
	and Elgamal encryption.
Differential Geome-	Representations of manifolds, metrics, tensor products, and coordinate
try Geometry	systems in Riemannian and pseudo-Riemannian geometries [41]. Representations of 2D geometrical entities, such as lines and circles. Enables queries on these entities, such as asking the area of an ellipse, checking for collinearity of a set of points, or finding the intersection
T· 41 1	between objects.
Lie Algebras	Representations of Lie algebras and root systems.
Logic	Boolean expressions, equivalence testing, satisfiability, and normal forms.
Matrices	Tools for creating matrices of symbols and expressions. Both sparse and
	dense representations, as well as symbolic linear algebraic operations
	(e.g., inversion and factorization), are supported.
Matrix Expressions	Matrices with symbolic dimensions (unspecified entries). Block matri-
	ces.
Number Theory	Prime number generation, primality testing, integer factorization, con-
	tinued fractions, Egyptian fractions, modular arithmetic, quadratic
	residues, partitions, binomial and multinomial coefficients, prime num-
Diotting	ber tools, hexidecimal digits of π , and integer factorization.
Plotting	Hooks for visualizing expressions via matplotlib [21] or as text drawings when lacking a graphical back-end. 2D function plotting, 3D function
	plotting, and 2D implicit function plotting are supported.
Polynomials	Polynomial algebras over various coefficient domains. Functionality
1 ory nonnais	ranges from simple operations (e.g., polynomial division) to advanced
	computations (e.g., Gröbner bases [1] and multivariate factorization
	over algebraic number domains).
Printing	Functions for printing SymPy expressions in the terminal with ASCII
0	or Unicode characters and converting SymPy expressions to LATEX and
	MathML.
Quantum Mechan-	Quantum states, bra-ket notation, operators, basis sets, representa-
ics	tions, tensor products, inner products, outer products, commutators,
	anticommutators, and specific quantum system implementations.
Series	Series expansion, sequences, and limits of sequences. This includes
	Taylor, Laurent, and Puiseux series as well as special series, such as
_	Fourier and formal power series.
Sets	Representations of empty, finite, and infinite sets. This includes special
	sets such as for all natural, integer, and complex numbers. Operations
	on sets such as union, intersection, Cartesian product, and building
0:	sets from other sets are supported.
Simplification	Functions for manipulating and simplifying expressions. Includes
	algorithms for simplifying hypergeometric functions, trigonometric expressions, rational functions, combinatorial functions, square root
	denesting, and common subexpression elimination.
Solvers	Functions for symbolically solving equations, systems of equations,
	both linear and non-linear, inequalities, ordinary differential equations,
	partial differential equations, Diophantine equations, and recurrence
	relations.

7/19

Special Functions	Implementations of a number of well known special functions, including Dirac delta, Gamma, Beta, Gauss error functions, Fresnel integrals, Exponential integrals, Logarithmic integrals, Trigonometric integrals, Bessel, Hankel, Airy, B-spline, Riemann Zeta, Dirichlet eta, polyloga-
	rithm, Lerch transcendent, hypergeometric, elliptic integrals, Mathieu,
	Jacobi polynomials, Gegenbauer polynomial, Chebyshev polynomial, Legendre polynomial, Hermite polynomial, Laguerre polynomial, and
	spherical harmonic functions.
Statistics	Support for a random variable type as well as the ability to declare
	this variable from prebuilt distribution functions such as Normal,
	Exponential, Coin, Die, and other custom distributions [36].
Tensors	Symbolic manipulation of indexed objects.
Vectors	Basic operations on vectors and differential calculus with respect to
	3D Cartesian coordinate systems.

282 3.1 Simplification

The generic way to simplify an expression is by calling the simplify function. It must be emphasized that simplification is not a rigorously defined mathematical operation [8]. The simplify function applies several simplification routines along with heuristics to make the output expression "simple".⁵

It is often preferable to apply more directed simplification functions. These apply very specific rules to the input expression and are typically able to make guarantees about the output. For instance, the factor function, given a polynomial with rational coefficients in several variables, is guaranteed to produce a factorization into irreducible factors. Table 2 lists common simplification

²⁹¹ functions.

 Table 2. Some SymPy Simplification Functions

expand	expand the expression
factor	factor a polynomial into irreducibles
collect	collect polynomial coefficients
cancel	rewrite a rational function as p/q with common factors canceled
apart	compute the partial fraction decomposition of a rational function
trigsimp	simplify trigonometric expressions [14]
hyperexpand	expand hypergeometric functions [34, 35]

292 3.2 Calculus

SymPy provides all the basic operations of calculus, such as calculating limits, derivatives,
 integrals, or summations.

Limits are computed with the limit function, using the Gruntz algorithm [18] for computing symbolic limits and heuristics (a description of the Gruntz algorithm may be found in the supplement). For example, the following computes $\lim_{x\to\infty} x\sin(\frac{1}{x}) = 1$. Note that SymPy denotes ∞ as oo.

```
299 >>> limit(x*sin(1/x), x, oo)
300 1
```

As a more complex example, SymPy computes

$$\lim_{x \to 0} \left(2e^{\frac{1 - \cos(x)}{\sin(x)}} - 1 \right)^{\frac{\sinh(x)}{\tan^2(x)}} = e.$$

⁵The measure parameter of the simplify function lets specify the Python function used to determine how complex an expression is. The default measure function returns the total number of operations in the expression.

