

# SYMPY: SYMBOLIC COMPUTING IN PYTHON

ONDŘEJ ČERTÍK\*, ISURU FERNANDO†, AND ASHUTOSH SABOO‡

**1. Introduction.** SymPy is a full featured computer algebra system (CAS) written in the Python programming language. It is open source, licensed under the extremely permissive 3-clause BSD license. SymPy was started by Ondřej Čertík in 2005, and it has since grown into a large open source project, with over 500 contributors. SymPy is developed on GitHub using a bazaar community model [36]. The accessibility of the codebase and the open community model allow SymPy to rapidly respond to the needs of the community of users, and has made the large contributor count possible.

SymPy is written entirely in the Python programming language. Python is a popular dynamically typed programming language that has a focus on ease of use and readability. It also a very popular language for scientific computing and data science, with a wide range of useful libraries [31]. SymPy is itself used by many libraries and tools across many domains, such as Sage [39] (pure mathematics), yt [42] (astronomy and astrophysics), PyDi (multibody dynamics), and SfePy [16] (finite elements).

Unlike many CASs, SymPy does not invent its own programming language. Python is used both for the internal implementation and the user interaction. Exclusively using Python in this way makes it easier for people already familiar with the language to use or develop SymPy. It also lets the SymPy developers focus on mathematics, rather than language design.

SymPy is designed with a strong focus that it be usable as a library. This means that extensibility is important in its application program interface (API) design. This is also one of the reasons SymPy makes no attempt to extend the Python language itself. The goal is for users of SymPy to be able to import SymPy alongside other Python libraries in their workflow, whether that is an interactive workflow or programmatic use as part of a larger system.

SymPy does not have a built in graphical user interface (GUI), however, when used in the Jupyter Notebook SymPy expressions will pretty print using MathJax.

Section 2 discusses the architecture of SymPy. Following that, Section 3 looks at the numerical features of SymPy and its dependency library, mpmath. Section 4 enumerates the features of SymPy and takes a closer look at some of the important ones. Section 5 looks at the domain specific submodules for doing classical mechanics and quantum mechanics. Finally, Section 6 concludes the paper and discusses future work.

## 2. Architecture.

**2.1. Basic Usage.** Being built on Python, SymPy requires that all variable names be defined before they can be used. The statement

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

will import all SymPy functions into the global Python namespace. All the examples in this paper assume that this has been run.

---

\*Los Alamos National Laboratory ([ondrej.certik@gmail.com](mailto:ondrej.certik@gmail.com)).

†University of Moratuwa ([isuru.11@cse.mrt.ac.lk](mailto:isuru.11@cse.mrt.ac.lk)).

‡Birla Institute of Technology and Science, Pilani, K.K. Birla Goa Campus ([ashutosh.saboo@gmail.com](mailto:ashutosh.saboo@gmail.com)).

Additionally, symbolic variables, called symbols, must be assigned to Python variables before they can be used. This is typically done through the `symbols` function, which creates multiple symbols at once. For instance,

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

creates three symbols named `x`, `y`, and `z`, assigned to Python variables of the same name. The Python variable names that symbols are assigned to are immaterial—we could have just as well have written `a, b, c = symbol('x y z')`. All the examples in this paper will assume that the symbols `x`, `y`, and `z` have been assigned as above.

Expressions are created from symbols using Python syntax, which mirrors usual mathematical notation. Note that in Python, exponentiation is `**`.

```
>>> (x**2 - 2*x + 3)/y
(x**2 - 2*x + 3)/y
```

All SymPy expressions are immutable. This simplifies the design by allowing interning. It also allows expressions to be hashed and stored in a Python dictionary, which enables caching and other features.

**2.2. The Core.** The core of a computer algebra system (CAS) refers to the module that is in charge of resending symbolic expressions and performing basic manipulations with them. In SymPy, every symbolic expression is an instance of a Python class. Expressions are represented by expression trees. The operators are represented by the type of an expression and the child nodes are stored in the `args` attribute. A leaf node in the expression tree has an empty `args`. The `args` attribute is provided by the class `Basic`, which is a superclass of all SymPy objects and provides common methods to all SymPy tree-elements. For example, consider the expression  $xy + 2$ :

```
>>> from sympy import *
>>> x, y = symbols('x y')
>>> expr = x*y + 2
```

The expression `expr` is an addition, so it is of type `Add`. The child nodes of `expr` are `x*y` and `2`.

```
>>> type(expr)
<class 'sympy.core.add.Add'>
>>> expr.args
(2, x*y)
```

We can dig further into the expression tree to see the full expression. For example, the first child node, given by `expr.args[0]` is `2`. Its class is `Integer`, and it has empty `args`, indicating that it is a leaf node.

```
>>> expr.args[0]
2
>>> type(expr.args[0])
<class 'sympy.core.numbers.Integer'>
>>> expr.args[0].args
()
```

The function `srepr` gives a string representing a valid Python code, containing all the nested class constructor calls to create the given expression.

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

Every SymPy expression satisfies a key invariant, namely `expr.func(*expr.args) == expr`. This means that expressions are rebuildable from their `args`<sup>1</sup>. Here, we note that

---

<sup>1</sup>`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.

in SymPy, the `==` operator represents exact structural equality, not mathematical equality. This allows one to test if any two expressions are equal to one another as expression trees.

Python allows classes to overload operators. The Python interpreter translates the above `x*y + 2` to, roughly, `(x.__mul__(y)).__add__(2)`. `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 builtin `int` type. When `2` is called by the `__add__` method, it is converted to the SymPy type `Integer(2)`. In this way, SymPy expressions can be built in the natural way using Python operators and numeric literals.

One must be careful in one particular instance. Python does not have a builtin rational literal type. Given a fraction of integers such as `1/2`, Python will perform floating point division and produce `0.5`<sup>2</sup>. Python uses eager evaluation, so expressions like `x + 1/2` will produce `x + 0.5`, and by the time any SymPy function sees the `1/2` it has already been converted to `0.5` by Python. However, for a CAS like SymPy, one typically wants to work with exact rational numbers whenever possible. Working around this is simple, however: one can wrap one of the integers with `Integer`, like `x + Integer(1)/2`, or using `x + Rational(1, 2)`. SymPy provides a function `S` which can be used to convert objects to SymPy types with minimal typing, such as `x + S(1)/2`. This gotcha is a small downside to using Python directly instead of a custom domain specific language (DSL), and we consider it to be worth it for the advantages listed above.

**2.3. Assumptions.** An important feature of the SymPy core is the assumptions system. The assumptions system allows users to specify that symbols have certain common mathematical properties, such as being positive, imaginary, or integer. SymPy is careful to never perform simplifications on an expression unless the assumptions allow them. For instance, the identity  $\sqrt{x^2} = x$  holds if  $x$  is nonnegative ( $x \geq 0$ ). If  $x$  is real, the identity  $\sqrt{x^2} = |x|$  holds. However, for general complex  $x$ , no such identity holds.

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

```
>>> x = Symbol('x')
>>> sqrt(x**2)
sqrt(x**2)
```

By assuming symbols are complex by default, SymPy avoids performing mathematically invalid operations. However, in many cases users will wish to simplify expressions containing terms like  $\sqrt{x^2}$ .

Assumptions are set on `Symbol` objects when they are created. For instance `Symbol('x', positive=True)` will create a symbol named `x` that is assumed to be positive.

```
>>> x = Symbol('x', positive=True)
>>> sqrt(x**2)
x
```

Some common assumptions that SymPy allows are `positive`, `negative`, `real`, `nonpositive`, `nonnegative`, `real`, `integer`, and `commutative`<sup>3</sup>. Assumptions on

<sup>2</sup>This is the behavior in Python 3. In Python 2, `1/2` will perform integer division and produce `0`, unless one uses `from __future__ import division`.

<sup>3</sup>If  $A$  and  $B$  are Symbols created with `commutative=False` then SymPy will keep  $A \cdot B$  and  $B \cdot A$

any object can be checked with the `is_assumption` attributes, like `x.is_positive`.

Assumptions are only needed to restrict a domain so that certain simplifications can be performed. It is 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 builtin objects `True`, `False`, and `None`. `None` represents the “unknown” case. This could mean that the given assumption could be either true or false under the given information, for instance, `Symbol('x', real=True).is_positive` will give `None` because a real symbol might be positive or it might not. It could also mean not enough is implemented to compute the given fact, for instance, `(pi + E).is_irrational` gives `None`, because SymPy does not know how to determine if  $\pi + e$  is rational or irrational, indeed, it is an open problem in mathematics.

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`.

SymPy also has an experimental assumptions system where facts are stored separate from objects, and deductions are made with a SAT solver. We will not discuss this system here.

**2.4. Extensibility.** Extensibility is an important feature for SymPy. Because the same language, Python, is used both for the internal implementation and the external usage by users, all the extensibility capabilities available to users are also used by functions that are part of SymPy.

The typical way to create a custom SymPy object is to subclass an existing SymPy class, generally either `Basic`, `Expr`, or `Function`. All SymPy classes used for expression trees<sup>4</sup> 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 values for which the function should be automatically evaluated, and `None` for arguments that shouldn’t be automatically evaluated.

