DOI 10.1007/s00165-016-0367-1
BCS © 2016
Formal Aspects of Computing (2016) 28: 881–907

Formal Aspects of Computing

On the formal analysis of Gaussian optical systems in HOL

Umair Siddique[1] and Sofiène Tahar[1]

[1]Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada

Abstract. Optics technology is being increasingly used in mainstream industrial and research domains such as terrestrial telescopes, biomedical imaging and optical communication. One of the most widely used modeling approaches for such systems is Gaussian optics, which describes light as a beam. In this paper, we propose to use higher-order-logic theorem proving for the analysis of Gaussian optical systems. In particular, we present the formalization of Gaussian beams and verify the corresponding properties such as beam transformation, beam waist radius and location. Consequently, we build formal reasoning support for the analysis of quasi-optical systems. In order to demonstrate the effectiveness of our approach, we present a case study about the receiver module of a real-world Atacama Pathfinder Experiment (APEX) telescope.

Keywords: Geometrical optics, Gaussian beams, Quasi-optical systems, Theorem proving, HOL light

1. Introduction

In recent decades, optics technology has found applications in a variety of critical domains such as space missions, laser surgeries, remote sensing and high-speed computing. The designing of different optical systems is highly dependent on the modeling choices for the light and optical components (e.g., mirrors and lenses). In fact, light can be modeled at different levels of abstraction such as ray, beam, electromagnetic and quantum optics.Ray optics [ST07] characterizes light as a straight line which linearly traverses in the optical system. In Gaussian optics [Wel91], light is considered as Gaussian beams which have a small spread around the axis of propagation. Electromagnetic optics [Trä07] describes the vectorial wave nature of light. On the other hand, quantum optics [Gri05] characterizes light as a stream of photons and helps to tackle situations where it is necessary to consider both the wave-like and particle-like behavior of light. In general, each of these theories has been used to model different aspects of the same or different optical components. For example, a phase-conjugate mirror [HW05a] can be modeled using the ray, electromagnetic and quantum optics.

Correspondence and offprint requests to: U. Siddique, E-mail: muh_sidd@ece.concordia.ca

The application of each theory is dependent on the type of system properties which need to be verified. For instance, ray optics provides a convenient way to verify the stability of optical resonators [HW05b]. Gaussian beam transformation is widely used to evaluate the coupling efficiency of optical fibers [Dam05] and the mode analysis of optical resonators [HW05a]. On the other hand, ensuring that no energy is lost when light travels through a waveguide and the modeling of so-called active elements require electromagnetic and quantum optics [ST07], respectively. In this paper, we consider the modeling and analysis of optical systems using Gaussian optics which has widespread applications such as laser resonators, telescopes and predictions for the design parameters of physical experiments, e.g., a soliton transmission [Moo01].

Due to the delays and costs associated with the manufacturing process of optical systems, it is not practical to analyze the impact of design parameters on the system behavior by the successive fabrication and characterization of prototypes only. The comprehensive characterization of optical systems is also a time-consuming process and does not unveil all the internal behaviors of the design under test, since all properties cannot be directly measured. Therefore, detailed mathematical models and exhaustive analysis are required to develop an understanding of the underlying operations and the dependence on the parameters of an optical system. One natural step is to identify some fundamental building-blocks (e.g., mirrors or lenses) used in practical optical systems and then develop universal models characterizing the associated behavior to process light. Consequently, a significant portion of time is spent in the analysis and verification of these models so that bugs in the design can be detected prior to the manufacturing of the actual system. Even minor bugs in optical systems can lead to disastrous consequences such as the loss of human life because of their use in biomedical devices (e.g., refractive index measurement of cancer cells [SZL+06]), or financial loss because of their use in high budget space missions.

In order to build an accurate and reliable optical system analysis framework, a project[1] [ASM+14] was initiated in our research group at Concordia University in 2009. The main scope of this project includes the higher-order logic (HOL) [Har09a] formalization of different theories of optics which provides the basis to conduct more accurate analysis than traditional paper-and-pencil based proofs [KL66, NKST98, Moo01], numerical simulation [reZ15, SXS+11, LAS15] and computer algebra systems (CAS) [Opt15]. So far, the formalization of ray optics [SAT13b, SAT13a], electromagnetic optics [KAHT14] and quantum optics [MT14] has been implemented in the HOL light [Har09b] proof assistant. Recently, an extension of process calculus has been reported for the verification of linear optical quantum circuits [FAGP13]. However, the above-mentioned framework lacks the formalization of the basic building-blocks of Gaussian optics, which are based on the notion of Gaussian beams and their corresponding link to ray optics [Trä07].

The main focus of this paper is to bridge the above-mentioned gap and strengthen the formal reasoning support in the area of Gaussian optical systems. We build upon our preliminary work about the formalization of ray optics [SAT13b]. In particular, we formalize the notion of optical systems and Gaussian beams in higher-order logic. Building on this formalization, we develop a library of optical components (such as spherical mirror and thick lens) along with their corresponding behaviorial properties with respect to light beams. We also consider the notion of the composition of optical systems and show that all basic properties of a single optical system also hold for an arbitrary composed system. Consequently, we formalize widely used quasi-optical systems [Gol98] along with the formal verification of their generic properties about the Gaussian beams' transformation. In order to demonstrate the utilization of our work, we consider a case study about the verification of the system magnification of a receiver module of the Atacama Pathfinder Experiment (APEX) telescope[2] [NLD+09]. In this work, we use the HOL Light theorem prover [Har09b] due its rich multivariate analysis libraries [Har13]. Another motivation for using HOL light is to complement our work with other related developments of optics theories (i.e., ray, electromagnetic and quantum optics). The source code of our formalization is available for download [Sid15] and can be used by other researchers and optical engineers for further developments and the analysis of practical Gaussian optical systems.

The rest of the paper is organized as follows: Sect. 2 describes some fundamentals of ray optics and Gaussian beams. The proposed formal analysis framework is outlined in Sect. 3. We present the formalization of geometrical optics in Sect. 4 followed by the formalization of Gaussian beams in Sect. 5. We demonstrate the use of our work by formalizing generic quasi-optical systems and conducting the formal analysis of the receiver module of the APEX telescope in Sect. 6. Finally, Sect. 7 concludes the paper and provides some hints for future research directions.

---

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

[2] http://www.apex-telescope.org/

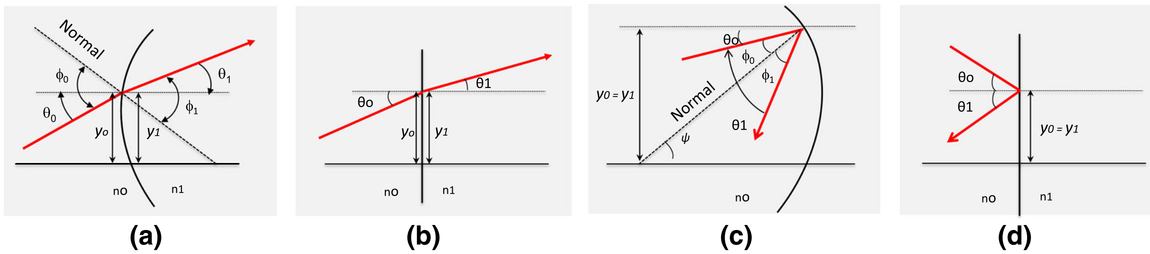

Fig. 1. Refraction and reflection of a ray

2. Preliminaries

In this section, we provide a brief introduction to geometrical optics and Gaussian beams. The intent is to introduce the basic theories along with some notations we use in the rest of the paper.

2.1. Geometrical optics

Ray optics or geometrical optics characterizes light as rays and is based on a set of postulates used to derive the rules for the propagation of light through an optical medium. These postulates are given as follows [ST07]:

- Light travels in the form of rays emitted by a source.
- An optical medium is characterized by its refractive index.
- Light rays follow Fermat's principle of least time.

Generally, the main components of an optical system are lenses, mirrors and a propagating medium which is either a free space or some material such as glass. These components are usually centered about an optical axis, around which rays travel at small inclinations (angle with the optical axis). Such rays are called *paraxial rays* and this assumption provides the basis of *paraxial optics* which is the simplest framework of geometrical optics. When a ray passes through optical components, it undergoes *translation* or *refraction*. In translation, the ray simply travels in a straight line from one component to the next and we only need to know the thickness of the translation. On the other hand, refraction takes place at the boundary of two regions with different refractive indices and the ray obeys the law of refraction, i.e., the angle of refraction relates to the angle of incidence by the relation $n_0\phi_0 = n_1\phi_1$, called *Paraxial Snell's law* [ST07], where n_0, n_1 are the refractive indices of both regions and ϕ_0, ϕ_1 are the angles of the incident and refracted rays, respectively, with the normal to the surface. The refraction and reflection of a single ray from plane and spherical interfaces are shown in Fig. 1.