301 >>> limit((2*E**((1-cos(x))/sin(x))-1)**(sinh(x)/atan(x)**2), x, 0)
302 E

Derivatives are computed with the diff function, which recursively uses the various differentiation rules.

```
305 >>> diff(sin(x)*exp(x), x)
```

```
_{306} \exp(x) * \sin(x) + \exp(x) * \cos(x)
```

Integrals are calculated with the integrate function. SymPy implements a combination of the Risch algorithm [6], table lookups, a reimplementation of Manuel Bronstein's "Poor Man's Integrator" [5], and an algorithm for computing integrals based on Meijer G-functions [34, 35]. These allow SymPy to compute a wide variety of indefinite and definite integrals. The Meijer G-function algorithm and the Risch algorithm are respectively demonstrated below by the computation of

$$\int_{0}^{\infty} e^{-st} \log\left(t\right) dt = -\frac{\log\left(s\right) + \gamma}{s}$$

and

$$\int \frac{-2x^2 \left(\log \left(x\right)+1\right) e^{x^2} + \left(e^{x^2}+1\right)^2}{x \left(e^{x^2}+1\right)^2 \left(\log \left(x\right)+1\right)} \, dx = \log \left(\log \left(x\right)+1\right) + \frac{1}{e^{x^2}+1}.$$

```
307 >>> s, t = symbols('s t', positive=True)
308 >>> integrate(exp(-s*t)*log(t), (t, 0, oo)).simplify()
309 -(log(s) + EulerGamma)/s
310 >>> integrate((-2*x**2*(log(x) + 1)*exp(x**2) +
311 ... (exp(x**2) + 1)**2)/(x*(exp(x**2) + 1)**2*(log(x) + 1)), x)
312 log(log(x) + 1) + 1/(exp(x**2) + 1)
```

Summations are computed with summation using a combination of Gosper's algorithm [17], an algorithm that uses Meijer G-functions [34, 35], and heuristics. Products are computed with product function via a suite of heuristics.

```
316 >>> i, n = symbols('i n')
317 >>> summation(2**i, (i, 0, n - 1))
318 2**n - 1
319 >>> summation(i*factorial(i), (i, 1, n))
320 n*factorial(n) + factorial(n) - 1
```

Integrals, derivatives, summations, products, and limits that cannot be computed return unevaluated objects. These can also be created directly if the user chooses.

```
323 >>> integrate(x**x, x)
324 Integral(x**x, x)
325 >>> Sum(2**i, (i, 0, n - 1))
326 Sum(2**i, (i, 0, n - 1))
```

327 3.3 Polynomials

SymPy implements a suite of algorithms for polynomial manipulation, which ranges from
 relatively simple algorithms for doing arithmetic of polynomials, to advanced methods for
 factoring multivariate polynomials into irreducibles, symbolically determining real and complex
 root isolation intervals, or computing Gröbner bases.

Polynomial manipulation is useful in its own right. Within SymPy, though, it is mostly used indirectly as a tool in other areas of the library. In fact, many mathematical problems in symbolic computing are first expressed using entities from the symbolic core, preprocessed, and then transformed into a problem in the polynomial algebra, where generic and efficient algorithms are used to solve the problem. The solutions to the original problem are subsequently
 recovered from the results. This is a common scheme in symbolic integration or summation
 algorithms.

SymPy implements dense and sparse polynomial representations.⁶ Both are used in the uni-339 variate and multivariate cases. The dense representation is the default for univariate polynomials. 340 For multivariate polynomials, the choice of representation is based on the application. The most 341 common case for the sparse representation is algorithms for computing Gröbner bases (Buchberger, 342 F4, and F5) [7, 11, 12]. This is because different monomial orderings can be expressed easily in 343 this representation. However, algorithms for computing multivariate GCDs or factorizations, at 344 least those currently implemented in SymPy [29], are better expressed when the representation 345 is dense. The dense multivariate representation is specifically a recursively-dense representation, 346 where polynomials in $K[x_0, x_1, \ldots, x_n]$ are viewed as a polynomials in $K[x_0][x_1] \ldots [x_n]$. Note 347 that despite this, the coefficient domain K, can be a multivariate polynomial domain as well. 348 The dense recursive representation in Python gets inefficient as the number of variables increases. 349 Some examples for the sympy.polys module can be found in the supplement. 350

351 3.4 Printers

SymPy has a rich collection of expression printers. By default, an interactive Python session will render the str form of an expression, which has been used in all the examples in this paper so far. The str form of an expression is valid Python and roughly matches what a user would type to enter the expression.

```
356 >>> phi0 = Symbol('phi0')
```

357 >>> str(Integral(sqrt(phi0), phi0))

```
358 'Integral(sqrt(phi0), phi0)'
```

Expressions can be printed in 2D with monospace fonts via pprint. Unicode characters are used for rendering mathematical symbols such as integral signs, square roots, and parentheses. Greek letters and subscripts in symbol names that have Unicode code points associated are also rendered automatically.