The behavior of classes in SymPy with various other SymPy functions is defined by defining a relevant `_eval_*` method on the class. For instance, an object can tell SymPy’s `diff` function how to take the derivative of itself by defining the `_eval_derivative(self, x)` method. The most common `_eval_*` methods relate to the assumptions. `_eval_is_assumption` defines the assumptions for *assumption*.

---

distinct.

<sup>4</sup>Some internal classes, such as those used in the polynomial module, do not follow this rule for efficiency reasons.

181 As an example of the notions presented in this section, we present below a stripped  
 182 down version of the gamma function  $\Gamma(x)$  from SymPy, which evaluates itself on  
 183 positive integer arguments, has the positive and real assumptions defined, can be  
 184 rewritten in terms of factorial with `gamma(x).rewrite(factorial)`, and can be dif-  
 185 ferentiated. `fdiff` is a convenience method for subclasses of `Function`. `fdiff` returns  
 186 the derivative of the function without worrying about the chain rule. `self.func` is  
 187 used throughout instead of referencing `gamma` explicitly so that potential subclasses  
 188 of `gamma` can reuse the methods.

```
189 from sympy import Integer, Function, floor, factorial, polygamma
190
191 class gamma(Function)
192     @classmethod
193     def eval(cls, arg):
194         if isinstance(arg, Integer) and arg.is_positive:
195             return factorial(arg - 1)
196
197     def _eval_is_real(self):
198         x = self.args[0]
199         # noninteger means real and not integer
200         if x.is_positive or x.is_noninteger:
201             return True
202
203     def _eval_is_positive(self):
204         x = self.args[0]
205         if x.is_positive:
206             return True
207         elif x.is_noninteger:
208             return floor(x).is_even
209
210     def _eval_rewrite_as_factorial(self, z):
211         return factorial(z - 1)
212
213     def fdiff(self, argindex=1):
214         from sympy.core.function import ArgumentIndexError
215         if argindex == 1:
216             return self.func(self.args[0])*polygamma(0, self.args[0])
217         else:
218             raise ArgumentIndexError(self, argindex)
```

219 The actual gamma function defined in SymPy has many more capabilities, such  
 220 as evaluation at rational points and series expansion.

221 **3. Numerics.** The `Float` class holds an arbitrary-precision binary floating-point  
 222 value and a precision in bits. An operation between two `Float` inputs is rounded to  
 223 the larger of the two precisions. Since Python floating-point literals automatically  
 224 evaluate to `double` (53-bit) precision, strings should be used to input precise decimal  
 225 values:

```
226 >>> Float(1.1)
227 1.1000000000000000
228 >>> Float(1.1, 30) # precision equivalent to 30 digits
229 1.100000000000000008881784197001
```



improving performance.

The mpmath library includes support for 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 support computing hundreds of digits rapidly, except in degenerate cases.

The double exponential (tanh-sinh) quadrature is used for numerical integration by default. For smooth integrands, this algorithm usually converges extremely rapidly, even when the integration interval is infinite or singularities are present at the end-points [40, 9]. However, for good performance, singularities in the middle of the interval must be specified by the user. To evaluate slowly converging limits and infinite series, mpmath automatically attempts to apply Richardson extrapolation and the Shanks transformation (Euler-Maclaurin summation can also be used) [10]. A function to evaluate oscillatory integrals by means of convergence acceleration is also available.

A wide array of higher mathematical functions are implemented with full support for complex values of all parameters and arguments, including complete and incomplete gamma functions, Bessel functions, orthogonal polynomials, elliptic functions and integrals, zeta and polylogarithm functions, the generalized hypergeometric function, and the Meijer G-function.

Most special functions are implemented as linear combinations of the generalized hypergeometric function  ${}_pF_q$ , which is computed by a combination of direct summation, argument transformations (for  ${}_2F_1$ ,  ${}_3F_2$ , ...) and asymptotic expansions (for  ${}_0F_1$ ,  ${}_1F_1$ ,  ${}_1F_2$ ,  ${}_2F_2$ ,  ${}_2F_3$ ) to cover the whole complex domain. Numerical integration and generic convergence acceleration are also used in a few special cases.

In general, linear combinations and argument transformations give rise to singularities that have to be removed for certain combinations of parameters. A typical example is the modified Bessel function of the second kind

$$K_\nu(z) = \frac{1}{2} \left[ \left( \frac{z}{2} \right)^{-\nu} \Gamma(\nu) {}_0F_1 \left( 1 - \nu, \frac{z^2}{4} \right) - \left( \frac{z}{2} \right)^\nu \frac{\pi}{\nu \sin(\pi\nu) \Gamma(\nu)} {}_0F_1 \left( \nu + 1, \frac{z^2}{4} \right) \right]$$

where the limiting value  $\lim_{\epsilon \rightarrow 0} K_{n+\epsilon}(z)$  has to be computed when  $\nu = n$  is an integer. A generic algorithm is used to evaluate hypergeometric-type linear combinations of the above type. This algorithm automatically detects cancellation problems, and computes limits numerically by perturbing parameters whenever internal singularities occur (the perturbation size is automatically decreased until the result is detected to converge numerically).

Due to this generic approach, particular combinations of hypergeometric functions can be specified easily. The implementation of the Meijer G-function takes only a few dozen lines of code, yet covers the whole input domain in a robust way. The Meijer G-function instance  $G_{1,3}^{3,0} \left( 0; \frac{1}{2}, -1, -\frac{3}{2} | x \right)$  is a good test case [41]; past versions of both Maple and Mathematica produced incorrect numerical values for large  $x > 0$ . Here, mpmath automatically removes the internal singularity and compensates for cancellations (amounting to 656 bits of precision when  $x = 10000$ ), giving correct values:

```
>>> mpmath.mp.dps = 15
>>> mpmath.meijerg([], [0], [[-0.5, -1, -1.5], []], 10000)
mpf('2.4392576907199564e-94')
```

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

```

325 >>> meijerg([], [0], [[-S(1)/2, -1, -S(3)/2], []], 10000).evalf()
326 2.43925769071996e-94

```

327 We highlight the generalized hypergeometric functions and the Meijer G-function,  
328 due to those functions' frequent appearance in closed forms for integrals and sums  
329 Via mpmath, SymPy has relatively good support for evaluating sums and integrals  
330 numerically, using two complementary approaches: direct numerical evaluation, or  
331 first computing a symbolic closed form involving special functions.

332 **3.2. Numerical simplification.** The `nsimplify` function in SymPy (a wrapper  
333 of `identify` in mpmath) attempts to find a simple symbolic expression that evaluates  
334 to the same numerical value as the given input. It works by applying a few simple  
335 transformations (including square roots, reciprocals, logarithms and exponentials) to  
336 the input and, for each transformed value, using the PSLQ algorithm [18] to search  
337 for a matching algebraic number or optionally a linear combination of user-provided  
338 base constants (such as  $\pi$ ).

```

339 >>> x = 1 / (sin(pi/5)+sin(2*pi/5)+sin(3*pi/5)+sin(4*pi/5))*2
340 >>> nsimplify(x)
341 -2*sqrt(5)/5 + 1
342 >>> nsimplify(pi, tolerance=0.01)
343 22/7
344 >>> nsimplify(1.783919626661888, [pi], tolerance=1e-12)
345 pi/(-1/3 + 2*pi/3)

```

346 **4. Features.** SymPy has an extensive feature set that encompasses too much to  
347 cover in-depth here. Bedrock areas, such as calculus, receive their own sub-sections  
348 below. Additionally, Table 1 describes other capabilities present in the SymPy code  
349 base. This gives a sampling from the breadth of topics and application domains that  
350 SymPy services.

Table 1: SymPy Features and Descriptions

Feature	Description
Discrete Math	Summations, products, binomial coefficients, prime number tools, integer factorization, Diophantine equation solving, and boolean logic representation, equivalence testing, and inference.
Concrete Math	Tools for determining whether summation and product expressions are convergent, absolutely convergent, hypergeometric, and other properties. May also compute Gosper's normal form [35] for two univariate polynomials.
Plotting	Hooks for visualizing expressions via matplotlib [?] or as text drawings when lacking a graphical back-end.
Geometry	Allows the creation of 2D geometrical entities, such as lines and circles. Enables queries on these entities, including asking the area of an ellipse, checking for collinearity of a set of points, or finding the intersection between two lines.