The change in the position and inclination of a paraxial ray as it travels through an optical system can be conveniently described by the use of matrix algebra [KL66]. This matrix formalism (called *ray-transfer matrices*) of geometrical optics provides an efficient, scalable and systematic analysis of real-world complex optical and laser systems. This is because each optical component can be described as a 2 × 2 matrix and all linear algebraic properties can be used in the analysis of optical systems. For example, the general optical system with an input and output ray vector can be described as follows:

$$\begin{bmatrix} y_n \\ \theta_n \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} y_0 \\ \theta_0 \end{bmatrix}$$

where y_0 and θ_0 represent ray height and ray angle with optical axis, respectively. The parameters A, B, C and D are the components of the ray-transfer matrix which relates the input ray angle and ray height to those of the corresponding output.

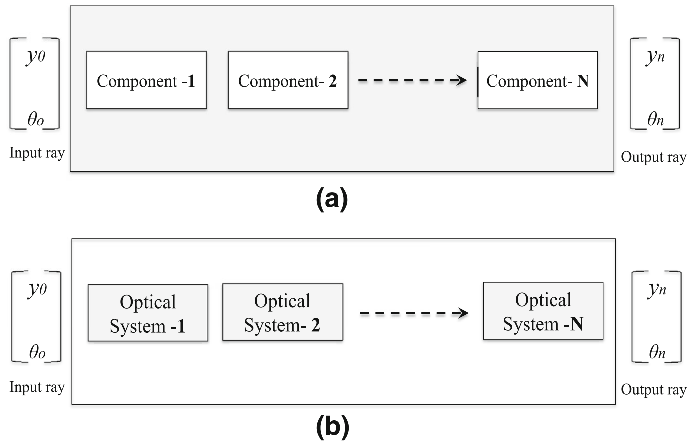

Fig. 2. Optical system and composed optical system

Finally, if we have an optical system consisting of N optical components, then we can trace the input ray R_0 through all optical components using the composition of matrices of each optical component as follows:

$$R_n = (M_k \cdot M_{k-1} \ldots M_1).R_0 \tag{1}$$

where M_i represents the ith component of the system. We can write $R_n = M_s R_0$, where $M_s = \Pi_{i=k}^{1} M_i$. Here, R_n is the output ray and R_0 is the input ray. Similarly, a composed optical system consists of N optical subsystems which inherits the same properties as of a single optical system, as shown in Fig. 2. This is a very useful modeling notion for the systems that consist of small subsystems, as we can use already available infrastructure with minimal effort.

Typical applications of ray-transfer matrices are the stability analysis of optical resonators [MPM$^+$11], optical pulse transmission [NKST98], and analysis of micro optoelectro-mechanical systems [WA05].

2.2. Gaussian optics

In Gaussian optics, light is abstracted as beams and its intensity follows the distribution of a Gaussian function [ST07]. Mathematically, a Gaussian beam is a solution of the paraxial wave equation in which wave front normals (i.e., the locus of points having the same phase) make very small angles with the axis of propagation. Figure 3 describes the wavefronts and wavefront normals for a paraxial wave. One of the most commonly used methods for constructing a paraxial wave is to consider a plane wave $A \exp(-jkz)$ (where $j = \sqrt{-1}$, k is the wave-number and z is the direction of propagation) and modify its complex amplitude A, by making it a slowly varying function of the position, i.e., $A(x, y, z)$. Mathematically, the complex amplitude of a paraxial wave becomes:

$$U(x, y, z) = A(x, y, z) \exp(-jkz) \tag{2}$$

For a paraxial wave to be valid in the context of geometrical optics, it should satisfy the paraxial Helmholtz equation [ST07], which is given as follows:

$$\nabla_T^2 A(x, y, z) - j2k \frac{\partial A(x, y, z)}{\partial z} = 0 \tag{3}$$

where $\nabla_T^2 = \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$ is the transverse Laplacian operator.

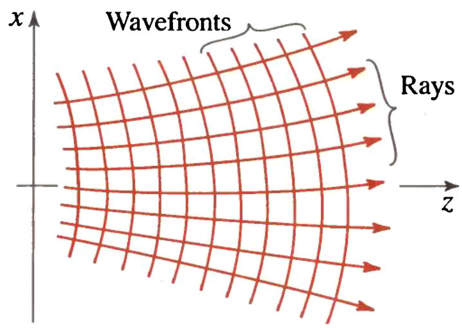

Fig. 3. The wavefronts and wavefront normal of a paraxial wave [ST07]

In general, different solutions can be found which satisfy Eq. (3). For example, a paraboloidal wave is a solution for which the complex envelope is given as below:

$$A(x, y, z) = \frac{A_0}{z} \exp\left[-jk\frac{x^2 + y^2}{2z}\right] \tag{4}$$

where $A_0 \in \mathbb{C}$ is a complex-valued constant.

Another solution of the Helmholtz equation is provided by the Gaussian beam [ST07] which is obtained from the paraboloidal wave by a simple transformation. Indeed, the complex envelope of the paraboloidal wave is a solution of the paraxial Helmholtz equation, and a shifted version is also a solution, i.e., replacing z with $z - \zeta$ in Eq. (4):

$$A(x, y, z) = \frac{A_0}{z - \zeta} \exp\left[-jk\frac{x^2 + y^2}{2(z - \zeta)}\right] \tag{5}$$

where $\zeta \in \mathbb{C}$ is a constant. Physically, it provides a paraboloidal wave centered about the point $z = \zeta$, rather than $z = 0$. The parameter ζ is very important and produces different properties depending upon the variation of its value, e.g., $\zeta = -jz_R$, provides the complex envelope of a Gaussian beam [ST07], which can be compactly described as follows:

$$A(x, y, z) = \frac{A_0}{q(z)} \exp\left[-jk\frac{x^2 + y^2}{2q(z)}\right] \tag{6}$$

where $q(z) = z + jz_R$ is called q-parameter of Gaussian beams. The parameter $z_R \in \mathbb{R}$ is known as the Rayleigh range.

In order to study the properties (e.g., phase and amplitude) of Gaussian beams, the above-mentioned complex valued q-parameter is expressed as $\frac{1}{q(z)} = \frac{1}{z + jz_R}$. In the optics literature, this expression is further transformed into a new form by defining two real-valued functions $R(z)$ and $W(z)$, as follows:

$$\frac{1}{q(z)} = \frac{1}{R(z)} - j\frac{\lambda}{\pi W^2(z)} \tag{7}$$

where $W(z)$ and $R(z)$ are measures of the beam width and wavefront radius of curvature, respectively (shown in Fig. 4). Mathematically, these parameters can be expressed as follows:

$$R(z) = z\left[1 + \left(\frac{z_R}{z}\right)^2\right] \tag{8}$$

$$W(z) = w_0\left[1 + \left(\frac{z_R}{z}\right)^2\right]^{\frac{1}{2}} \tag{9}$$

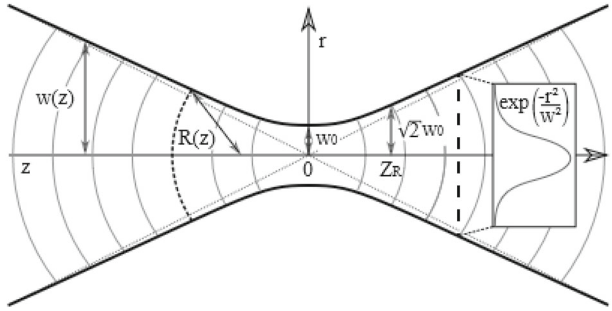

Fig. 4. Gaussian beam

The parameter $w_0 \in \mathbb{R}$ represents the value of the beam width at $z = 0$ which is also called beam waist size or beam waist radius. Finally, substituting Eqs. (7) in (6) and using Eq. (2), we obtain the following complex amplitude $U(x, y, z)$:

$$U(x, y, z) = A_0 \frac{w_0}{W(z)} \exp\left[-\frac{x^2 + y^2}{W^2(z)}\right] \exp\left[-jkz - jk\frac{x^2 + y^2}{2R(Z)} + j\xi(z)\right] \tag{10}$$

where $\xi(z) = \tan^{-1}(\frac{z}{z_R})$. The above equation is the main representation of Gaussian beams and describes the important properties of light when it travels from one component to another. For example, the optical intensity, $I(x, y, z) =\mid U(x, y, z) \mid^2$ can be expressed as follows:

$$I(x, y, z) = \left|\frac{A_0}{jz_R}\right|^2 \left(\frac{w_0}{W(z)}\right)^2 \exp\left[-\frac{k\frac{\lambda}{\pi}(x^2 + y^2)}{W^2(z)}\right] \tag{11}$$

Note that at each value of z, the intensity is a Gaussian function of the radial distance, which leads to the name Gaussian beams.