>>> pprint(Integral(sqrt(phi0 + 1), phi0))

$$\sqrt{\phi_0 + 1} d(\phi_0)$$

363

Alternately, the use_unicode=False flag can be set, which causes the expression to be printed using only ASCII characters.

```
366 >>> pprint(Integral(sqrt(phi0 + 1), phi0), use_unicode=False)
367 /
368 |
369 |
370 | \/ phi0 + 1 d(phi0)
371 |
372 /
```

 $_{373}$ The function latex returns a LAT_EX representation of an expression.

```
374 >>> print(latex(Integral(sqrt(phi0 + 1), phi0)))
```

```
375 \int \sqrt{\phi_{0} + 1}\, d\phi_{0}
```

Users are encouraged to run the init_printing function at the beginning of interactive sessions, which automatically enables the best pretty printing supported by their environment. In the Jupyter Notebook or Qt Console [30], the LATFX printer is used to render expressions

⁶In a dense representation, the coefficients for all terms up to the degree of each variable are stored in memory. In a sparse representation, only the nonzero coefficients are stored.

³⁷⁹ using MathJax or LATEX, if it is installed on the system. The 2D text representation is used ³⁸⁰ otherwise.

Other printers such as MathML are also available. SymPy uses an extensible printer subsystem for customizing any given printer, and allows custom objects to define their printing behavior for any printer. The code generation functionality of SymPy relies on this subsystem to convert expressions into code in various target programming languages.

385 3.5 Solvers

SymPy has a module of equation solvers that can handle ordinary differential equations, recurrence
 relationships, Diophantine equations, and algebraic equations. There is also rudimentary support
 for simple partial differential equations.

There are two functions for solving algebraic equations in SymPy: solve and solveset. solveset has several design changes with respect to the older solve function. This distinction is present in order to resolve the usability issues with the previous solve function API while maintaining backward compatibility with earlier versions of SymPy. solveset only requires essential input information from the user. The function signatures of solve and solveset are

```
394 solve(f, *symbols, **flags)
```

395 solveset(f, symbol, domain=S.Complexes)

The domain parameter is typically either S.Complexes (the default) or S.Reals; the latter causes solveset to only return real solutions.

An important difference between the two functions is that the output API of solve varies with input (sometimes returning a Python list and sometimes a Python dictionary) whereas solveset always returns a SymPy set object.

Both functions implicitly assume that expressions are equal to 0. For instance, solveset(x - 1 - y) as y = 1 - 0 for x = 1 - 0

402 1, x) solves x - 1 = 0 for x.

solveset is under active development as a planned replacement for solve. There are certain
 features which are implemented in solve that are not yet implemented in solveset, including
 multivariate systems, and some transcendental equations.

406 More examples of solveset and solve can be found in the supplement.

407 3.6 Matrices

Besides being an important feature in its own right, computations on matrices with symbolic
entries are important for many algorithms within SymPy. The following code shows some basic
usage of the Matrix class.

411 >>> A = Matrix(2, 2, [x, x + y, y, x])

412 >>> A

413 Matrix([

 $_{414}$ [x, x + y],

415 [y, x]])

SymPy matrices support common symbolic linear algebra manipulations, including matrix
addition, multiplication, exponentiation, computing determinants, solving linear systems, and
computing inverses using LU decomposition, LDL decomposition, Gauss-Jordan elimination,
Cholesky decomposition, Moore-Penrose pseudoinverse, and adjugate matrix.

All operations are performed symbolically. For instance, eigenvalues are computed by generating the characteristic polynomial using the Berkowitz algorithm and then solving it using polynomial routines.

```
423 >>> A.eigenvals()
```

424 {x - $sqrt(y^{*}(x + y)): 1, x + sqrt(y^{*}(x + y)): 1$ }

⁴²⁵ Internally these matrices store the elements as lists of lists, making it a dense representation.⁷

⁴²⁶ For storing sparse matrices, the SparseMatrix class can be used. Sparse matrices store their ⁴²⁷ elements as a dictionary of keys.

⁷Similar to the polynomials module, dense here means that all entries are stored in memory, contrasted with a sparse representation where only nonzero entries are stored.

⁴²⁸ SymPy also supports matrices with symbolic dimension values. MatrixSymbol represents ⁴²⁹ a matrix with dimensions $m \times n$, where m and n can be symbolic. Matrix addition and ⁴³⁰ multiplication, scalar operations, matrix inverse, and transpose are stored symbolically as matrix ⁴³¹ expressions.

Block matrices are also implemented in SymPy. BlockMatrix elements can be any matrix expression, including explicit matrices, matrix symbols, and other block matrices. All functionalities
of matrix expressions are also present in BlockMatrix.

When symbolic matrices are combined with the assumptions module for logical inference, they provide powerful reasoning over invertibility, semi-definiteness, orthogonality, etc., which are valuable in the construction of numerical linear algebra systems.