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.
Polynomials	Computes polynomial algebras over various coefficient domains ranging from the simple (e.g., polynomial division) to the advanced (e.g., Gröbner bases [8] and multivariate factorization over algebraic number domains).
Sets	Representations of empty, finite, and infinite sets. This includes special sets such as for all natural, integer, and complex numbers.
Series	Implements series expansion, sequences, and limit of sequences. This includes special series, such as Fourier and power series.
Vectors	Provides basic vector math and differential calculus with respect to 3D Cartesian coordinate systems.
Matrices	Tools for creating matrices of symbols and expressions. This is capable of both sparse and dense representations and performing symbolic linear algebraic operations (e.g., inversion and factorization).
Combinatorics & Group Theory	Implements permutations, combinations, partitions, subsets, various permutation groups (such as polyhedral, Rubik, symmetric, and others), Gray codes [30], and Prufer sequences [11].
Code Generation	Enables generation of compilable and executable code in a variety of different programming languages directly from expressions. Target languages include C, Fortran, Julia, JavaScript, Mathematica, Matlab and Octave, Python, and Theano.
Tensors	Symbolic manipulation of indexed objects.
Lie Algebras	Represents Lie algebras and root systems.
Cryptography	Represents block and stream ciphers, including shift, Affine, substitution, Vigenere's, Hill's, bifid, RSA, Kid RSA, linear-feedback shift registers, and Elgamal encryption

## Special Functions

Implements 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, polylogarithm, Lerch transcendent, hypergeometric, elliptic integrals, Mathieu, Jacobi polynomials, Gegenbauer polynomial, Chebyshev polynomial, Legendre polynomial, Hermite polynomial, Laguerre polynomial, and spherical harmonic functions.

**4.1. Simplification.** The generic way to simplify an expression is by calling the `simplify` function. It must be emphasized that simplification is not an unambiguously defined mathematical operation [15]. The `simplify` function applies several simplification routines along with some heuristics to make the output expression as “simple” as possible.

It is often preferable to apply more directed simplification functions. These apply very specific rules to the input expression, and are often 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 some common simplification functions.

Table 2: SymPy Simplification Functions

<code>expand</code>	expand the expression
<code>factor</code>	factor a polynomial into irreducibles
<code>collect</code>	collect polynomial coefficients
<code>cancel</code>	rewrite a rational function as $p/q$ with common factors canceled
<code>apart</code>	compute the partial fraction decomposition of a rational function
<code>trigsimp</code>	simplify trigonometric expressions [19]

Substitutions are performed through the `.subs` method, which is sensible to some mathematical properties while matching, such as associativity, commutativity, additive and multiplicative inverses, and matching of powers.

**4.2. Calculus.** Derivatives can be computed with the `diff` function.

```
>>> diff(sin(x), x)
```

```
cos(x)
```

Unevaluated `Derivative` objects are also supported.

```
>>> expr = Derivative(sin(x), x)
```

```
>>> expr
```

```
Derivative(sin(x), x)
```

Unevaluated expressions can be evaluated with the `doit` method.

```
>>> expr.doit()
```

```
cos(x)
```

Integrals can be analogously calculated either with the `integrate` function, or the unevaluated `Integral` objects.

```

376 >>> integrate(sin(x), x)
377 -cos(x)
378 >>> expr = Integral(sin(x), x)
379 >>> expr
380 Integral(sin(x), x)
381 >>> expr.doit()
382 -cos(x)
383 Definite integration can be calculated with the same method, by specifying a range
384 of the integration variable. The following computes  $\int_0^1 \sin(x) dx$ .
385 >>> integrate(sin(x), (x, 0, 1))
386 -cos(1) + 1
387 SymPy implements a combination of the Risch algorithm [14], table lookups, a
388 reimplement of Manuel Bronstein's "Poor Man's Integrator" [13], and an algo-
389 rithm for computing integrals based on Meijer G-functions. These allow SymPy to
390 compute a wide variety of indefinite and definite integrals.
391 Summations and products are also supported, via the evaluated summation and
392 product and unevaluated Sum and Product, and use the same syntax as integrate.
393 Summations are computed using a combination of Gosper's algorithm and an algo-
394 rithm that uses Meijer G-functions. Products are computed via some heuristics.
395 The limit module implements the Gruntz algorithm [22] for computing symbolic
396 limits. For example, the following computes  $\lim_{x \rightarrow \infty} x \sin(\frac{1}{x}) = 1$  (note that  $\infty$  is oo in
397 SymPy).
398 >>> limit(x*sin(1/x), x, oo)
399 1
400 As a more complicated example, SymPy computes  $\lim_{x \rightarrow 0} \left( 2e^{\frac{1-\cos(x)}{\sin(x)}} - 1 \right)^{\frac{\sinh(x)}{\operatorname{atan}^2(x)}} = e$ .
401 >>> limit((2*E**((1-cos(x))/sin(x))-1)**(sinh(x)/atan(x)**2), x, 0)
402 E

```

**4.3. Printers.** SymPy has a rich collection of expression printers for displaying expressions to the user. 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.

```

406 >>> phi0 = Symbol('phi0')
407 >>> str(Integral(sqrt(phi0), phi0))
408 Integral(sqrt(phi0 + 1), x)
409 Expressions can be printed with 2D monospace text with pprint. This uses
410 Unicode characters to render mathematical symbols such as integral signs, square
411 roots, and parentheses. Greek letters and subscripts in symbol names are rendered
412 automatically.

```

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

```

415 >>> pprint(Integral(sqrt(phi0 + 1), phi0), use_unicode=False)
416 /
417 |
418 | -----
419 | \ / phi0 + 1  d(phi0)
420 |
421 /

```

The function `latex` returns a  $\text{\LaTeX}$  representation of an expression.

```

422 >>> print(latex(Integral(sqrt(phi0 + 1), phi0)))
423

```

424  $\int \sqrt{\phi_0 + 1} dx$

425 Users are encouraged to run the `init_printing` function at the beginning of  
 426 interactive sessions, which automatically enables the best pretty printing supported  
 427 by their environment. In the Jupyter notebook or qtconsole [33] the L<sup>A</sup>T<sub>E</sub>X printer is  
 428 used to render expressions using MathJax or L<sup>A</sup>T<sub>E</sub>X if it is installed on the system.  
 429 The 2D text representation is used otherwise.

430 Other printers such as MathML are also available. SymPy uses an extensible  
 431 printer subsystem which allows users to customize the printing for any given printer,  
 432 and for custom objects to define their printing behavior for any printer. SymPy's  
 433 code generation capabilities, which we will not discuss in-depth here, use the same  
 434 printer model.

435 **4.4. Solvers.** SymPy has module of equation solvers for symbolic equations.  
 436 There are two submodules to solve algebraic equations in SymPy, referred to as old  
 437 solve function, `solve`, and new solve function, `solveset`. Solveset is introduced with  
 438 several design changes with respect to old `solve` function to resolve the issues with  
 439 old `solve` function, for example old `solve` function's input API has many flags which  
 440 are not needed and they make it hard for the user and the developers to work on  
 441 solvers. In contrast to old solve function, the `solveset` has a clean input API, It  
 442 only asks for the much needed information from the user, following are the function  
 443 signatures of old and new solve function:

444 `solve(f, *symbols, **flags)` # old solve function

445 `solveset(f, symbol, domain)` # new solve function

446 The old `solve` function has an inconsistent output API for various types of inputs,  
 447 whereas the `solveset` has a canonical output API which is achieved using sets. It  
 448 can consistently return various types of solutions.

449 • Single solution

450 `>>> solveset(x - 1)`

451 `{1}`

452 • Finite set of solution, quadratic equation

453 `>>> solveset(x**2 - pi**2, x)`

454 `{-pi, pi}`

455 • No Solution

456 `>>> solveset(1, x)`

457 `EmptySet()`

458 • Interval of solution

459 `>>> solveset(x**2 - 3 > 0, x, domain=S.Reals)`

460 `(-oo, -sqrt(3)) U (sqrt(3), oo)`

461 • Infinitely many solutions

462 `>>> solveset(sin(x) - 1, x, domain=S.Reals)`

463 `ImageSet(Lambda(_n, 2*_n*pi + pi/2), Integers())`

464 `>>> solveset(x - x, x, domain=S.Reals)`

465 `(-oo, oo)`

466 `>>> solveset(x - x, x, domain=S.Complexes)`

467 `S.Complexes`

468 • Linear system: finite and infinite solution for determined, under determined  
 469 and over determined problems.

470 `>>> A = Matrix([[1, 2, 3], [4, 5, 6], [7, 8, 10]])`

471 `>>> b = Matrix([3, 6, 9])`

472 `>>> linsolve((A, b), x, y, z)`

```

473 {(-1,2,0)}
474 >>> linsolve(Matrix([[1, 1, 1, 1], [1, 1, 2, 3]]), (x, y, z))
475 {(-y - 1, y, 2)}

```

476 The new solve i.e. **solveset** is under active development and is a planned replacement for **solve**. Hence there are some features which are implemented in solve and is not yet implemented in solveset. The table below show the current state of old and new solve functions.