2.2.1. ABCD-law of beam transformation

We can completely characterize a Gaussian beam by its q-parameter ($q(z)$) (i.e., Eq. 7 [ST07]). This provides a convenient way to study the behavior of a Gaussian beam when it passes through an optical system. Indeed, it is sufficient to just consider the variations of the input q-parameter at each optical component. In paraxial geometrical optics, an optical system is completely characterized by the 2×2 ray-transfer matrix relating the position and inclination of the transmitted ray to the incident ray. Similarly, it is important to find out the effect of an arbitrary optical system (characterized by a matrix M of elements A, B, C, D) on the parameters of an input beam. This can be described by a well-known ABCD-law of Gaussian beam transformation [Trä07], given as follows:

$$q_0 = \frac{A.q_i + B}{C.q_i + D} \tag{12}$$

where q_i and q_o represent the input and output beam q-parameters, respectively. The elements A, B, C, and D correspond to the final ray transfer matrix of a geometrical optical system (which indeed represent the composition of the matrices of individual optical components, as shown in Fig. 5).

The main applications of beam transformation are in the analysis of laser cavities [ST07], telescopes [NLD+09] and the prediction of design parameters for physical experiments [Moo01].

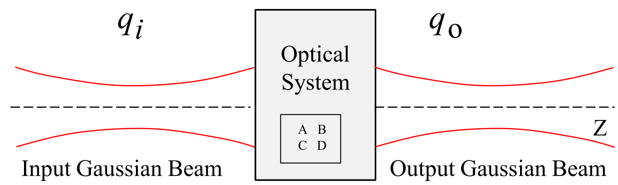

Fig. 5. Gaussian beam transformation

2.2.2. Quasi-optical systems

Quasi-optics [Gol98] deals with the propagation of a beam of radiation which is reasonably well collimated (i.e., rays are parallel and spread is minimal during the propagation, e.g., laser light) and the wavelength is relatively small along the axis of propagation. At a first glance, this looks like a restrictive notion of light but it has extraordinarily diverse applications ranging from compact systems in which all components are only a few wavelengths in size to antenna feed systems that illuminate an aperture of thousands or more wavelengths in diameter (e.g., space receiving stations) [Gol98]. It is important to note that geometrical optics deals with light rays with wavelength $\lambda \rightarrow 0$ and no diffraction effects. On the other hand, quasi-optics considers a wavelength which is approximately equal to the dimensions of the systems along with diffraction effects. In practice, quasi-optics is based on the Gaussian beams theory as it provides the necessary foundation to tackle the properties of the diffraction of light. Some of the successful applications of quasi-optics are in critical domains, e.g., military setups, radars, remote sensing, materials measurement systems [Gol98], radio frequency and radiometric optical systems [CSBE10].

3. Formal analysis framework for Gaussian optics

The proposed framework, given in Fig. 6, outlines the main idea and contribution of the current work. The two inputs to the framework are the description of the optical system and specification, i.e., the spatial organization of various components and their parameters (e.g., radius of the curvature of mirrors and distance between the component, etc.). In order to build the formal model of the given Gaussian optical system, we require a formalization of optical system structures (either a single system or a composed one) which consists of definitions of optical interfaces (e.g., plane or spherical) and optical components (e.g., lenses and mirrors). The next step is to formalize the physical concepts of rays and Gaussian beams. Building on these fundamentals, the next step is to derive the matrix model of the optical system, which is basically a composition of the matrix models of individual optical components. This step also includes the formalization of the ABCD-law of Gaussian optics, which describes the input-output relation of the given ray-transfer matrix and Gaussian beam parameters. Consequently, we model the general notion of quasi-optical systems and derive the important properties such as beam transformation and system magnification, which provide the basis for deriving the suitable parameters of Gaussian beams for a given system structure. We then develop a library of frequently used optical components such as thin lenses, thick lenses and mirrors. Since such components are the most basic blocks of optical systems, this library will help to formalize new optical systems as shown in Fig. 6. Finally, we can apply these developed theories to formally verify a variety of practical systems such as telescopes, laser devices and optical fiber-based systems. The output of the proposed framework is the formal proof that certifies that the system implementation meets its specification. The verified systems will then be made available in the library for future use either independently or as a part of a larger optical system.

4. Formalization of geometrical optics

In this section, we present a brief overview of our higher-order logic formalization of geometrical optics. The formalization consists of two parts: (1) fundamental concepts of optical systems' structures and light rays; and (2) verification of the ray-transfer matrix model of an arbitrary optical system.

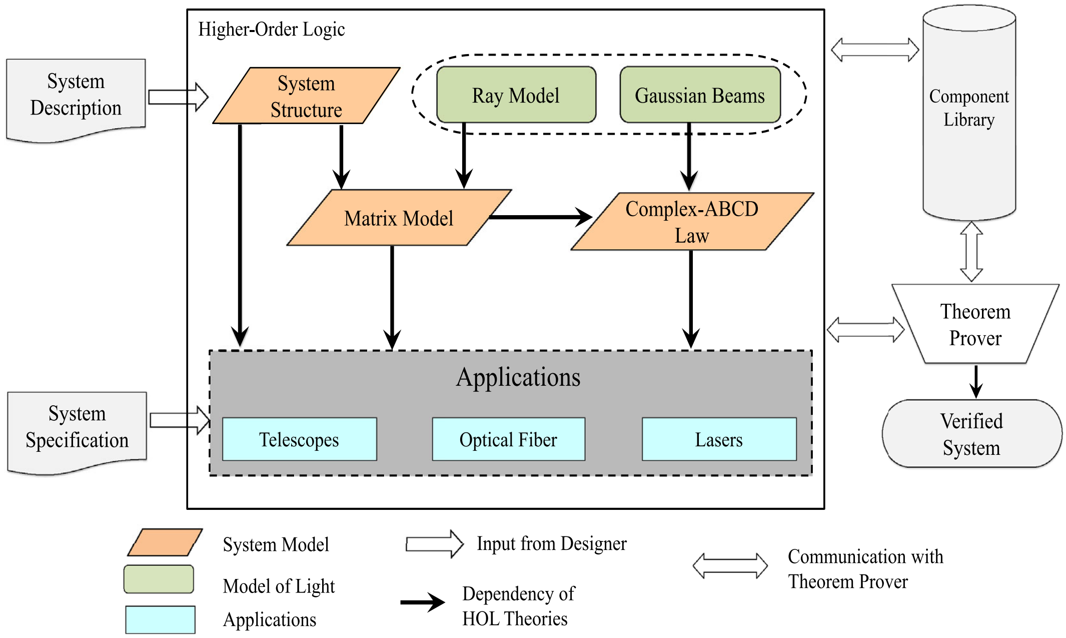

Fig. 6. Formal analysis framework

4.1. Optical system structure and ray

Ray optics explains the behavior of light when it passes through a free space and interacts with different optical interfaces. We can model free space using a pair of real numbers (n, d), which are essentially the refractive index and the total width. In the optics literature, light can be traced through an optical system by two techniques: sequential and non-sequential. In this paper, we only consider sequential ray tracing, which is based on the following modeling criteria [Trä07]:

- The type of each interface (e.g., plane or spherical, etc.) is known.
- The parameters of the corresponding interface (e.g., the radius of a spherical interface) are known in advance.
- The spacing between the optical components and the misalignment with respect to optical axis are provided by the system specification.
- The refractive indices of all materials and their dependence on wavelength are available.

We consider two optical interfaces, i.e., plane and spherical, which can be further classified as transmitted or reflected. In geometrical optics, we can describe a spherical interface by its radius of curvature (R). We model an optical component as a triplet (fs, i, ik), i.e., free space, an optical interface and its kind, i.e., transmitted or reflected. This modeling choice is based on the physical behavior as we consider a free space between two consecutive optical interfaces. Consequently, we define an optical system as a list of optical components followed by a free space. In HOL Light, we can use the available types (e.g., Real, Complex, etc.) to abbreviate new types. We use this feature to define a type abbreviation for a free space, as follows:

Definition 4.1 (*Free Space*)

```
new_type_abbrev ("free_space",`:ℝ × ℝ`)
```

In many situations, it is convenient to define new types in addition to those which are already available in HOL Light theories. One common way is to use enumerated types, where one gives an exhaustive list of members of the new type. In our formalization, we package different optical interfaces and their corresponding types as follows:

Definition 4.2 (*Optical Interfaces*)
```
define_type "optical_interface = plane | spherical ℝ"
define_type "interface_kind = transmitted | reflected"
```

where a spherical interface takes a real number representing its radius of curvature. Note that this datatype can easily be extended to many other optical interfaces if needed.