438 More examples for Matrix and BlockMatrix may be found in the supplement.

439 4 NUMERICS

Floating point numbers in SymPy are implemented by the Float class, which represents an
arbitrary-precision binary floating-point number by storing its value and precision (in bits).
This representation is distinct from the Python built-in float type, which is a wrapper around
machine double types and uses a fixed precision (53-bit).

Because Python float literals are limited in precision, strings should be used to input precise decimal values:

```
446 >>> Float(1.1)
```

```
447 1.10000000000000
```

```
448 >>> Float(1.1, 30) # precision equivalent to 30 digits
```

```
449 1.1000000000000008881784197001
```

```
450 >>> Float("1.1", 30)
```

The evalf method converts a constant symbolic expression to a Float with the specified precision, here 25 digits:

```
454 >>> (pi + 1).evalf(25)
455 4.141592653589793238462643
```

Float numbers do not track their accuracy, and should be used with caution within symbolic
expressions since familiar dangers of floating-point arithmetic apply [16]. A notorious case is
that of catastrophic cancellation:

```
459 >>> cos(exp(-100)).evalf(25) - 1
460 0
```

⁴⁶¹ Applying the evalf method to the whole expression solves this problem. Internally, evalf
⁴⁶² estimates the number of accurate bits of the floating-point approximation for each sub-expression,
⁴⁶³ and adaptively increases the working precision until the estimated accuracy of the final result
⁴⁶⁴ matches the sought number of decimal digits:

```
465 >>> (cos(exp(-100)) - 1).evalf(25)
466 -6.919482633683687653243407e-88
```

⁴⁶⁷ The evalf method works with complex numbers and supports more complicated expressions, such ⁴⁶⁸ as special functions, infinite series, and integrals. The internal error tracking does not provide ⁴⁶⁹ rigorous error bounds (in the sense of interval arithmetic) and cannot be used to accurately track ⁴⁷⁰ uncertainty in measurement data; the sole purpose is to mitigate loss of accuracy that typically ⁴⁷¹ occurs when converting symbolic expressions to numerical values.

472 4.1 The mpmath library

 $_{\tt 473}$ $\,$ The implementation of arbitrary-precision floating-point arithmetic is supplied by the mpmath

 $_{474}$ $\,$ library. Originally, it was developed as a SymPy module but has subsequently been moved to

 $_{475}\,$ a standalone pure-Python package. The basic datatypes in mpmath are mpf and mpc, which

respectively act as multiprecision substitutes for Python's float and complex. The floating-point
 precision is controlled by a global context:⁸

```
478 >>> import mpmath
```

479 >>> mpmath.mp.dps = 30 # 30 digits of precision

```
480 >>> mpmath.mpf("0.1") + mpmath.exp(-50)
```

```
482 >>> print(_) # pretty-printed
```

483 0.100000000000000000000192874985

For pure numerical computing, it is convenient to use mpmath directly with from mpmath import *. Nevertheless, it is best to avoid such an import statement when using SymPy simultaneously, since the names of numerical functions such as exp will collide the symbolic counterparts in SymPy.

Like SymPy, mpmath is a pure Python library. Internally, mpmath represents a floating-point number $(-1)^{s}x \cdot 2^{y}$ by a tuple (s, x, y, b) where x and y are arbitrary-size Python integers and the redundant integer b stores the bit length of x for quick access. If GMPY [19] is installed, mpmath automatically uses the gmpy.mpz type for x, and GMPY methods for rounding-related operations, improving performance.

The mpmath library supports special functions, root-finding, linear algebra, polynomial approximation, and numerical computation of limits, derivatives, integrals, infinite series, and ODE solutions. All features work in arbitrary precision and use algorithms that allow computing hundreds of digits rapidly (except in degenerate cases).

The double exponential (tanh-sinh) quadrature is used for numerical integration by default. 497 For smooth integrands, this algorithm usually converges extremely rapidly, even when the 498 integration interval is infinite or singularities are present at the endpoints [43, 2]. However, for 499 good performance, singularities in the middle of the interval must be specified by the user. To 500 evaluate slowly converging limits and infinite series, mpmath automatically tries Richardson 501 extrapolation and the Shanks transformation (Euler-Maclaurin summation can also be used) [3]. 502 A function to evaluate oscillatory integrals by means of convergence acceleration is also available. 503 A wide array of higher mathematical functions are implemented with full support for complex 504 values of all parameters and arguments, including complete and incomplete gamma functions, 505 Bessel functions, orthogonal polynomials, elliptic functions and integrals, zeta and polylogarithm 506 functions, the generalized hypergeometric function, and the Meijer G-function. The Meijer 507 G-function instance $G_{1,3}^{3,0}(0;\frac{1}{2},-1,-\frac{3}{2}|x)$ is a good test case [44]; past versions of both Maple and 508 Mathematica produced incorrect numerical values for large x > 0. Here, mpmath automatically 509 removes an internal singularity and compensates for cancellations (amounting to 656 bits of 510 precision when x = 10000), giving correct values: 511

512 >>> mpmath.mp.dps = 15

```
s13 >>> mpmath.meijerg([[],[0]],[[-0.5,-1,-1.5],[]],10000)
```

⁵¹⁴ mpf('2.4392576907199564e-94')

⁵¹⁵ Equivalently, with SymPy's interface this function can be evaluated as:

si6 >>> meijerg([[],[0]],[[-S(1)/2,-1,-S(3)/2],[]],10000).evalf()

```
517 2.43925769071996e-94
```

⁵¹⁸ Symbolic integration and summation often produces hypergeometric and Meijer G-function ⁵¹⁹ closed forms (see Subsection 3.2); numerical evaluation of such special functions is a useful ⁵²⁰ complement to direct numerical integration and summation.