480

Solveset vs Solve		
Feature	solve	solveset
Consistent Output API	No	Yes
Consistent Input API	No	Yes
Univariate	Yes	Yes
Linear System	Yes	Yes (linsolve)
Non Linear System	Yes	Not yet
Transcendental	Yes	Not yet

481

482

483

484 Below are some of the examples of old **solve** function:

485 • Non Linear (multivariate) System of Equation: Intersection of a circle and a parabola.

```

486
487 >>> solve([x**2 + y**2 - 16, 4*x - y**2 + 6], x, y)
488 [(-2 + sqrt(14), -sqrt(-2 + 4*sqrt(14))),
489  (-2 + sqrt(14), sqrt(-2 + 4*sqrt(14))),
490  (-sqrt(14) - 2, -I*sqrt(2 + 4*sqrt(14))),
491  (-sqrt(14) - 2, I*sqrt(2 + 4*sqrt(14)))]

```

492 • Transcendental Equation

```

493 >>> solve(x + log(x)**2 - 5*(x + log(x)) + 6, x)
494 [LambertW(exp(2)), LambertW(exp(3))]
495 >>> solve(x**3 + exp(x))
496 [-3*LambertW((-1)**(2/3)/3)]

```

497 **4.5. Matrices.** SymPy supports matrices with symbolic expressions as elements.

```

498 >>> x, y = symbols('x y')
499 >>> A = Matrix(2, 2, [x, x + y, y, x])
500 >>> A
501 Matrix([
502 [ x, x + y],
503 [ y, x]])

```

504 All SymPy matrix types can do linear algebra including matrix addition, multiplication, exponentiation, computing determinant, solving linear systems and computing inverses using LU decomposition, LDL decomposition, Gauss-Jordan elimination, Cholesky decomposition, Moore-Penrose pseudoinverse, and adjugate matrix.

508 All operations are computed are computed symbolically. Eigenvalues are computed by generating the characteristic polynomial using the Berkowitz algorithm and then solving it using polynomial routines. Diagonalizable matrices can be diagonalized first to compute the eigenvalues.

```

512 >>> A.eigenvals()
513 {x - sqrt(y*(x + y)): 1, x + sqrt(y*(x + y)): 1}

```

514 Internally these matrices store the elements as a list making it a dense representation. For storing sparse matrices, the **SparseMatrix** class can be used. Sparse

515

```

516 matrices store the elements in a dictionary of keys (DoK) format.
517     SymPy also supports matrices with symbolic dimension values. MatrixSymbol
518 represents a matrix with dimensions  $m \times n$ , where  $m$  and  $n$  can be symbolic. Matrix
519 addition and multiplication, scalar operations, matrix inverse and transpose are stored
520 symbolically as matrix expressions.
521 >>> m, n, p = symbols("m, n, p", integer=True)
522 >>> R = MatrixSymbol("R", m, n)
523 >>> S = MatrixSymbol("S", n, p)
524 >>> T = MatrixSymbol("t", m, p)
525 >>> U = R*S + 2*T
526 >>> u.shape
527 (m, p)
528 >>> U[0, 1]
529 2*T[0, 1] + Sum(R[0, _k]*S[_k, 1], (_k, 0, n - 1))
530     Block matrices are also supported in SymPy. BlockMatrix elements can be any
531 matrix expression which includes explicit matrices, matrix symbols, and block matrices. All functionalities of matrix expressions are also present in BlockMatrix.
532 >>> n, m, l = symbols('n m l')
533 >>> X = MatrixSymbol('X', n, n)
534 >>> Y = MatrixSymbol('Y', m, m)
535 >>> Z = MatrixSymbol('Z', n, m)
536 >>> B = BlockMatrix([[X, Z], [ZeroMatrix(m, n), Y]])
537 >>> B
538 Matrix([
539 [X, Z],
540 [0, Y]])
541 >>> B[0, 0]
542 X[0, 0]
543 >>> B.shape
544 (m + n, m + n)

```

546 **5. Domain Specific Submodules.** 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 classical mechanics, optics, and quantum mechanics along with support for manipulating physical quantities with units.

## 551 **5.1. Classical Mechanics.**

552 **5.1.1. Vector Algebra.** The `sympy.physics.vector` package provides reference frame, time, and space aware vector and dyadic objects that allow for three dimensional operations such as addition, subtraction, scalar multiplication, inner and outer products, cross products, etc. Both of these objects can be written in very compact notation that make it easy to express the vectors and dyadics in terms of multiple reference frames with arbitrarily defined relative orientations. The vectors are used to specify the positions, velocities, and accelerations of points, orientations, angular velocities, and angular accelerations of reference frames, and force and torques. The dyadics are essentially reference frame aware  $3 \times 3$  tensors. The vector and dyadic objects can be used for any one-, two-, or three-dimensional vector algebra and they provide a strong framework for building physics and engineering tools.

563 The following Python interpreter session showing how a vector is created using

```

564 the orthogonal unit vectors of three reference frames that are oriented with respect
565 to each other and the result of expressing the vector in the  $A$  frame. The  $B$  frame
566 is oriented with respect to the  $A$  frame using Z-X-Z Euler Angles of magnitude  $\pi$ ,  $\frac{\pi}{2}$ ,
567 and  $\frac{\pi}{3}$ rad, respectively whereas the  $C$  frame is oriented with respect to the  $B$  frame
568 through a simple rotation about the  $B$  frame's X unit vector through  $\frac{\pi}{2}$ rad.
569 >>> from sympy import pi
570 >>> from sympy.physics.vector import ReferenceFrame
571 >>> A = ReferenceFrame('A')
572 >>> B = ReferenceFrame('B')
573 >>> C = ReferenceFrame('C')
574 >>> B.orient(A, 'body', (pi, pi / 3, pi / 4), 'zxz')
575 >>> C.orient(B, 'axis', (pi / 2, B.x))
576 >>> v = 1 * A.x + 2 * B.z + 3 * C.y
577 >>> v
578 A.x + 2*B.z + 3*C.y
579 >>> v.express(A)
580 A.x + 5*sqrt(3)/2*A.y + 5/2*A.z

```

581 **5.1.2. Mechanics.** The `sympy.physics.mechanics` package utilizes the `sympy.`  
582 `physics.vector` package to populate time aware particle and rigid body objects to  
583 fully describe the kinematics and kinetics of a rigid multi-body system. These objects  
584 store all of the information needed to derive the ordinary differential or differential al-  
585 gebraic equations that govern the motion of the system, i.e., the equations of motion.  
586 These equations of motion abide by Newton's laws of motion and can handle any ar-  
587 bitrary kinematical constraints or complex loads. The package offers two automated  
588 methods for formulating the equations of motion based on Lagrangian Dynamics [26]  
589 and Kane's Method [25]. Lastly, there are automated linearization routines for con-  
590 strained dynamical systems based on [34].

591 **5.2. Quantum Mechanics.** The `sympy.physics.quantum` package provides  
592 quantum functions, states, operators, and computation of standard quantum mod-  
593 els.

594 **6. Conclusion and future work.** SymPy is a robust CAS that provides a wide  
595 array of features. It is written in a general purpose programming language, Python,  
596 which allows it to be used in a first-class way with other Python projects, including  
597 the scientific Python stack. It is designed to be used in an extensible way. Unlike  
598 many other CASs, it is designed to be used both as a end-user application and as a  
599 library.

600 SymPy expressions are built from immutable trees of Python classes. It uses  
601 Python both as the internal language and the user language, meaning users can use the  
602 same methods that the library implements to extend it. SymPy has an assumptions  
603 system for declaring and deducing mathematical properties on expressions.

604 The numerics of SymPy are implemented in the `mpmath` library, which uses  
605 arbitrary precision floating point arithmetic implemented in pure Python. This allows  
606 expressions to be evaluated with concrete data as needed.

607 SymPy has submodules for many areas of mathematics. It has functions for sim-  
608 plifying expressions, doing common calculus operations, pretty printing expressions,  
609 solving equations, and symbolic matrices. Other areas also included are discrete  
610 math, concrete math, plotting, geometry, statistics, polynomials, sets, series, vectors,  
611 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.

Some of the planned future work for SymPy includes work on improving code generation, improvements to the speed of SymPy, and improving the solvers module.