We next define an optical system as a sequence of optical components which are composed of free space, optical interface and its type (plane or spherical) yielding the corresponding constructors:

Definition 4.3 (*Optical Component and System*)
```
new_type_abbrev ("optical_component",
             `:free_space × optical_interface × interface_kind`)
new_type_abbrev ("optical_system",
             `:optical_component list × free_space`)
```

In the above formalization, we defined the basic constructors of optical systems which can be used to build the formal model of different applications, e.g., an optical resonator or a thick lens. In order to ensure the correct physical behavior of these models, we need to formalize some constraints which we call system specification. A value of type `free_space` represents a real space only if the refractive index is greater than zero. In addition, in order to have a fixed order in the representation of an optical system, we ensure that the distance of an optical interface relative to the previous interface is greater or equal to zero. We encode this requirement in the following predicate:

Definition 4.4 (*Valid Free Space*)
```
⊢ is_valid_free_space ((n,d):free_space) ⇔ 0 < n ∧ 0 ≤ d
```

We also need to assert the validity of a value of type `optical_interface` by ensuring that the radius of the curvature of the spherical interfaces is never equal to zero. This yields the following predicate:

Definition 4.5 (*Valid Optical Interface*)
```
⊢ (is_valid_interface plane ⇔ T) ∧
  (is_valid_interface (spherical R) ⇔ R ≠ 0)
```

Note that the radius of curvature of spherical interface can be positive or negative depending upon the concavity or convexity of the interface.

We use the above formalization to characterize valid optical systems by ensuring the validity of each component as follows:

Definition 4.6 (*Valid Optical Component*)
```
⊢∀fs i ik. is_valid_optical_component ((fs,i,ik):optical_component)
  ⇔ is_valid_free_space fs ∧ is_valid_interface i
```

Definition 4.7 (*Valid Optical System*)
```
⊢∀cs fs. is_valid_optical_system ((cs,fs):optical_system) ⇔
  ALL is_valid_optical_component cs ∧ is_valid_free_space fs
```

where `ALL` is a HOL light library function which checks that a predicate holds for all the elements of a list.

We can now formalize the physical behavior of a ray when it passes through an optical system. We only model the points where it hits an optical interface (instead of modeling all the points constituting the ray). So it is sufficient to just provide the distance of each of these hitting points to the axis and the angle taken by the ray at these points. Consequently, we should have a list of such pairs ($distance$, $angle$) for every component of a system. In addition, the same information should be provided for the source of the ray. For the sake of simplicity, we define a type for a pair ($distance$, $angle$) as `ray_at_point` as follows:

Definition 4.8 (*Ray*)

```
new_type_abbrev ("ray_at_point",‘:ℝ×ℝ‘)
new_type_abbrev ("ray",
  ‘:ray_at_point × ray_at_point ×
   (ray_at_point × ray_at_point) list‘)
```

where the first ray_at_point is the pair (*distance*, *angle*) for the source of the ray, the second one is the one after the first free space, and the list of ray_at_point pairs represents the same information for the interfaces and free spaces at every hitting point of an optical system.

Once again, we specify a valid ray by using some predicates. First of all, we define the behavior of a ray when it is traveling through a free space. This requires the position and orientation of the ray at the previous and current points of observation, and the free space itself.

Definition 4.9 (*Behavior of a Ray in Free Space*)

```
⊢ is_valid_ray_in_free_space
  (y0,θ0) (y1,θ1) ((n,d):free_space) ⇔ y1 = y0 + dθ0 ∧ θ0 = θ1
```

We next provide the specification of the valid behavior of a ray when hitting a particular interface. This requires the position and orientation of the ray at the previous and current interfaces, and the refractive indices before and after the component. Then, the predicate is defined by case analysis on the interface type as follows:

Definition 4.10 (*Behavior of a Ray at Given Interface*)

```
⊢ (is_valid_ray_at_interface (y0,θ0) (y1,θ1) n0 n1 plane transmitted ⇔
   y1 = y0 ∧ n0θ0 = n1θ1) ∧
  (is_valid_ray_at_interface (y0,θ0) (y1,θ1) n0 n1 (spherical R) transmitted ⇔
   let φi= θ0 + y1/R and φt = θ1 + y1/R in
   y1 = y0 ∧ n0φi = n1φt) ∧
  (is_valid_ray_at_interface (y0,θ0) (y1,θ1) n0 n1 plane reflected ⇔
   y1 = y0 ∧ n0θ0 = n0θ1) ∧
  (is_valid_ray_at_interface (y0,θ0) (y1,θ1) n0 n1 (spherical R) reflected ⇔
   let φi = y1/R - θ0 in y1 = y0 ∧ θ1 = -(θ0 + 2 φi))
```

where n_0 and n_1 represent the refractive indices before and after the interfaces, respectively. The above definition states some basic geometrical facts about the distance to the axis, and applies the paraxial Snell's law and the law of reflection [Trä07] to the orientation of the ray as shown in Fig. 1. Finally, we can recursively apply these predicates to define the behavior of a ray going through a series of optical components in an arbitrary optical system where more details can be found in [SAT13b].

4.2. Ray-transfer matrices of optical components

The main strength of ray optics is its matrix formalism [Trä07], which provides an efficient way to model all optical components in the form of a matrix. Indeed, a matrix relates the input and the output ray by a linear relation. For example, in the case of free space, the input and output ray parameters are related by two linear equations, i.e., $y_1 = y_0 + d\theta_0$ and $\theta_1 = \theta_0$, which further can be described as a matrix (also called ray-transfer matrix of free space). We can use the specification of the behavior of a ray in a free space to verify its ray-transfer-matrix as follows:

Table 1. Ray-transfer matrices of optical components

Component	HOL light formalization
Plane interface (reflection)	$\vdash \forall$ $\mathtt{n_0}$ $\mathtt{n_1}$ $\mathtt{y_0}$ θ_0 $\mathtt{y_1}$ θ_1. 0 < $\mathtt{n_0}$ $\wedge$ 0 < $\mathtt{n_1}$ $\wedge$ is_valid_ray_at_interface $(\mathtt{y_0},\theta_0)$ $(\mathtt{y_1},\theta_1)$ $\mathtt{n_0}$ $\mathtt{n_1}$ plane reflected $\Longrightarrow$ $\begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ** \begin{bmatrix} y_0 \\ \theta_0 \end{bmatrix}$
Plane interface (transmission)	$\vdash \forall$$\mathtt{n_0}$ $\mathtt{n_1}$ $\mathtt{y_0}$ θ_0 $\mathtt{y_1}$ θ_1. 0 < $\mathtt{n_0}$ $\wedge$ 0 < $\mathtt{n_1}$ $\wedge$ is_valid_ray_at_interface $(\mathtt{y_0},\theta_0)$ $(\mathtt{y_1},\theta_1)$ $\mathtt{n_0}$ $\mathtt{n_1}$ plane transmitted $\Longrightarrow$ $\begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{n_0}{n_1} \end{bmatrix} ** \begin{bmatrix} y_0 \\ \theta_0 \end{bmatrix}$
Spherical interface (reflection)	$\vdash \forall$$\mathtt{n_0}$ $\mathtt{n_1}$ $\mathtt{y_0}$ θ_0 $\mathtt{y_1}$ θ_1 R. 0 < $\mathtt{n_0}$ $\wedge$ 0 < $\mathtt{n_1}$ $\wedge$ (is_valid_interface (spherical R) $\wedge$ is_valid_ray_at_interface $(\mathtt{y_0},\theta_0)$ $(\mathtt{y_1},\theta_1)$ $\mathtt{n_0}$ $\mathtt{n_1}$ (spherical R) reflected $\Longrightarrow$ $\begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -\frac{2}{R} & 1 \end{bmatrix} ** \begin{bmatrix} y_0 \\ \theta_0 \end{bmatrix}$
Spherical interface (transmission)	$\vdash \forall$$\mathtt{n_0}$ $\mathtt{n_1}$ $\mathtt{y_0}$ θ_0 $\mathtt{y_1}$ θ_1 R. 0 < $\mathtt{n_0}$ $\wedge$ 0 < $\mathtt{n_1}$ $\wedge$ (is_valid_interface (spherical R) $\wedge$ is_valid_ray_at_interface $(\mathtt{y_0},\theta_0)$ $(\mathtt{y_1},\theta_1)$ $\mathtt{n_0}$ $\mathtt{n_1}$ (spherical R) transmitted $\Longrightarrow$ $\begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{n_0-n_1}{Rn_1} & \frac{n_0}{n_1} \end{bmatrix} ** \begin{bmatrix} y_0 \\ \theta_0 \end{bmatrix}$

Theorem 4.1 (Ray-Transfer-Matrix for Free Space)

$\vdash \forall$n d $\mathtt{y_0}$ θ_0 $\mathtt{y_1}$ θ_1.
  is_valid_free_space (n,d) $\wedge$
  is_valid_ray_in_free_space $(\mathtt{y_0},\theta_0)$ $(\mathtt{y_1},\theta_1)$ (n,d)) $\Longrightarrow$ $\begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix} = \begin{bmatrix} 1 & d \\ 0 & 1 \end{bmatrix} ** \begin{bmatrix} y_0 \\ \theta_0 \end{bmatrix}$

where ** describes vector-vector or vector-matrix multiplication in HOL light. The first assumption ensures the validity of free space and the second assumption ensures the valid behavior of a ray in free space. We proved the above theorem using the definitions is_valid_free_space and is_valid_ray_in_free_space along with the properties of the vectors. Similarly, we proved the ray-transfer matrices of plane and spherical interfaces for the case of transmission and reflection, as listed in Table 1. The availability of these theorems in our formalization is quite handy and it helps to reduce the interactive verification efforts for applications. The proof steps for these theorems are similar to the ones for Theorem 4.1. In order to make the proof of these theorems automatic, we build a tactic common_prove which is mainly based on the simplification with the above-mentioned definitions and the application of matrix operations.