521 5 DOMAIN SPECIFIC SUBMODULES

522 SymPy includes several packages that allow users to solve domain specific problems. For example,

a comprehensive physics package is included that is useful for solving problems in mechanics,
 optics, and quantum mechanics along with support for manipulating physical quantities with
 units.

⁸All examples in this section use mpmath version 0.19.

526 5.1 Classical Mechanics

527 One of the core domains that SymPy suports is the physics of classical mechanics. This is in

⁵²⁸ turn separated into two distinct components: vector algebra symbolics and mechanics.

529 5.1.1 Vector Algebra

The sympy.physics.vector package provides reference frame-, time-, and space-aware vector 530 and dyadic objects that allow for three-dimensional operations such as addition, subtraction, 531 scalar multiplication, inner and outer products, and cross products. Both of these objects can 532 be written in very compact notation that make it easy to express the vectors and dyadics in 533 terms of multiple reference frames with arbitrarily defined relative orientations. The vectors 534 are used to specify the positions, velocities, and accelerations of points; orientations, angular 535 velocities, and angular accelerations of reference frames; and forces and torques. The dyadics are 536 essentially reference frame-aware 3×3 tensors [42]. The vector and dyadic objects can be used 537 for any one-, two-, or three-dimensional vector algebra, and they provide a strong framework for 538 building physics and engineering tools. 539

The following Python code demonstrates how a vector is created using the orthogonal unit vectors of three reference frames that are oriented with respect to each other, and the result of expressing the vector in the A frame. The B frame is oriented with respect to the A frame using Z-X-Z Euler Angles of magnitude π , $\frac{\pi}{2}$, and $\frac{\pi}{3}$ rad, respectively, whereas the C frame is oriented with respect to the B frame through a simple rotation about the B frame's X unit vector through $\frac{\pi}{2}$ rad.

```
>>> from sympy.physics.vector import ReferenceFrame
546
    >>> A = ReferenceFrame('A')
547
    >>> B = ReferenceFrame('B')
548
    >>> C = ReferenceFrame('C')
549
    >>> B.orient(A, 'body', (pi, pi/3, pi/4), 'zxz')
550
    >>> C.orient(B, 'axis', (pi/2, B.x))
551
    >>> v = 1*A.x + 2*B.z + 3*C.y
552
    >>> v
553
    A.x + 2*B.z + 3*C.y
554
    >>> v.express(A)
555
    A.x + 5*sqrt(3)/2*A.y + 5/2*A.z
556
```

557 5.1.2 Mechanics

The sympy.physics.mechanics package utilizes the sympy.physics.vector package to populate 558 time-aware particle and rigid-body objects to fully describe the kinematics and kinetics of a 559 rigid multi-body system. These objects store all of the information needed to derive the ordinary 560 differential or differential algebraic equations that govern the motion of the system, i.e., the 561 equations of motion. These equations of motion abide by Newton's laws of motion and can handle 562 arbitrary kinematic constraints or complex loads. The package offers two automated methods for 563 formulating the equations of motion based on Lagrangian Dynamics [23] and Kane's Method [22]. 564 Lastly, there are automated linearization routines for constrained dynamical systems [31]. 565

566 5.2 Quantum Mechanics

The sympy.physics.quantum package has extensive capabilities for performing symbolic quantum mechanics, using Python objects to represent the different mathematical objects relevant in quantum theory [38]: states (bras and kets), operators (unitary, Hermitian, etc.), and basis sets, as well as operations on these objects such as representations, tensor products, inner products, outer products, commutators, and anticommutators. The base objects are designed in the most general way possible to enable any particular quantum system to be implemented by subclassing the base operators and defining the relevant class methods to provide system-specific logic.