## 7. References.

## REFERENCES

- [1] <https://github.com/sympy/sympy/blob/master/doc/src/modules/polys/ringseries.rst>.
- [2] <https://reference.wolfram.com/language/ref/Flat.html>.
- [3] <https://reference.wolfram.com/language/ref/Orderless.html>.
- [4] <https://reference.wolfram.com/language/ref/OneIdentity.html>.
- [5] <https://reference.wolfram.com/language/tutorial/FlatAndOrderlessFunctions.html>.
- [6] *The software engineering of the wolfram system*, 2016, <https://reference.wolfram.com/language/tutorial/TheSoftwareEngineeringOfTheWolframSystem.html>.
- [7] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York, NY, USA, ninth printing ed., 1964, <http://www.math.ucla.edu/~cbm/aands/>.
- [8] W. W. ADAMS AND P. LOUSTAUNAU, *An introduction to Gröbner bases*, no. 3, American Mathematical Soc., 1994.
- [9] D. H. BAILEY, K. JEYABALAN, AND X. S. LI, *A comparison of three high-precision quadrature schemes*, *Experimental Mathematics*, 14 (2005), pp. 317–329.
- [10] C. M. BENDER AND S. A. ORSZAG, *Advanced Mathematical Methods for Scientists and Engineers*, Springer, 1st ed., October 1999.
- [11] N. BIGGS, E. K. LLOYD, AND R. J. WILSON, *Graph Theory, 1736-1936*, Oxford University Press, 1976.
- [12] R. P. BRENT AND P. ZIMMERMANN, *Modern Computer Arithmetic*, Cambridge University Press, version 0.5.1 ed.
- [13] M. BRONSTEIN, *Poor Man's Integrator*, <http://www-sop.inria.fr/cafe/Manuel.Bronstein/pmint>.
- [14] M. BRONSTEIN, *Symbolic Integration I: Transcendental Functions*, Springer-Verlag, New York, NY, USA, 2005.
- [15] J. CARETTE, *Understanding Expression Simplification*, in ISSAC '04: Proceedings of the 2004 International Symposium on Symbolic and Algebraic Computation, New York, NY, USA, 2004, ACM Press, pp. 72–79, <http://dx.doi.org/http://doi.acm.org/10.1145/1005285.1005298>.
- [16] R. CIRMANN, *SfePy - write your own FE application*, in Proceedings of the 6th European Conference on Python in Science (EuroSciPy 2013), P. de Buyl and N. Varoquaux, eds., 2014, pp. 65–70. <http://arxiv.org/abs/1404.6391>.
- [17] R. J. FATEMAN, *A review of Mathematica*, *Journal of Symbolic Computation*, 13 (1992), pp. 545–579, [http://dx.doi.org/DOI:10.1016/S0747-7171\(10\)80011-2](http://dx.doi.org/DOI:10.1016/S0747-7171(10)80011-2).
- [18] H. R. P. FERGUSON, D. H. BAILEY, AND S. ARNO, *Analysis of PSLQ, an integer relation finding algorithm*, *Mathematics of Computation*, 68 (1999), pp. 351–369.
- [19] H. FU, X. ZHONG, AND Z. ZENG, *Automated and Readable Simplification of Trigonometric Expressions*, *Mathematical and Computer Modelling*, 55 (2006), pp. 1169–1177.
- [20] Y. C. FUNG, *A first course in continuum mechanics*, Pearson, third edition ed., 1993.
- [21] D. GOLDBERG, *What every computer scientist should know about floating-point arithmetic*, *ACM Computing Surveys (CSUR)*, 23 (1991), pp. 5–48.
- [22] D. GRUNTZ, *On Computing Limits in a Symbolic Manipulation System*, PhD thesis, Swiss Federal Institute of Technology, Zürich, Switzerland, 1996.
- [23] D. GRUNTZ AND W. KOEPF, *Formal power series*, (1993).
- [24] C. V. HORSSEN, *GMPY*. <https://pypi.python.org/pypi/gmpy2>, 2015.
- [25] T. R. KANE AND D. A. LEVINSON, *Dynamics, Theory and Applications*, McGraw Hill, 1985.
- [26] J. LAGRANGE, *Mécanique analytique*, no. v. 1 in *Mécanique analytique*, Ve Courcier, 1811.
- [27] L. R. U. MANSSUR, R. PORTUGAL, AND B. F. SVAITER, *Group-theoretic approach for symbolic tensor manipulation*, *Int. J. Mod. Phys. C*, 13 (2002), <http://dx.doi.org/http://dx.doi.org/10.1142/S0129183102004571>.
- [28] J. MARTÍN-GARCÍA, *xact, efficient tensor computer algebra*, 2002-2016, <http://metric.iem.csic.es/Martin-Garcia/xAct/>.
- [29] M. MOSKEWICZ, C. MADIGAN, AND S. MALIK, *Method and system for efficient implementa-*



- tion of boolean satisfiability, Aug. 26 2008, <http://www.google.co.in/patents/US7418369>.  
US Patent 7,418,369.
- [30] A. NIJENHUIS AND H. S. WILF, *Combinatorial Algorithms: For Computers and Calculators*, Academic Press, New York, NY, USA, second ed., 1978.
  - [31] T. E. OLIPHANT, *Python for scientific computing*, Computing in Science & Engineering, 9 (2007), pp. 10–20.
  - [32] K. PEETERS, *Cadabra: a field-theory motivated symbolic computer algebra system*, Computer Physics Communications, (2007).
  - [33] F. PÉREZ AND B. E. GRANGER, *Ipython: a system for interactive scientific computing*, Computing in Science & Engineering, 9 (2007), pp. 21–29.
  - [34] D. L. PETERSON, G. GEDE, AND M. HUBBARD, *Symbolic linearization of equations of motion of constrained multibody systems*, Multibody System Dynamics, 33 (2014), pp. 143–161, <http://dx.doi.org/10.1007/s11044-014-9436-5>.
  - [35] M. PETKOVŠEK, H. S. WILF, AND D. ZEILBERGER, *A = bak peters*, Wellesley, MA, (1996).
  - [36] E. RAYMOND, *The cathedral and the bazaar*, Knowledge, Technology & Policy, 12 (1999), pp. 23–49.
  - [37] M. SOFRONIOU AND G. SPALETTA, *Precise numerical computation*, Journal of Logic and Algebraic Programming, 64 (2005), pp. 113–134.
  - [38] P. SOLIN, K. SEGETH, AND I. DOLEZEL, *Higher-Order Finite Element Methods*, Chapman & Hall / CRC Press, 2003.
  - [39] W. STEIN AND D. JOYNER, *SAGE: System for Algebra and Geometry Experimentation*, 2005.
  - [40] H. TAKAHASHI AND M. MORI, *Double exponential formulas for numerical integration*, Publications of the Research Institute for Mathematical Sciences, 9 (1974), pp. 721–741.
  - [41] V. T. TOTH, *Maple and meijer’s g-function: a numerical instability and a cure*. <http://www.vttoth.com/CMS/index.php/technical-notes/67>, 2007.
  - [42] M. J. TURK, B. D. SMITH, J. S. OISHI, S. SKORY, S. W. SKILLMAN, T. ABEL, AND M. L. NORMAN, *yt: A Multi-code Analysis Toolkit for Astrophysical Simulation Data*, The Astrophysical Journal Supplement Series, 192 (2011), pp. 9–+, <http://dx.doi.org/10.1088/0067-0049/192/1/9>, arXiv:1011.3514.
  - [43] O. ZIENKIEWICZ, R. TAYLOR, AND J. ZHU, *The Finite Element Method: Its Basis and Fundamentals*, Butterworth-Heinemann, seventh edition ed., 2013, <http://dx.doi.org/http://dx.doi.org/10.1016/B978-1-85617-633-0.00019-8>.

## 8. Supplement.

**8.1. Limits: The Gruntz Algorithm.** SymPy calculates limits using the Gruntz algorithm, as described in [22]. The basic idea is as follows: any limit can be converted to a limit  $\lim_{x \rightarrow \infty} f(x)$  by substitutions like  $x \rightarrow \frac{1}{x}$ . Then the most varying subexpression  $\omega$  (that converges to zero as  $x \rightarrow \infty$  the fastest from all subexpressions) is identified in  $f(x)$ , and  $f(x)$  is expanded into a series with respect to  $\omega$ . Any positive powers of  $\omega$  converge to zero. If there are negative powers of  $\omega$ , then the limit is infinite. The constant term (independent of  $\omega$ , but could depend on  $x$ ) then determines the limit (one might need to recursively apply the Gruntz algorithm on this term to determine the limit).

To determine the most varying subexpression, the comparability classes must first be defined, by calculating  $L$ :

$$(1) \quad L \equiv \lim_{x \rightarrow \infty} \frac{\log |f(x)|}{\log |g(x)|}$$

And then operations  $<$ ,  $>$  and  $\sim$  are defined as follows:  $f > g$  when  $L = \pm\infty$  (it is said that  $f$  is more rapidly varying than  $g$ , i.e.,  $f$  goes to  $\infty$  or 0 faster than  $g$ ,  $f$  is greater than any power of  $g$ ),  $f < g$  when  $L = 0$  ( $f$  is less rapidly varying than  $g$ ) and  $f \sim g$  when  $L \neq 0, \pm\infty$  (both  $f$  and  $g$  are bounded from above and below by suitable integral powers of the other). Here are some examples of comparability classes:

$$2 < x < e^x < e^{x^2} < e^{e^x}$$

$$2 \sim 3 \sim -5$$

$$x \sim x^2 \sim x^3 \sim \frac{1}{x} \sim x^m \sim -x$$

$$e^x \sim e^{-x} \sim e^{2x} \sim e^{x+e^{-x}}$$

$$f(x) \sim \frac{1}{f(x)}$$

715 The Gruntz algorithm is now illustrated on the following example:

716 (2) 
$$f(x) = e^{x+2e^{-x}} - e^x + \frac{1}{x}.$$

717 The goal is to calculate  $\lim_{x \rightarrow \infty} f(x)$ . First the set of most rapidly varying subexpressions  
 718 is determined, the so called *mrv set*. For (2), the following mrv set  $\{e^x, e^{-x}, e^{x+2e^{-x}}\}$   
 719 is obtained. These are all subexpressions of (2) and they all belong to the same  
 720 comparability class. This calculation can be done using SymPy as follows:

```
721 >>> from sympy.series.gruntz import mrv
722 >>> mrv(exp(x+2*exp(-x))-exp(x) + 1/x, x)[0].keys()
723 dict_keys([exp(x + 2*exp(-x)), exp(x), exp(-x)])
```

724 Next any item  $\omega$  is taken from mrv that converges to zero for  $x \rightarrow \infty$ . The item  
 725  $\omega = e^{-x}$  is obtained. If such a term is not present in the mrv set (i.e., all terms  
 726 converge to infinity instead of zero), the relation  $f(x) \sim \frac{1}{f(x)}$  can be used.