Our next goal is to formally prove that any optical interface can be described by a general ray-transfer-matrix relation. Mathematically, this relation is described in the following theorem:

Theorem 4.2 (Ray-Transfer-Matrix any Interface)

$\vdash \forall$$\mathtt{n_0}$ $\mathtt{n_1}$ $\mathtt{y_0}$ θ_0 $\mathtt{y_1}$ θ_1 i ik. is_valid_interface i $\wedge$
  is_valid_ray_at_interface $(\mathtt{y_0},\theta_0)$ $(\mathtt{y_1},\theta_1)$ $\mathtt{n_0}$ $\mathtt{n_1}$ i ik $\wedge$
  0 < $\mathtt{n_0}$ $\wedge$ 0 < $\mathtt{n_1}$ $\Longrightarrow$ $\begin{bmatrix} y_1 \\ \theta_1 \end{bmatrix}$ = interface_matrix $\mathtt{n_0}$ $\mathtt{n_1}$ i ik ** $\begin{bmatrix} y_0 \\ \theta_0 \end{bmatrix}$

where interface_matrix accepts the refractive indices ($\mathtt{n_0}$ and $\mathtt{n_1}$), interface (i) and the type of the interface (ik) and returns the corresponding matrix of the system as described in Table 2. In the above theorem, both assumptions ensure the validity of the interface and behavior of the ray at the interface, respectively. This theorem is mainly proved by case splitting on i and ik.

Now, equipped with the above theorem, the next step is to formally verify the ray-transfer-matrix relation for a complete optical system as given in Eq. 1. It is important to note that in this equation, individual matrices of optical components are composed in reverse order, which indeed represents the implementation of an optical system. We formalize systems composition in the following recursive definition:

Definition 4.11 (*System Composition*)
```
⊢ system_composition ([],n,d) ⇔ free_space_matrix d ∧
  system_composition (CONS ((nt,dt),i,ik) cs,n,d) ⇔
  (system_composition (cs,n,d) **
  interface_matrix nt (head_index (cs,n,d)) i ik) **
  free_space_matrix dt
```

where the type of `system_composition` is : `optical_system` $\rightarrow \mathbb{R}^{2x2}$, i.e., it takes an optical system and returns a (2×2) matrix.

We verify the generalized ray-transfer-matrix relation for an optical system as follows:

Theorem 4.3 (Ray-Transfer-Matrix for an Optical System)
```
⊢ ∀sys ray. is_valid_optical_system sys ∧
  is_valid_ray_in_system ray sys ⟹
  let (y0,θ0),(y1,θ1),rs = ray in
  let yn,θn = last_ray_at_point ray in
```
$$\begin{bmatrix} y_n \\ \theta_n \end{bmatrix} = \texttt{system\_composition sys **} \begin{bmatrix} y_0 \\ \theta_0 \end{bmatrix}$$

where the parameters `sys` and `ray` represent the optical system and an arbitrary ray, respectively. The predicate `is_valid_ray_in_system` takes two parameters, i.e., `ray` and `sys` and ensures that the behavior of `ray` is valid in the system (`sys`). The function `last_ray_at_point` returns the last `ray_at_point` of the ray in the system. The theorem is proved by induction on the length of the system and by using previous results and definitions.

The above described implementation of optical systems and corresponding ray-transfer matrix relation only hold for a single optical system consisting of different optical components. Our main requirement is to extend this model for a general system which is composed of n optical subsystems as shown in Fig. 2. Introducing another layer of optical systems has two advantages: (1) we can use the formalization and properties of already available optical systems if they are being used in another system; (2) we can derive the properties of an optical system without the prior knowledge of actual parameters, i.e., replacing it with an unknown system. For example, we can compose two known and one unknown systems as a list $[system_1;\ unknown;\ system_2]$. We formalize the notion of a composed optical system as follows:

Definition 4.12 (*Composed Optical System Model*)
```
⊢ composed_system [] = I ∧
  composed_system (CONS sys cs) =
  composed_system cs ** system_composition sys
```

where `I` represents the identity matrix and the function `composed_system` accepts a list of optical systems `:(optical_system)list` and returns the overall system model by the recursive application of the function `system_composition` (Definition 4.11). We define the validity of a composed optical system by ensuring the validity of each involved optical system as follows:

Definition 4.13 (*Valid Composed Optical System*)

```
⊢∀(sys:optical_system list). is_valid_composed_system sys ⇔
    ALL is_valid_optical_system sys
```

In order to reason about composed optical systems, we need to give some new definitions about the ray behavior inside a composed optical system. One of the easiest ways is to consider n rays corresponding to n optical systems individually and then make sure each ray is the same as the one applied at the input. This can be done by ensuring that the starting point of each ray is the same as the ending point of the previous ray as shown in Fig. 2. We encode this physical behavior of a ray as follows:

Definition 4.14 (*Valid General Ray*)

```
⊢ is_valid_genray ([]:ray list) ⇔ F ∧
  is_valid_genray (CONS h t) ⇔
  (last_single_ray h = fst_single_ray (HD t) ∧
  is_valid_genray t)
```

where `fst_single_ray`, `last_single_ray` and `HD`, provide the first and last single ray at a point and first element of a list, respectively. Along the same lines, we also specify the behavior of a ray when it passes through each optical system by the function `is_valid_gray_in_system`. Finally, we verify that the ray-transfer-matrix relation holds for composed optical systems which ensures that all valid properties for a single optical system can be generalized to the composed system as well.

Theorem 4.4 (Ray-Transfer-Matrix for a Composed Optical System)

$$\begin{array}{l}
\vdash \forall(\texttt{sys: optical_system list})\ (\texttt{ray: ray list}).\\
\quad \texttt{is_valid_composed_system sys} \wedge\\
\quad \texttt{is_valid_gray_in_system ray sys} \wedge\\
\quad \texttt{is_valid_genray ray} \Longrightarrow\\
\quad \texttt{let } (\mathrm{y}_0,\theta_0) = \texttt{fst_single_ray (HD ray) in}\\
\quad \texttt{let } (\mathrm{y}_n,\theta_n) = \texttt{last_single_ray (LAST ray) in}\\
\quad \begin{bmatrix} \mathrm{y_n} \\ \theta_\mathrm{n} \end{bmatrix} = \texttt{composed_system sys **} \begin{bmatrix} \mathrm{y_0} \\ \theta_0 \end{bmatrix}
\end{array}$$

We next present the formalization of Gaussian beams and their physical behavior in HOL.

5. Formalization of Gaussian beams

In this section, we present the formalization related to Gaussian beams, which can be divided into three parts: (1) Formalization of the q-parameters and verification of some related properties; (2) Formalization of the paraxial Helmholtz equation and the verification that the envelope of a Gaussian beam [Eq. (6)] satisfies the paraxial Helmholtz equation; and 3) Formalization of q-parameters transformation in optical systems and verification of the ABCD-law [Eq. (12)].

5.1. Formalization of q-parameters

In the optics literature, Gaussian beams are defined in different forms depending upon the application of the beam transformation. In any case, a Gaussian beam can be characterized by the corresponding q-parameter, i.e., ($q(z) = z + jz_R$). Furthermore, the Rayleigh range ($z_R = \frac{\pi w_0^2}{\lambda}$), can be described by two parameters, i.e., value of the beam width at $z = 0$, (w_0) and wavelength (λ). Thus, the q-parameter can be completely characterized by a triplet (w_0, λ, z) and hence the Gaussian beam. We define the Rayleigh range and q-parameter in HOL Light as follows:

Definition 5.1 (*Rayleigh Range and q-parameter*)

$\vdash \forall$ $\mathtt{w_0\ lam.\ rayleigh\_range\ w_0\ lam} = \dfrac{\pi\, \mathtt{w_0^2}}{\mathtt{lam}}$

$\vdash \forall$ $\mathtt{z\ w_0\ lam.\ q\ (z,w_0,lam) = z + j\ (rayleigh\_range\ w_0\ lam)}$

where j represents the imaginary unit ($j = \sqrt{-1}$).