574 Symbolic quantum operators and states may be defined, and one can perform a full range of 575 operations with them.

576 >>> from sympy.physics.quantum import Commutator, Dagger, Operator

577 >>> from sympy.physics.quantum import Ket, qapply

```
>>> A = Operator('A')
578
    >>> B = Operator('B')
579
    >>> C = Operator('C')
580
    >>> D = Operator('D')
581
    >>> a = Ket('a')
582
    >>> comm = Commutator(A, B)
583
    >>> comm
584
    [A,B]
585
    >>> gapply(Dagger(comm*a)).doit()
586
587
    -<a|*(Dagger(A)*Dagger(B) - Dagger(B)*Dagger(A))</pre>
    Commutators can be expanded using common commutator identities:
588
    >>> Commutator(C+B, A*D).expand(commutator=True)
589
    -[A,B]*D - [A,C]*D + A*[B,D] + A*[C,D]
590
        On top of this set of base objects, a number of specific quantum systems have been implemented
591
    in a fully symbolic framework. These include:
592
       • Many of the exactly solvable quantum systems, including simple harmonic oscillator states
593
          and raising/lowering operators, infinite square well states, and 3D position and momentum
594
          operators and states.
595
       • Second quantized formalism of non-relativistic many-body quantum mechanics [13].
596
       • Quantum angular momentum [46]. Spin operators and their eigenstates can be represented
597
          in any basis and for any quantum numbers. A rotation operator representing the Wigner-D
598
          matrix, which may be defined symbolically or numerically, is also implemented to rotate
599
          spin eigenstates. Functionality for coupling and uncoupling of arbitrary spin eigenstates is
600
          provided, including symbolic representations of Clebsch-Gordon coefficients and Wigner
601
          symbols.
602
         Quantum information and computing [26]. Multidimensional qubit states, and a full
603
          set of one- and two-qubit gates are provided and can be represented symbolically or as
604
          matrices/vectors. With these building blocks, it is possible to implement a number of basic
605
          quantum algorithms including the quantum Fourier transform, quantum error correction,
606
          quantum teleportation, Grover's algorithm, dense coding, etc. In addition, any quantum
607
          circuit may be plotted using the circuit plot function (Figure 1).
608
       Here are a few short examples of the quantum information and computing capabilities in
609
    sympy.physics.quantum. Start with a simple four-qubit state and flip the second qubit from the
610
    right using a Pauli-X gate:
611
    >>> from sympy.physics.quantum.qubit import Qubit
612
    >>> from sympy.physics.quantum.gate import XGate
613
    >>> q = Qubit('0101')
614
    >>> q
615
    |0101>
616
    >>> X = XGate(1)
617
    >>> qapply(X*q)
618
    |0111>
619
    Qubit states can also be used in adjoint operations, tensor products, inner/outer products:
620
    >>> Dagger(q)
621
    <0101|
622
    >>> ip = Dagger(q)*q
623
    >>> ip
624
625
    <0101|0101>
    >>> ip.doit()
626
    1
627
```

Quantum gates (unitary operators) can be applied to transform these states and then classical measurements can be performed on the results:

```
>>> from sympy.physics.quantum.qubit import measure all
630
    >>> from sympy.physics.quantum.gate import H, X, Y, Z
631
    >>> c = H(0)*H(1)*Qubit('00')
632
    >>> c
633
    H(0) * H(1) * |00>
634
    >>> q = qapply(c)
635
    >>> measure all(q)
636
    [(|00>, 1/4), (|01>, 1/4), (|10>, 1/4), (|11>, 1/4)]
637
```

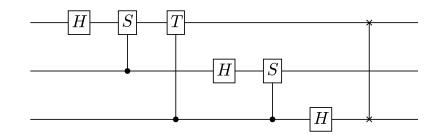


Figure 1. The circuit diagram for a three-qubit quantum Fourier transform generated by SymPy.

Lastly, the following example demonstrates creating a three-qubit quantum Fourier transform,

decomposing it into one- and two-qubit gates, and then generating a circuit plot for the sequence
 of gates (see Figure 1).

```
641 >>> from sympy.physics.quantum.qft import QFT
642 >>> from sympy.physics.quantum.circuitplot import circuit_plot
643 >>> fourier = QFT(0,3).decompose()
644 >>> fourier
645 SWAP(0,2)*H(0)*C((0),S(1))*H(1)*C((0),T(2))*C((1),S(2))*H(2)
646 >>> c = circuit_plot(fourier, nqubits=3)
```

647 6 CONCLUSION AND FUTURE WORK

⁶⁴⁸ SymPy is a robust computer algebra system that provides a wide spectrum of features both in ⁶⁴⁹ traditional computer algebra and in a plethora of scientific disciplines. This allows SymPy to be ⁶⁵⁰ used in a first-class way with other Python projects, including the scientific Python stack. Unlike ⁶⁵¹ many other CASs, SymPy is designed to be used in an extensible way: both as an end-user ⁶⁵² application and as a library.

SymPy expressions are immutable trees of Python objects. SymPy uses Python both as the internal language and the user language. This permits users to access to the same methods that the library implements in order to extend it for their needs. Additionally, SymPy has a powerful assumptions system for declaring and deducing mathematical properties of expressions.

SymPy has submodules for many areas of mathematics. This includes functions for simplifying expressions, performing common calculus operations, pretty printing expressions, solving equations, and representing symbolic matrices. Other included areas are discrete math, concrete math, plotting, geometry, statistics, polynomials, sets, series, vectors, combinatorics, group theory, code generation, tensors, Lie algebras, cryptography, and special functions. Additionally, SymPy contains submodules targeting certain specific domains, such as classical mechanics and quantum mechanics. This breadth of domains has been engendered by a strong and vibrant user
 community. Anecdotally, these users likely chose SymPy because of its ease of access.