727 Next step is to rewrite the mrv in terms of  $\omega$ :  $\{\frac{1}{\omega}, \omega, \frac{1}{\omega}e^{2\omega}\}$ . Then the original  
 728 subexpressions are substituted back into  $f(x)$  and expanded with respect to  $\omega$ :

729 (3) 
$$f(x) = \frac{1}{x} - \frac{1}{\omega} + \frac{1}{\omega}e^{2\omega} = 2 + \frac{1}{x} + 2\omega + O(\omega^2)$$

730 Since  $\omega$  is from the mrv set, then in the limit  $x \rightarrow \infty$  it is  $\omega \rightarrow 0$  and so  
 731  $2\omega + O(\omega^2) \rightarrow 0$  in (3):

732 (4) 
$$f(x) = \frac{1}{x} - \frac{1}{\omega} + \frac{1}{\omega}e^{2\omega} = 2 + \frac{1}{x} + 2\omega + O(\omega^2) \rightarrow 2 + \frac{1}{x}$$

733 Since the result  $(2 + \frac{1}{x})$  still depends on  $x$ , the above procedure is iterated on the  
 734 result until just a number (independent of  $x$ ) is obtained, which is the final limit. In  
 735 the above case the limit is 2, as can be verified by SymPy:

```
736 >>> limit(exp(x+2*exp(-x))-exp(x) + 1/x, x, oo)
737 2
```

738 In general, when  $f(x)$  is expanded in terms of  $\omega$ , it is obtained:

739 (5) 
$$f(x) = \underbrace{O\left(\frac{1}{\omega^3}\right)}_{\infty} + \underbrace{\frac{C_{-2}(x)}{\omega^2}}_{\infty} + \underbrace{\frac{C_{-1}(x)}{\omega}}_{\infty} + C_0(x) + \underbrace{C_1(x)\omega}_0 + \underbrace{O(\omega^2)}_0$$

740 The positive powers of  $\omega$  are zero. If there are any negative powers of  $\omega$ , then the  
 741 result of the limit is infinity, otherwise the limit is equal to  $\lim_{x \rightarrow \infty} C_0(x)$ . The expression  
 742  $C_0(x)$  is simpler than  $f(x)$  and so the algorithm always converges. A proof of this, as  
 743 well as further details are given in Gruntz's Ph.D. thesis [22].

## 744 8.2. Series.

**8.2.1. Series Expansion.** SymPy is able to calculate the symbolic series expansion of an arbitrary series or expression involving elementary and special functions and multiple variables. For this it has two different implementations- the `series` method and Ring Series.

The first approach stores a series as an object of the `Basic` class. Each function has its specific implementation of its expansion which is able to evaluate the Puiseux series expansion about a specified point. For example, consider a Taylor expansion about 0:

```
>>> from sympy import symbols, series
>>> x, y = symbols('x, y')
>>> series(sin(x+y) + cos(x*y), x, 0, 2)
1 + sin(y) + x*cos(y) + O(x**2)
```

The newer and much faster[1] approach called Ring Series makes use of the observation that a truncated Taylor series, is in fact a polynomial. Ring Series uses the efficient representation and operations of sparse polynomials. The choice of sparse polynomials is deliberate as it performs well in a wider range of cases than a dense representation. Ring Series gives the user the freedom to choose the type of coefficients he wants to have in his series, allowing the use of faster operations on certain types.

For this, several low level methods for expansion of trigonometric, hyperbolic and other elementary functions like inverse of a series, calculating  $n$ th root, etc, are implemented using variants of the Newton[12] Method. All these support Puiseux series expansion. The following example demonstrates the use of an elementary function that calculates the Taylor expansion of the sine of a series.

```
>>> from sympy import ring
>>> from sympy.polys.ring_series import rs_sin
>>> R, x = ring('x', QQ)
>>> rs_sin(x**2 + x, x, 5)
-1/2*x**4 - 1/6*x**3 + x**2 + x
```

The function `sympy.polys.rs_series` makes use of these elementary functions to expand an arbitrary SymPy expression. It does so by following a recursive strategy of expanding the lower most functions first and then composing them recursively to calculate the desired expansion. Currently it only supports expansion about 0 and is under active development. Ring Series is several times faster than the default implementation with the speed difference increasing with the size of the series. The `sympy.polys.rs_series` takes as input any SymPy expression and hence there is no need to explicitly create a polynomial ring. An example:

```
>>> from sympy.polys.ring_series import rs_series
>>> from sympy.abc import a, b
>>> from sympy import sin, cos
>>> rs_series(sin(a + b), a, 4)
-1/2*(sin(b))*a**2 + (sin(b)) - 1/6*(cos(b))*a**3 + (cos(b))*a
```

**8.2.2. Formal Power Series.** SymPy can be used for computing the Formal Power Series of a function. The implementation is based on the algorithm described in the paper on Formal Power Series[23]. The advantage of this approach is that an explicit formula for the coefficients of the series expansion is generated rather than just computing a few terms.

The following example shows how to use `fps`:

```
>>> f = fps(sin(x), x, x0=0)
```

```

794 >>> f.truncate(6)
795 x - x**3/6 + x**5/120 + O(x**6)
796 >>> f[15]
797 -x**15/1307674368000

```

8.2.3. **Fourier Series.** SymPy provides functionality to compute Fourier Series of a function using the `fourier_series` function. Under the hood it just computes  $a_0$ ,  $a_n$ ,  $b_n$  using standard integration formulas.

Here's an example on how to compute Fourier Series in SymPy:

```

802 >>> L = symbols('L')
803 >>> f = fourier_series(2 * (Heaviside(x/L) - Heaviside(x/L - 1)) - 1, (x, 0, 2*L))
804 >>> f.truncate(3)
805 4*sin(pi*x/L)/pi + 4*sin(3*pi*x/L)/(3*pi) + 4*sin(5*pi*x/L)/(5*pi)

```

8.3. **Logic.** SymPy supports construction and manipulation of boolean expressions through the `logic` module. SymPy symbols can be used as propositional variables and also be substituted as `True` or `False`. A good number of manipulation features for boolean expressions have been implemented in the `logic` module.

8.3.1. **Constructing boolean expressions.** A boolean variable can be declared as a SymPy symbol. Python operators `&`, `|` and `~` are overloaded for logical `And`, `Or` and `negate`. Several others like `Xor`, `Implies` can be constructed with `^`, `»` respectively. The above are just a shorthand, expressions can also be constructed by directly calling `And()`, `Or()`, `Not()`, `Xor()`, `Nand()`, `Nor()`, etc.

```

815 >>> from sympy import *
816 >>> x, y, z = symbols('x y z')
817 >>> e = (x & y) | z
818 >>> e.subs({x: True, y: True, z: False})
819 True

```

8.3.2. **CNF and DNF.** Any boolean expression can be converted to conjunctive normal form, disjunctive normal form and negation normal form. The API also permits to check if a boolean expression is in any of the above mentioned forms.

```

823 >>> from sympy import *
824 >>> x, y, z = symbols('x y z')
825 >>> to_cnf((x & y) | z)
826 And(Or(x, z), Or(y, z))
827 >>> to_dnf(x & (y | z))
828 Or(And(x, y), And(x, z))
829 >>> is_cnf((x | y) & z)
830 True
831 >>> is_dnf((x & y) | z)
832 True

```

8.3.3. **Simplification and Equivalence.** The module supports simplification of given boolean expression by making deductions on it. Equivalence of two expressions can also be checked. If so, it is possible to return the mapping of variables of two expressions so as to represent the same logical behaviour.

```

837 >>> from sympy import *
838 >>> a, b, c, x, y, z = symbols('a b c x y z')
839 >>> e = a & (~a | ~b) & (a | c)
840 >>> simplify(e)

```

```

841 And(Not(b), a)
842 >>> e1 = a & (b | c)
843 >>> e2 = (x & y) | (x & z)
844 >>> bool_map(e1, e2)
845 (And(Or(b, c), a), {b: y, a: x, c: z})

```

846 **8.3.4. SAT solving.** The module also supports satisfiability checking of a given  
847 boolean expression. If satisfiable, it is possible to return a model for which the ex-  
848 pression is satisfiable. The API also supports returning all possible models. The SAT  
849 solver has a clause learning DPLL algorithm implemented with watch literal scheme  
850 and VSIDS heuristic[29].

```

851 >>> from sympy import *
852 >>> a, b, c = symbols('a b c')
853 >>> satisfiable(a & (~a | b) & (~b | c) & ~c)
854 False
855 >>> satisfiable(a & (~a | b) & (~b | c) & c)
856 {b: True, a: True, c: True}

```

857 **8.4. Diophantine Equations.** Diophantine equations play a central and an im-  
858 portant role in number theory. A Diophantine equation has the form,  $f(x_1, x_2, \dots, x_n) =$   
859  $0$  where  $n \geq 2$  and  $x_1, x_2, \dots, x_n$  are integer variables. If we can find  $n$  integers  
860  $a_1, a_2, \dots, a_n$  such that  $x_1 = a_1, x_2 = a_2, \dots, x_n = a_n$  satisfies the above equation, we  
861 say that the equation is solvable.