One of the most important forms of the q-parameters is given in Eq. 7, i.e., in the form of $R(z)$ (Eq. 8) and $W(z)$ (Eq. 9). We define $R(z)$ and $W(z)$ to make the reasoning simpler, as given in the following definitions:

Definition 5.2 (*Wavefront Radius and Beam Width*)

$\vdash \forall$ $\mathtt{z\ w_0\ lam.\ R\ z\ w_0\ lam = z} \left[1 + \left(\dfrac{\mathtt{rayleigh\_range\ w_0\ lam}}{\mathtt{z}}\right)^2\right]$

$\vdash \forall$ $\mathtt{z\ w_0\ lam.\ W\ z\ w_0\ lam = w_0} \left[1 + \left(\dfrac{\mathtt{rayleigh\_range\ w_0\ lam}}{\mathtt{z}}\right)^2\right]^{\frac{1}{2}}$

where the functions R and W are both of type $\mathbb{R}\rightarrow\mathbb{R}\rightarrow\mathbb{R}\rightarrow\mathbb{R}$, and take three parameters z, $\mathtt{w_0}$ and lam and return a real number corresponding to Eqs. (8) and (9), respectively. Next, we use these definitions to verify Eq. (7) as follows:

Theorem 5.1 (q-Parameter Alternative Form)

$\vdash \forall$ $\mathtt{z\ w_0\ lam.}\ 0 < \mathtt{w_0} \ \wedge\ 0 < \mathtt{lam}\ \wedge\ \mathtt{z} \neq 0$

$$\frac{1}{\mathtt{q(z, w_0, lam)}} = \frac{1}{\mathtt{R\ z\ w_0\ lam}} - \mathtt{j}\ \frac{\mathtt{lam}}{\pi(\mathtt{W\ z\ w_0\ lam})^2}$$

The proof of this theorem mainly involves complex analysis and some properties of q, R and W, some of which are listed here:

Lemma 5.1 (Properties)

$\vdash \forall\ \mathtt{z\ w_0\ lam.}$
$\quad \mathtt{z} \neq 0\ \wedge (\mathtt{rayleigh\_range\ w_0\ lam})^2 = \mathtt{z}^2 \implies (\mathtt{R\ z\ w_0\ lam}) \neq 0$

$\vdash \forall\ \mathtt{z\ w_0\ lam.}$
$\quad 0 < \mathtt{w_0}\ \wedge\ 0 < \mathtt{lam}\ \wedge\ 0 \leq \mathtt{z} \implies 0 < \mathtt{W\ z\ w_0\ lam}$

$\vdash \forall\ \mathtt{z\ w_0\ lam.}$
$\quad 0 < \mathtt{w_0}\ \wedge\ 0 < \mathtt{lam} \implies \mathtt{q\ (z, w_0, lam)} \neq 0$

$\vdash \forall\ \mathtt{z\ w_0\ lam.}$
$\quad 0 < \mathtt{w_0}\ \wedge\ 0 < \mathtt{lam} \implies 0 < \mathtt{rayleigh\_range\ z\ w_0\ lam}$

The alternative form of q-parameter (proved in Theorem 5.1) is quite helpful to verify the general form of a Gaussian beam [Eq. (10)] and the corresponding intensity (Eq. 11).

5.2. Formalization of the paraxial Helmholtz equation

Our first step is to formalize the notion of the transverse Laplacian operator ($\nabla_T^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$) for arbitrary functions as follows:

Definition 5.3 (*Laplacian*)
$\vdash \forall$ `f x y.`
`laplacian f (x, y) = higher_complex_derivative 2 (λx. f (x, y)) x +`
`higher_complex_derivative 2 (λy. f (x, y)) y`

where `higher_complex_derivative` represents the nth-order complex derivative of a function. We use `laplacian` to formalize the Paraxial Helmholtz equation as follows:

Definition 5.4 (*Paraxial Helmholtz Equation*)
$\vdash$ `Paraxial_Helmholtz_eq A (x,y,z) k` $\Leftrightarrow$
`laplacian(λ(x, y). A (x, y, z)) (x, y) −`
`2 j k complex_derivative (λz. A (x, y, z)) z = 0`

where `Paraxial_Helmholtz_eq` accepts a function `A` of type $((\mathbb{C}\times\mathbb{C}\times\mathbb{C}) \rightarrow \mathbb{C})$, a triplet `(x,y,z)` and a wave number `k` and returns the Paraxial Helmholtz equation. Next, we formalize the paraxial wave, given in Eq. (2), as follows:

Definition 5.5 (*Paraxial Wave*)
$\vdash \forall$ `A x y k z.`
`paraxial_wave A x y z k = A(x, y, z) exp(−jkz)`

where `A`$:((\mathbb{C}\times\mathbb{C}\times\mathbb{C}) \rightarrow \mathbb{C})$ represents the complex amplitude of the paraxial wave. The function `exp` represents the complex-valued exponential function in HOL Light.

We define the q-parameter-based amplitude of the paraxial wave [Eq. (6)] as follows:

Definition 5.6 (*q-parameter-Based Solution*)
$\vdash \forall$ $\mathtt{A_0}$ `k x y z` $\mathtt{w_0}$ `lam.`
$\mathtt{q\_parameter\_amplitude}\ \mathtt{A_0}\ \mathtt{z\ x\ y\ k}\ \mathtt{w_0}\ \mathtt{lam} = \dfrac{\mathtt{A_0}}{\mathtt{q}\ (\mathtt{z}, \mathtt{w_0}, \mathtt{lam})} \exp\left(\dfrac{-\mathtt{jk}(\mathtt{x}^2+\mathtt{y}^2)}{2\mathtt{q}\ (\mathtt{z}, \mathtt{w_0}, \mathtt{lam})}\right)$ where $\mathtt{A_0}$ is a complex-valued constant. The function `q` represents the q-parameter as described in Definition 5.1.

Now equipped with the above-mentioned formal definitions, an important requirement is to verify that the q-parameters based solution (Definition 5.6) satisfies the paraxial Helmholtz equation (Definition 5.4). We verify this requirement in the following theorem:

Theorem 5.2 (Helmholtz Equation Verified)
$\vdash \forall$ $\mathtt{A_0}$ `x y z` $\mathtt{w_0}$ `lam k.`
$0 < \mathtt{w_0} \ \wedge\ 0 < \mathtt{lam} \implies$
`Paraxial_Helmholtz_eq (λ(x, y, z). q_parameter_amplitude` $\mathtt{A_0}$ `z x y k` $\mathtt{w_0}$ `lam) (x, y, z) k`

where both assumptions ensure that the value of `q` (q-parameter) is not zero. The proof of this theorem is mainly based on three lemmas about the complex differentiation of `q_parameter_amplitude` with respect to parameters `x`, `y` and `z`. The proof of these lemmas is mainly done using the automated tactic called `COMPLEX_DIFF_TAC` (already available in HOL light and developed by Harrison), which can automatically compute the complex differentiation of complicated functions. Indeed, this tactic saves a lot of user interaction time while proving theorems which involve complex differentiation.

Our next step is to derive the expression representing a paraxial wave as a Gaussian beam (i.e., Eq. 10), as follows:

Theorem 5.3 (Gaussian Beam)
$\vdash \forall$ x y z w_0 lam A_0 k.
$0 < w_0 \;\wedge\; 0 < \mathtt{lam} \;\wedge\; z \neq 0 \implies$
paraxial_wave (λ(x, y, z). q_parameter_amplitude A_0 z x y k w_0 lam) x y z k =

$$\mathtt{let}\ A_c = \frac{A_0}{j(\mathtt{rayleigh\_range}\ w_0\ \mathtt{lam})}\ \mathtt{in}$$
$$A_c \frac{w_0}{\mathtt{W}\ z\ w_0\ \mathtt{lam}} \exp\left[-\frac{k\cdot \mathtt{lam}}{2\pi}\frac{x^2+y^2}{(\mathtt{W}\ z\ w_0\ \mathtt{lam})^2}\right]$$
$$\exp\left[-jkz - jk\frac{x^2+y^2}{2(\mathtt{R}\ z\ w_0\ \mathtt{lam})} + j\arctan\left(\frac{z}{\mathtt{rayleigh\_range}\ w_0\ \mathtt{lam}}\right)\right]$$

where `arctan` represents the inverse tangent function in HOL Light. The proof of this theorem mainly requires two lemmas, i.e., expressing $q-$parameter in equivalent form (Theorem 5.1) and expressing `arctan` as an argument of `exp`, given as follows:

Lemma 5.2 (arctanas an Argument of exp)
$\vdash \forall$ z w_0 lam.
$0 < w_0 \;\wedge\; 0 < \mathtt{lam} \implies$

$$\exp\left[j\arctan\left(\frac{z}{\mathtt{rayleigh\_range}\ w_0\ \mathtt{lam}}\right)\right] = \frac{(jz\ +\ \mathtt{rayleigh\_range}\ w_0\ \mathtt{lam})}{\sqrt{z^2\ +\ (\mathtt{rayleigh\_range}\ w_0\ \mathtt{lam})^2}}$$

Finally, we define the intensity of a paraxial wave as follows:

Definition 5.7 (*Beam Intensity*)
$\vdash \forall$ A x y z k.
beam_intensity A x y z k = $\|$ (paraxial_wave A x y z k)^2 $\|$

where $\mathtt{A}:((\mathbb{C}\times\mathbb{C}\times\mathbb{C}) \to \mathbb{C})$ represents the complex amplitude of the paraxial wave. The function $\|\ .\ \|$ represents the complex norm of a function in HOL light.