Some of the planned future work for SymPy includes work on improving code generation,

improvements to the speed of SymPy (one area of work in this direction is SymEngine, a C++

⁶⁶⁷ symbolic manipulation library that is planned to be usable as a alternative core for SymPy),

 $_{\rm 668}$ $\,$ improving the assumptions system, and improving the solvers module.

669 7 ACKNOWLEDGEMENTS

The Los Alamos National Laboratory is operated by Los Alamos National Security, LLC, for the

National Nuclear Security Administration of the U.S. Department of Energy under Contract No.
 DE-AC52-06NA25396.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin
 Company, for the United States Department of Energy's National Nuclear Security Administration
 under Contract DE-AC04-94AL85000.

Google Summer of Code is an international annual program in which Google awards stipends to all students who successfully complete a requested free and open-source software coding project during the summer.

The author of this paper Francesco Bonazzi thanks the Deutsche Forschungsgemeinschaft (DFG) for its financial support via the International Research Training Group 1524 "Self-Assembled Soft Matter Nano-Structures at Interfaces."

682 **REFERENCES**

[1] Adams, W. W. and Loustaunau, P. (1994). An introduction to Gröbner bases. Number 3.
 American Mathematical Society.

- ⁶⁸⁵ ^[2] Bailey, D. H., Jeyabalan, K., and Li, X. S. (2005). A comparison of three high-precision quadrature schemes. *Experimental Mathematics*, 14(3):317–329.
- ⁶⁸⁷ ^[3] Bender, C. M. and Orszag, S. A. (1999). Advanced Mathematical Methods for Scientists and ⁶⁸⁸ Engineers. Springer, 1st edition.
- ⁶⁸⁹ ^[4] Biggs, N., Lloyd, E. K., and Wilson, R. J. (1976). *Graph Theory*, 1736-1936. Oxford ⁶⁹⁰ University Press.
- ⁶⁹¹ ^[5] Bronstein, M. (2005a). pmint—The Poor Man's Integrator.
- ⁶⁹² ^[6] Bronstein, M. (2005b). Symbolic Integration I: Transcendental Functions. Springer-Verlag,
 ⁶⁹³ New York, NY, USA.
- ⁶⁹⁴ ^[7] Buchberger, B. (1965). Ein Algorithmus zum Auffinden der Basis Elemente des Restklassenrings nach einem nulldimensionalen Polynomideal. PhD thesis, University of Innsbruck,
- ⁶⁹⁶ Innsbruck, Austria.
- [8] Carette, J. (2004). Understanding Expression Simplification. In ISSAC '04: Proceedings of the 2004 International Symposium on Symbolic and Algebraic Computation, pages 72–79, New York, NY, USA. ACM Press.
- ⁷⁰⁰ [9] Cervone, D. (2012). Mathjax: a platform for mathematics on the web. Notices of the AMS, 59(2):312-316.
- [10] Cimrman, R. (2014). SfePy write your own FE application. In de Buyl, P. and Varoquaux,
 N., editors, *Proceedings of the 6th European Conference on Python in Science (EuroSciPy* 2013), pages 65-70. http://arxiv.org/abs/1404.6391.
- ⁷⁰⁵ [11] Faugère, J. C. (1999). A New Efficient Algorithm for Computing Gröbner Bases (F4).
 ⁷⁰⁶ Journal of Pure and Applied Algebra, 139(1-3):61-88.
- ⁷⁰⁷ ^[12] Faugère, J. C. (2002). A New Efficient Algorithm for Computing Gröbner Bases Without
 ⁷⁰⁸ Reduction To Zero (F5). In *ISSAC '02: Proceedings of the 2002 International Symposium on*

⁷⁰⁹ Symbolic and Algebraic Computation, pages 75–83, New York, NY, USA. ACM Press.

- ⁷¹⁰ [13] Fetter, A. and Walecka, J. (2003). *Quantum Theory of Many-Particle Systems*. Dover ⁷¹¹ Publications.
- ⁷¹² ^[14] Fu, H., Zhong, X., and Zeng, Z. (2006). Automated and Readable Simplification of
- Trigonometric Expressions. *Mathematical and Computer Modelling*, 55(11-12):1169–1177.