862 Currently, following five types of Diophantine equations can be solved using  
863 SymPy's Diophantine module.

- 864 • Linear Diophantine equations:  $a_1x_1 + a_2x_2 + \dots + a_nx_n = b$
- 865 • General binary quadratic equation:  $ax^2 + bxy + cy^2 + dx + ey + f = 0$
- 866 • Homogeneous ternary quadratic equation:  $ax^2 + by^2 + cz^2 + dxy + eyz + fzx = 0$
- 867 • Extended Pythagorean equation:  $a_1x_1^2 + a_2x_2^2 + \dots + a_nx_n^2 = a_{n+1}x_{n+1}^2$
- 868 • General sum of squares:  $x_1^2 + x_2^2 + \dots + x_n^2 = k$

869 When an equation is fed into Diophantine module, it factors the equation (if  
870 possible) and solves each factor separately. Then all the results are combined to create  
871 the final solution set. Following examples illustrate some of the basic functionalities  
872 of the Diophantine module.

```

873 >>> from sympy import symbols
874 >>> x, y, z = symbols("x, y, z", integer=True)
875
876 >>> diophantine(2*x + 3*y - 5)
877 set([(3*t_0 - 5, -2*t_0 + 5)])
878
879 >>> diophantine(2*x + 4*y - 3)
880 set()
881
882 >>> diophantine(x**2 - 4*x*y + 8*y**2 - 3*x + 7*y - 5)
883 set([(2, 1), (5, 1)])
884
885 >>> diophantine(x**2 - 4*x*y + 4*y**2 - 3*x + 7*y - 5)
886 set([(-2*t**2 - 7*t + 10, -t**2 - 3*t + 5)])
887
888 >>> diophantine(3*x**2 + 4*y**2 - 5*z**2 + 4*x*y - 7*y*z + 7*z*x)
889 set([(-16*p**2 + 28*p*q + 20*q**2, 3*p**2 + 38*p*q - 25*q**2, 4*p**2 - 24*p*q + 68*q**2)])

```

890

891 >>> from sympy.abc import a, b, c, d, e, f

892 >>> diophantine(9\*a\*\*2 + 16\*b\*\*2 + c\*\*2 + 49\*d\*\*2 + 4\*e\*\*2 - 25\*f\*\*2)

893 set([(70\*t1\*\*2 + 70\*t2\*\*2 + 70\*t3\*\*2 + 70\*t4\*\*2 - 70\*t5\*\*2, 105\*t1\*t5, 420\*t2\*t5, 60\*t3\*t5, 210\*t4\*t5, 4

894

895 >>> diophantine(a\*\*2 + b\*\*2 + c\*\*2 + d\*\*2 + e\*\*2 + f\*\*2 - 112)

896 set([(8, 4, 4, 4, 0, 0)])

897 **8.5. Sets.** SymPy supports representation of a wide variety of mathematical  
 898 sets. This is achieved by first defining abstract representations of atomic set classes  
 899 and then combining and transforming them using various set operations.

900 Each of the set classes inherits from the base class `Set` and defines methods to  
 901 check membership and calculate unions, intersections, and set differences. When these  
 902 methods are not able to evaluate to atomic set classes, they are represented as abstract  
 903 unevaluated objects.

904 SymPy has the following atomic set classes:

- 905 • `EmptySet` represents the empty set  $\emptyset$ .
- 906 • `UniversalSet` is an abstract “universal set” for which everything is a member.  
 907 The union of the universal set with any set gives the universal set and the  
 908 intersection gives to the other set itself.
- 909 • `FiniteSet` is functionally equivalent to Python’s built `inset` object. Its mem-  
 910 bers can be any SymPy object including other sets themselves.
- 911 • `Integers` represents the set of Integers  $\mathbb{Z}$ .
- 912 • `Naturals` represents the set of Natural numbers  $\mathbb{N}$ , i.e., the set of positive  
 913 integers.
- 914 • `Naturals0` represents the whole numbers, which are all the non-negative in-  
 915 tegers.
- 916 • `Range` represents a range of integers. A range is defined by specifying a start  
 917 value, an end value, and a step size. Range is functionally equivalent to  
 918 Python’s `range` except it supports infinite endpoints, allowing the represen-  
 919 tation of infinite ranges.
- 920 • `Interval` represents an interval of real numbers. It is specified by giving the  
 921 start and end point and specifying if it is open or closed in the respective  
 922 ends.

923 Other than unevaluated classes of Union, Intersection and Set Difference opera-  
 924 tions, we have following set classes.

- 925 • `ProductSet` defines the Cartesian product of two or more sets. The product  
 926 set is useful when representing higher dimensional spaces. For example to  
 927 represent a three-dimensional space we simply take the Cartesian product of  
 928 three real sets.
- 929 • `ImageSet` represents the image of a function when applied to a particular  
 930 set. In notation, the image set of a function  $F$  with respect to a set  $S$  is  
 931  $\{F(x)|x \in S\}$ . SymPy uses image sets to represent sets of infinite solutions  
 932 equations such as  $\sin(x) = 0$ .
- 933 • `ConditionSet` represents subset of a set whose members satisfies a particular  
 934 condition. In notation, the condition set of the set  $S$  with respect to the  
 935 condition  $H$  is  $\{x|H(x), x \in S\}$ . SymPy uses condition sets to represent  
 936 the set of solutions of equations and inequalities, where the equation or the  
 937 inequality is the condition and the set is the domain being solved over.

938 A few other classes are implemented as special cases of the classes described

above. The set of real numbers, `Reals` is implemented as a special case of `Interval`,  $(-\infty, \infty)$ . `ComplexRegion` is implemented as a special case of `ImageSet`. `ComplexRegion` supports both polar and rectangular representation of regions on the complex plane.

**8.6. SymPy Gamma.** SymPy Gamma is a simple web application that runs on Google App Engine. It executes and displays the results of SymPy expressions as well as additional related computations, in a fashion similar to that of Wolfram|Alpha. For instance, entering an integer will display its prime factors, digits in the base-10 expansion, and a factorization diagram. Entering a function will display its docstring; in general, entering an arbitrary expression will display its derivative, integral, series expansion, plot, and roots.

SymPy Gamma also has several additional features than just computing the results using SymPy.

- It displays integration steps, differentiation steps in detail, which can be viewed in Figure 1:

Integral Steps:

integrate(tan(x), x)

Fullscreen

1. Rewrite the integrand:

$$\tan(x) = \frac{\sin(x)}{\cos(x)}$$

2. Let  $u = \cos(x)$ .

Then let  $du = -\sin(x)dx$  and substitute  $du$ :

$$\int -\frac{1}{u} du$$

A. The integral of a constant times a function is the constant times the integral of the function:

$$\int -\frac{1}{u} du = - \int \frac{1}{u} du$$

I. The integral of  $\frac{1}{u}$  is  $\log(u)$ .

So, the result is:  $-\log(u)$

Now substitute  $u$  back in:

$$-\log(\cos(x))$$

3. Add the constant of integration:

$$-\log(\cos(x)) + \text{constant}$$

The answer is:

$$-\log(\cos(x)) + \text{constant}$$

Fig. 1: Integral steps of  $\tan(x)$

- It also displays the factor tree diagrams for different numbers.
  - SymPy Gamma also saves user search queries, and offers many such similar features for free, which Wolfram|Alpha only offers to its paid users.
- Every input query from the user on SymPy Gamma is first, parsed by its own parser,

which handles several different forms of function names, which SymPy as a library doesn't support. For instance, SymPy Gamma supports queries like `sin x`, whereas SymPy doesn't support this, and supports only `sin(x)`.

This parser converts the input query to the equivalent SymPy readable code, which is then eventually processed by SymPy and the result is finally formatted in LaTeX and displayed on the SymPy Gamma web-application.

**8.7. SymPy Live.** SymPy Live is an online Python shell, which runs on Google App Engine, that executes SymPy code. It is integrated in the SymPy documentation examples, located at this [link](#).

This is accomplished by providing a HTML/JavaScript GUI for entering source code and visualization of output, and a server part which evaluates the requested source code. It's an interactive AJAX shell, that runs SymPy code using Python on the server.

Certain Features of SymPy Live:

- It supports the exact same syntax as SymPy, hence it can be used easily, to test for outputs of various SymPy expressions.
- It can be run as a standalone app or in an existing app as an admin-only handler, and can also be used for system administration tasks, as an interactive way to try out APIs, or as a debugging aid during development.
- It can also be used to plot figures ([link](#)), and execute all kinds of expressions that SymPy can evaluate.
- SymPy Live also formats the output in LaTeX for pretty-printing the output.