We use the above definition to verify the general expression for the intensity of a Gaussian beam [Eq. (11)] in the following theorem:

Theorem 5.4 (Gaussian Beam Intensity)
$\vdash \forall$ A_0 x y z k w_0 lam.
$0 < w_0 \;\wedge\; 0 < \mathtt{lam} \;\wedge\; z \neq 0 \implies$
beam_intensity (λ(a, b, z). q_parameter_amplitude A_0 z a b k w_0 lam) x y z k =

$$\left\| \frac{A_0}{j(\mathtt{rayleigh\_range}\ w_0\ \mathtt{lam})} \right\|^2 \left(\frac{w_0}{\mathtt{W}\ z\ w_0\ \mathtt{lam}}\right)^2 \exp\left[-\frac{k\frac{\mathtt{lam}}{\pi}(x^2+y^2)}{(\mathtt{W}\ z\ w_0\ \mathtt{lam})^2}\right]$$

The proof of this theorem is mainly based on Theorem 5.3 and the properties of complex numbers.

We discuss the concepts and formalization behind the propagation of Gaussian beams in the next section.

5.3. Formalization of beam transformation for optical systems

In our formalization of the q-parameter of Gaussian beams, we consider that the size of the beam waist radius w_0 and its location z is already provided by the physicists or optical system design engineers. Indeed, these two parameters are sufficient to compute beam width $W(z)$ and wavefront radius of curvature $R(z)$ because the wavelength λ is fixed throughout the design life-cycle. Mathematically, this notion can be represented as a transformation $w_0, z \to W(z), R(z)$. However, in some practical situations, we may know only the beam width or wavefront radius of curvature. For example, such information can be taken from a physical experiment or if the beam is being generated from a known source [Gol98].

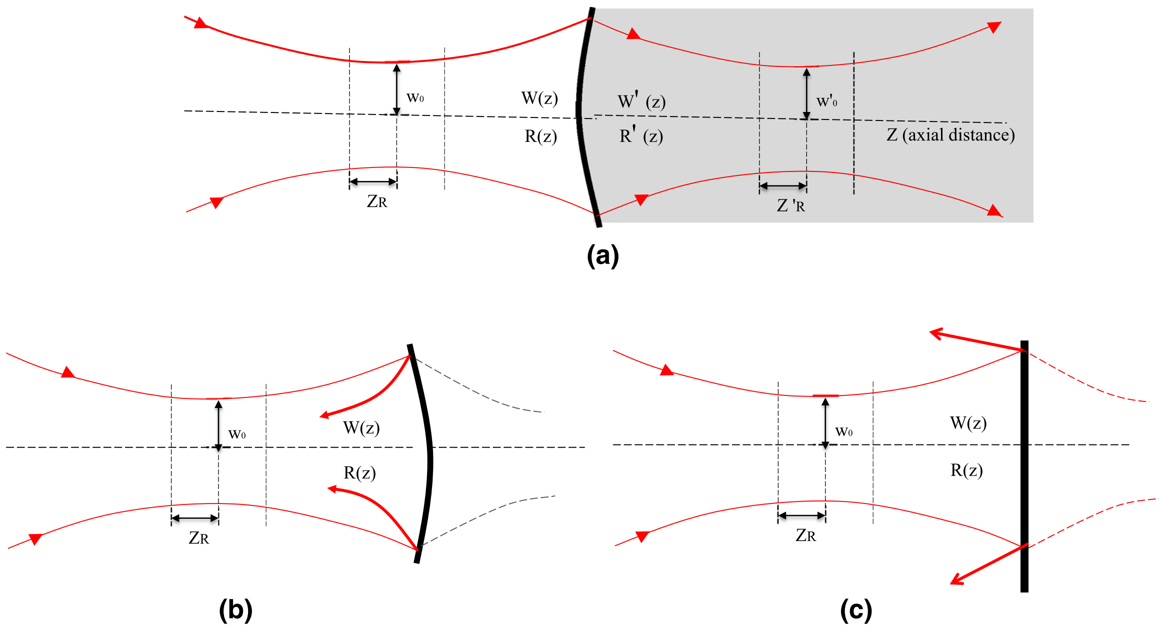

Fig. 7. Behavior of the gaussian beam at different interfaces

Our goal is to formalize the physical behavior of a Gaussian beam when it passes through an optical system. We only model the points where it hits an optical interface (e.g., spherical or plane interface). It is evident from the previous discussion that q-parameter is sufficient to characterize a Gaussian beam. Furthermore, λ is fixed which leads to the requirement of considering two parameters, i.e., w_0 and z. So we just need to provide the information about (z, w_0) at each interface. Consequently, we should have a list of such pairs for every component of a system. In addition, the same information should be provided for the source of the beam. We define a type for a pair $(z,\ w_0)$ as `single_q`. This yields the following definition:

Definition 5.8 (*Beam*)

```
new_type_abbrev ("single_q",`:ℝ × ℝ`)
new_type_abbrev ("beam",
              `:single_q × single_q ×
               (single_q × single_q) list`)
```

where the first `single_q` is the pair $(z,\ w_0)$ for the source of the beam, the second `single_q` is the one after the first free space, and the list of `single_q` pairs represents the same information for the interfaces and free spaces at every hitting point of an optical system.

The transmission of a Gaussian beam in an optical system depends on the nature of the components used in that system. It is known that a Gaussian beam remains a Gaussian beam when it is transmitted through a set of circularly symmetric optical components aligned with an optical axis and it is theoretically and experimentally proved for the case of real-world optical systems [ST07]. However, only the beam waist and curvature are modified so that the beam is only reshaped as compared to the input beam. If a Gaussian beam is subject to transmission in a free space of width d, only one parameter of the Gaussian beam is modified, i.e., z becomes $z + d$. When a beam transmits through a plane interface it only scales with respect to the refractive indices of input and output planes. However, in the case of transmission through a spherical interface, the beam width remains the same but the output beam has to satisfy the lens formula, i.e., $\frac{1}{q_2} = \frac{1}{q_1} - \frac{1}{f}$, where q_1, q_2 and f are the input and output beam q-parameters and the focal length of the spherical interface as shown in Fig. 7a. For the case of reflection from a plane interface, the input Gaussian beam bounces back without any change in its curvature (Fig. 7c). On the other hand, reflection from a curved interface results in a modified lens formula, i.e., $\frac{1}{q_2} = \frac{1}{q_1} + \frac{1}{f}$ with no alteration in the beam width as shown in Fig. 7b.

We specify a valid beam by using some predicates. First of all, we define the behavior of a beam when it is traveling through a free space. This requires the position of the beam at the previous and current points of observation, and the free space itself.

Definition 5.9 (*Beam in Free Space*)

$$\vdash \texttt{is\_valid\_beam\_in\_free\_space}\ (z, w_0)\ (z', w_0')\ (n_0, d) \Leftrightarrow w_0' = w_0 \ \wedge\ z' = z + d$$

where (z, w_0) and (z', w_0') represent the `single_q` at two points. The pair (n_0, d) represents a free space with refractive index n_0 and width d.