- ⁷¹⁴ ^[15] Gede, G., Peterson, D. L., Nanjangud, A. S., Moore, J. K., and Hubbard, M. (2013).
 ⁷¹⁵ Constrained multibody dynamics with Python: From symbolic equation generation to publica-
- tion. In ASME 2013 International Design Engineering Technical Conferences and Computers
- and Information in Engineering Conference, pages V07BT10A051–V07BT10A051. American
 Society of Mechanical Engineers.
- ⁷¹⁹ ^[16] Goldberg, D. (1991). What every computer scientist should know about floating-point ⁷²⁰ arithmetic. ACM Computing Surveys (CSUR), 23(1):5–48.
- ⁷²¹ ^[17] Gosper, R. W. (1978). Decision procedure for indefinite hypergeometric summation. *Pro-*⁷²² *ceedings of the National Academy of Sciences*, 75(1):40–42.
- [18] Gruntz, D. (1996). On Computing Limits in a Symbolic Manipulation System. PhD thesis,
 Swiss Federal Institute of Technology, Zürich, Switzerland.
- ⁷²⁵ ^[19] Horsen, C. V. (2015). GMPY. https://pypi.python.org/pypi/gmpy2.
- [20] Hudak, P. (1998). Domain specific languages. In Handbook of Programming Languages, Vol.
 III: Little Languages and Tools, chapter 3, pages 39–60. MacMillan, Indianapolis.
- [21] Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. Computing In Science & Engineering, 9(3):90-95.
- ⁷³⁰ ^[22] Kane, T. R. and Levinson, D. A. (1985). *Dynamics, Theory and Applications*. McGraw Hill.
- ⁷³¹ [23] Lagrange, J. (1811). Mécanique analytique. Number v. 1 in Mécanique analytique. Ve
 ⁷³² Courcier.
- ⁷³³ ^[24] Lang, S. (1966). Introduction to transcendental numbers. *Reading, Mass.*
- ⁷³⁴ ^[25] Lutz, M. (2013). *Learning Python*. O'Reilly Media, Inc.
- ⁷³⁵ ^[26] Nielsen, M. and Chuang, I. (2011). Quantum Computation and Quantum Information.
 ⁷³⁶ Cambridge University Press.
- ⁷³⁷ ^[27] Nijenhuis, A. and Wilf, H. S. (1978). *Combinatorial Algorithms: For Computers and* ⁷³⁸ *Calculators*. Academic Press, New York, NY, USA, second edition.
- ⁷³⁹ [28] Oliphant, T. E. (2007). Python for scientific computing. Computing in Science & Engineering,
 9(3):10-20.
- ⁷⁴¹ ^[29] Paprocki, M. (2010). Design and implementation issues of a computer algebra system in
 ⁷⁴² an interpreted, dynamically typed programming language. Master's thesis, University of
 ⁷⁴³ Technology of Wrocław, Poland.
- [30] Pérez, F. and Granger, B. E. (2007). IPython: a system for interactive scientific computing.
 Computing in Science & Engineering, 9(3):21-29.
- [31] Peterson, D. L., Gede, G., and Hubbard, M. (2014). Symbolic linearization of equations of motion of constrained multibody systems. *Multibody System Dynamics*, 33(2):143-161.
- ⁷⁴⁸ ^[32] Petkovšek, M., Wilf, H. S., and Zeilberger, D. (1996). A= BAK peters. Wellesley, MA.
- [33] Raymond, E. (1999). The cathedral and the bazaar. Knowledge, Technology & Policy, 12(3):23-49.
- ⁷⁵¹ ^[34] Roach, K. (1996). Hypergeometric function representations. In ISSAC '96: Proceedings of the 1996 International Symposium on Symbolic and Algebraic Computation, pages 301–308,
- ⁷⁵³ New York, NY, USA. ACM Press.
- ⁷⁵⁴ ^[35] Roach, K. (1997). Meijer G function representations. In *ISSAC '97: Proceedings of the* ⁷⁵⁵ *1997 international symposium on Symbolic and algebraic computation*, pages 205–211, New
- York, NY, USA. ACM.
- [36] Rocklin, M. and Terrel, A. R. (2012). Symbolic statistics with SymPy. Computing in Science and Engineering, 14.
- ⁷⁵⁹ ^[37] Rosen, L. (2005). Open source licensing, volume 692. Prentice Hall.
- ⁷⁶⁰ ^[38] Sakurai, J. and Napolitano, J. (2010). *Modern Quantum Mechanics*. Addison-Wesley.
- [39] Shaw, M. and Garlan, D. (1996). Software Architecture: Perspectives on an Emerging
 Discipline. Prentice Hall. Prentice Hall Ordering Information.
- [40] Stein, W. and Joyner, D. (2005). SAGE: System for Algebra and Geometry Experimentation.
 Communications in Computer Algebra, 39(2).
- ⁷⁶⁵ [41] Sussman, G. J. and Wisdom, J. (2013). Functional Differential Geometry. Massachusetts
 ⁷⁶⁶ Institute of Technology Press.
- [42] Tai, C.-T. (1997). Generalized vector and dyadic analysis: applied mathematics in field
 theory, volume 9. Wiley-IEEE Press.

- ^[43] Takahasi, H. and Mori, M. (1974). Double exponential formulas for numerical integration. 769 Publications of the Research Institute for Mathematical Sciences, 9(3):721-741. 770
- ^[44] Toth, V. T. (2007). Maple and Meijer's G-function: a numerical instability and a cure. 771 http://www.vttoth.com/CMS/index.php/technical-notes/67.
- 772
- ^[45] Turk, M. J., Smith, B. D., Oishi, J. S., Skory, S., Skillman, S. W., Abel, T., and Norman, 773 M. L. (2011). yt: A Multi-code Analysis Toolkit for Astrophysical Simulation Data. The 774
- Astrophysical Journal Supplement Series, 192:9-+. 775
- ^[46] Zare, R. (1991). Angular Momentum: Understanding Spatial Aspects in Chemistry and 776 Physics. Wiley. 777