**8.8. Comparison with Mathematica.** Wolfram Mathematica is a popular proprietary CAS. It features highly advanced algorithms. Mathematica has a core implemented in C++ [6] which interprets its own programming language (known as Wolfram language).

Analogously to Lisp's S-expressions, Mathematica uses its own style of M-expressions, which are arrays of either atoms or other M-expression. The first element of the expression identifies the type of the expression and is indexed by zero, whereas the first argument is indexed by one. Notice that SymPy expression arguments are stored in a Python tuple (that is, an immutable array), while the expression type is identified by the type of the object storing the expression.

Mathematica can associate attributes to its atoms. Attributes may define mathematical properties and behavior of the nodes associated to the atom. In SymPy, the usage of static class fields is roughly similar to Mathematica's attributes, though other programming patterns may also be used to achieve an equivalent behavior, such as class inheritance.

Unlike SymPy, Mathematica's expressions are mutable, that is one can change parts of the expression tree without the need of creating a new object. The reactivity of Mathematica allows for a lazy updating of any references to that data structure.

Products in Mathematica are determined by some builtin node types, such as `Times`, `Dot`, and others. `Times` is overloaded by the `*` operator, and is always meant to represent a commutative operator. The other notable product is `Dot`, overloaded by the `.` operator. This product represents matrix multiplication, it is not commutative. SymPy uses the same node for both scalar and matrix multiplication, the only exception being with abstract matrix symbols. Unlike Mathematica, SymPy determines commutativity with respect to multiplication from the factor's expression type. Mathematica puts the `Orderless` attribute on the expression type.

Regarding associative expressions, SymPy handles associativity by making asso-



ciative expressions inherit the class `AssocOp`, while Mathematica specifies the `Flat` attribute on the expression type.

Mathematica relies heavily on pattern matching: even the so-called equivalent of function declaration is in reality the definition of a pattern matching generating an expression tree transformation on input expressions. Mathematica's pattern matching is sensitive to associative, commutative, and one-identity properties of its expression tree nodes. SymPy has various ways to perform pattern matching. All of them play a lesser role in the CAS than in Mathematica and are basically available as a tool to rewrite expressions. The differential equation solver in SymPy somewhat relies on pattern matching to identify the kind of differential equation, but it is envisaged to replace that strategy with analysis of Lie symmetries in the future. Mathematica's real advantage is the ability to add new overloading to the expression builder at runtime, or for specific subnodes. Consider for example

```
In[1]:= Unprotect[Plus]
```

```
Out[1]= {Plus}
```

```
In[2]:= Sin[x_]^2 + Cos[y_]^2 := 1
```

```
In[3]:= x + Sin[t]^2 + y + Cos[t]^2
```

```
Out[3]= 1 + x + y
```

This expression in Mathematica defines a substitution rule that overloads the functionality of the `Plus` node (the node for additions in Mathematica). The trailing underscore after a symbol means that it is to be considered a wildcard. This example may not be practical, one may wish to keep this identity unevaluated, nevertheless it clearly illustrates the potentiality to define one's own immediate transformation rules. In SymPy the operations constructing the addition node in the expression tree are Python class constructors, and cannot be modified at runtime<sup>5</sup>. The way SymPy deals with extending the missing runtime overloadability functionality is by subclassing the node types. Subclasses may overload the class constructor to yield the proper extended functionality.

Unlike SymPy, Mathematica does not support type inheritance or polymorphism. SymPy relies heavily on class inheritance, but for the most part, class inheritance is used to make sure that SymPy objects inherit the proper methods and implement the basic hashing system. Associativity of expressions can be achieved by inheriting the class `AssocOp`, which may appear a more cumbersome operation than Mathematica's attribute setting.

Matrices in SymPy are types on their own. In Mathematica, nested lists are interpreted as matrices whenever the sublists have the same length. The main difference to SymPy is that ordinary operators and functions do not get generalized the same way as used in traditional mathematics. Using the standard multiplication in Mathematica performs an elementwise product, this is compatible with Mathematica's convention of commutativity of `Times` nodes. Matrix product is expressed by the `dot` operator, or the `Dot` node. The same is true for the other operators, and even functions, most notably calling the exponential function `Exp` on a matrix returns an elementwise exponentiation of its elements. The real matrix exponentiation is

---

<sup>5</sup>In reality, Python supports monkey patching, nonetheless it is a discouraged programming pattern.

available through the `MatrixExp` function.

Unevaluated expressions can be achieved in various ways, most commonly with the `HoldForm` or `Hold` nodes, that block the evaluation of subnodes by the parser. Note that such a node cannot be expressed in Python, because of greedy evaluation. Whenever needed in SymPy, it is necessary to add the parameter `evaluate=False` to all subnodes, or put the input expression in a string.

The operator `==` returns a boolean whenever it is able to immediately evaluate the truthness of the equality, otherwise it returns an `Equal` expression. In SymPy `==` means structural equality and is always guaranteed to return a boolean expression. To express an equality in SymPy it is necessary to explicitly construct the `Equality` class.

SymPy, in accordance with Python and unlike the usual programming convention, uses `**` to express the power operator, while Mathematica uses the more common `^`.

**8.9. Other Projects that use SymPy.** There are several projects that use SymPy as a library for implementing a part of their project, or even as a part of back-end for their application as well.

Some of them are listed below:

- **Cadabra**: Cadabra is a symbolic computer algebra system (CAS) designed specifically for the solution of problems encountered in field theory.
- **Octave Symbolic**: The Octave-Forge Symbolic package adds symbolic calculation features to GNU Octave. These include common Computer Algebra System tools such as algebraic operations, calculus, equation solving, Fourier and Laplace transforms, variable precision arithmetic and other features.
- **SymPy.jl**: Provides a Julia interface to SymPy using PyCall.
- **Mathics**: Mathics is a free, general-purpose online CAS featuring Mathematica compatible syntax and functions. It is backed by highly extensible Python code, relying on SymPy for most mathematical tasks.
- **Mathpix**: An iOS App, that uses Artificial Intelligence to detect handwritten math as input, and uses SymPy Gamma, to evaluate the math input and generate the relevant steps to solve the problem.
- **IKFast**: IKFast is a robot kinematics compiler provided by **OpenRAVE**. It analytically solves robot inverse kinematics equations and generates optimized C++ files. It uses SymPy for its internal symbolic mathematics.
- **Sage**: A CAS, visioned to be a viable free open source alternative to Magma, Maple, Mathematica and Matlab.
- **SageMathCloud**: SageMathCloud is a web-based cloud computing and course management platform for computational mathematics.
- **PyDy**: Multibody Dynamics with Python.
- **galgebra**: Geometric algebra (previously `sympy.galgebra`).
- **yt**: Python package for analyzing and visualizing volumetric data (`yt.units` uses SymPy).
- **SfePy**: Simple finite elements in Python, see Section 8.10.1.
- **Quameon**: Quantum Monte Carlo in Python.
- **Lcapy**: Experimental Python package for teaching linear circuit analysis.
- **Quantum Programming in Python**: Quantum 1D Simple Harmonic Oscillator and Quantum Mapping Gate.
- **LaTeX Expression project**: Easy LaTeX typesetting of algebraic expressions in symbolic form with automatic substitution and result computation.

- **Symbolic statistical modeling:** Adding statistical operations to complex physical models.

**8.10. Project Details.** Below we provide particular examples of SymPy use in some of the projects listed above.

**8.10.1. SfePy.** **SfePy** (Simple finite elements in Python), cf. [16], is a Python package for solving partial differential equations (PDEs) in 1D, 2D and 3D by the finite element (FE) method [43]. SymPy is used within this package mostly for code generation and testing, namely:

- generation of the hierarchical FE basis module, involving generation and symbolic differentiation of 1D Legendre and Lobatto polynomials, constructing the FE basis polynomials [38] and generating the C code;
- generation of symbolic conversion formulas for various groups of elastic constants [20] – provide any two of the Young’s modulus, Poisson’s ratio, bulk modulus, Lamé’s first parameter, shear modulus (Lamé’s second parameter) or longitudinal wave modulus and get the other ones;
- simple physical unit conversions, generation of consistent unit sets;
- testing FE solutions using method of manufactured (analytical) solutions – the differential operator of a PDE is symbolically applied and a symbolic right-hand side is created, evaluated in quadrature points, and subsequently used to obtain a numerical solution that is then compared to the analytical one;
- testing accuracy of 1D, 2D and 3D numerical quadrature formulas (cf. [7]) by generating polynomials of suitable orders, integrating them, and comparing the results with those obtained by the numerical quadrature.

**8.11. Tensors.** Ongoing work to provide the capabilities of tensor computer algebra has so far produced the `tensor` module. It is composed of three separated sub-modules, whose purposes are quite different: `tensor.indexed` and `tensor.indexed_methods` support indexed symbols, `tensor.array` contains facilities to operator on symbolic  $N$ -dimensional arrays and finally `tensor.tensor` is used to define abstract tensors. The abstract tensors subsection is inspired by xAct[28] and Cadabra[32]. Canonicalization based on the Butler-Portugal[27] algorithm is supported in SymPy. It is currently limited to polynomial tensor expressions.