Now we specify the valid behavior of a beam at the plane and spherical interfaces as follows:

Definition 5.10 (*Beam at Plane Interface*)

$$\begin{array}{l}\vdash (\texttt{is\_valid\_beam\_at\_plane\_interface}\ (z, w_0)\ (z', w_0')\ \texttt{lam}\ n\ n'\ \texttt{plane\_transmitted} \Leftrightarrow \\ \quad z' = z\dfrac{n'}{n} \ \wedge\ w_0' = w_0\sqrt{\dfrac{n'}{n}} \ \wedge\ 0 < n \ \wedge\ 0 < n') \ \wedge \\ \quad (\texttt{is\_valid\_beam\_at\_plane\_interface}\ (z, w_0)\ (z', w_0')\ \texttt{lam}\ n\ n'\ \texttt{plane reflected} \Leftrightarrow \\ \quad z' = z \ \wedge\ w_0' = w_0 \ \wedge\ 0 < n \ \wedge\ 0 < n')\end{array}$$

where `is_valid_beam_at_plane_interface` accepts two `single_q`, wavelength `lam` and two refractive indices n and n' before and after the plane interface (transmitted and reflected), and returns the physical behavior of the beam described above (as shown in Fig. 7). Similarly, we formally specify the physical behavior of the beam at the spherical interface in the following definition:

Definition 5.11 (*Beam at Spherical Interface*)

$$\begin{array}{l}\vdash \texttt{is\_valid\_beam\_at\_spherical\_interface}\ (z, w_0)\ (z', w_0')\ \texttt{lam}\ n\ n'\ (\texttt{spherical}\ R)\ \texttt{transmitted} \Leftrightarrow \\ \quad \texttt{valid\_single\_q}\ (z, w_0) \ \wedge\ \texttt{valid\_single\_q}\ (z', w0') \ \wedge \\ \quad 0 < n \ \wedge\ 0 < n' \ \wedge\ -\dfrac{n - n'}{nR} = \dfrac{1}{q\ (z, w_0, \texttt{lam})} \ \wedge \\ \quad \dfrac{1}{R\ z'\ w_0'\ \texttt{lam}} = \dfrac{n}{n'}\dfrac{1}{R\ z\ w_0\ \texttt{lam}} + \dfrac{n - n'}{n'R} \ \wedge\ (W\ z'\ w_0'\ \texttt{lam}) = \sqrt{\dfrac{n'}{n}}(W\ z\ w_0\ \texttt{lam}) \ \wedge \\ \quad \texttt{is\_valid\_beam\_at\_spherical\_interface}\ (z, w_0)\ (z', w_0')\ \texttt{lam}\ n\ n'\ (\texttt{spherical}\ R)\ \texttt{reflected} \Leftrightarrow \\ \quad \texttt{valid\_single\_q}\ (z, w_0) \ \wedge\ \texttt{valid\_single\_q}\ (z', w0') \ \wedge \\ \quad 0 < n \ \wedge\ 0 < n' \ \wedge\ q\ (z, w_0, \texttt{lam}) \neq \dfrac{R}{2} \ \wedge \\ \quad \dfrac{1}{R\ z'\ w_0'\ \texttt{lam}} = \dfrac{1}{R\ z\ w_0\ \texttt{lam}} - \dfrac{2}{R} \ \wedge\ (W\ z'\ w_0'\ \texttt{lam}) = (W\ z\ w_0\ \texttt{lam})\end{array}$$

where `valid_single_q` ensures that w_0 is positive and z is not equal to zero.

Note that we describe separately the valid behavior of beam at plane and spherical interface for the sake of convenience and finally we combine them into one definition called `is_valid_beam_at_interface`. Along the same lines, we also specify the behavior of the beam through an arbitrary optical system by ensuring its validity at each optical interface and free space. This function is named as `is_valid_beam_in_system`, where more details can be found in the source code [Sid15].

In order to ensure the correctness of our definitions and to facilitate the formal analysis of practical systems, we verify three classical results of Gaussian beams theory: (1) ABCD-law for each optical interface (i.e., free space, spherical and plane for both reflection and transmission); (2) ABCD-law for an arbitrary optical system; and (3) for composed optical systems as follows:

Theorem 5.5 (ABCD-Law for Interface)

$$\begin{array}{l}
\vdash \forall \mathtt{i\ ik\ z\ w_0\ z'\ w'_0\ lam\ n\ n'}. \\
\quad \mathtt{is_valid_beam_at_interface\ (z, w_0)\ (z', w'_0)\ lam\ n\ n'\ i\ ik} \ \wedge \\
\quad \mathtt{is_valid_interface\ i} \ \wedge\ 0 < \mathtt{lam} \implies \\
\quad \mathtt{let} \begin{bmatrix} \mathtt{A} & \mathtt{B} \\ \mathtt{C} & \mathtt{D} \end{bmatrix} = (\mathtt{interface_matrix\ n\ n'\ i\ ik})\ \mathtt{in} \implies \\
\quad \mathtt{q}\ (\mathtt{z'}, \mathtt{w'_0}, \mathtt{lam}) = \dfrac{\mathtt{Aq}\ (\mathtt{z}, \mathtt{w_0}, \mathtt{lam}) + \mathtt{B}}{\mathtt{Cq}\ (\mathtt{z}, \mathtt{w_0}, \mathtt{lam}) + \mathtt{D}}
\end{array}$$

where is_valid_beam_at_interface ensures the valid behavior at each interface i of type ik, i.e., reflected or transmitted. The function is_valid_interface ensures that each interface i is indeed a valid interface. The assumption $0 < \mathtt{lam}$ is required to ensure that the wavelength is greater than zero. Finally, the function interface_matrix represents the corresponding matrix of each optical component. We prove Theorem 5.5 by induction on i and ik along with the properties of complex numbers.

Next, we verify the complex ABCD law for an arbitrary optical system as follows:

Theorem 5.6 (ABCD-Law for Optical System)

$$\begin{array}{l}
\vdash \forall \mathtt{sys\ beam\ lam\ A\ B\ C\ D}. \\
\quad \mathtt{is_valid_beam_in_system\ beam\ lam\ sys} \ \wedge \\
\quad \mathtt{is_valid_optical_system\ sys} \ \wedge\ 0 < \mathtt{lam} \ \wedge \\
\quad \begin{bmatrix} \mathtt{A} & \mathtt{B} \\ \mathtt{C} & \mathtt{D} \end{bmatrix} = \mathtt{system_composition\ sys} \implies \\
\quad \mathtt{let}\ (\mathtt{z}, \mathtt{w_0}), (\mathtt{z'}, \mathtt{w'_0}), \mathtt{rs} = \mathtt{beam\ in} \\
\quad \mathtt{let}\ (\mathtt{z_n}, \mathtt{w_{0n}}) = \mathtt{last_single_beam\ beam\ in} \\
\quad \mathtt{q}\ (\mathtt{z_n}, \mathtt{w_{0n}}, \mathtt{lam}) = \dfrac{\mathtt{Aq}\ (\mathtt{z}, \mathtt{w_0}, \mathtt{lam}) + \mathtt{B}}{\mathtt{Cq}\ (\mathtt{z}, \mathtt{w_0}, \mathtt{lam}) + \mathtt{D}}
\end{array}$$

where is_valid_beam_in_system ensures the valid behavior of the beam in optical system sys. The function is_valid_system ensures the validity of the optical systems structure. Finally, the function system_composition represents the corresponding matrix of the optical system. We prove Theorem 5.6 by induction on sys and length of the beam along with some complex arithmetic reasoning.

Similarly, we verify the complex ABCD law for the composed systems where a system is composed of multiple optical systems, given as follows:

Theorem 5.7 (ABCD-Law for Composed System)

$$\begin{array}{l}
\vdash \forall \mathtt{c_sys\ gbeam\ lam\ A\ B\ C\ D}. \\
\quad \mathtt{is_valid_gbeam_in_c_system\ gbeam\ lam\ c_sys} \ \wedge\ \mathtt{is_valid_gen_beam\ gbeam} \ \wedge \\
\quad \mathtt{is_valid_composed_system\ c_sys} \ \wedge\ 0 < \mathtt{lam} \ \wedge \\
\quad \begin{bmatrix} \mathtt{A} & \mathtt{B} \\ \mathtt{C} & \mathtt{D} \end{bmatrix} = \mathtt{composed_system\ c_sys} \implies \\
\quad \mathtt{let}\ (\mathtt{z}, \mathtt{w_0}), (\mathtt{z'}, \mathtt{w'_0}) = \mathtt{beam_origin\ gbeam\ in} \\
\quad \mathtt{let}\ (\mathtt{z_n}, \mathtt{w_{0n}}) = \mathtt{beam_end\ gbeam\ in} \\
\quad \mathtt{q}\ (\mathtt{z_n}, \mathtt{w_{0n}}, \mathtt{lam}) = \dfrac{\mathtt{Aq}\ (\mathtt{z}, \mathtt{w_0}, \mathtt{lam}) + \mathtt{B}}{\mathtt{Cq}\ (\mathtt{z}, \mathtt{w_0}, \mathtt{lam}) + \mathtt{D}}
\end{array}$$

where is_valid_gbeam_in_c_system ensures the valid behavior of the gbeam in a composed optical system c_sys. The functions beam_origin and beam_end provide the starting and ending point of a general beam, respectively.

This concludes our formalization of the Gaussian beam transformation for arbitrary optical systems. In the next section, we present the formalization Quasi-optical systems.

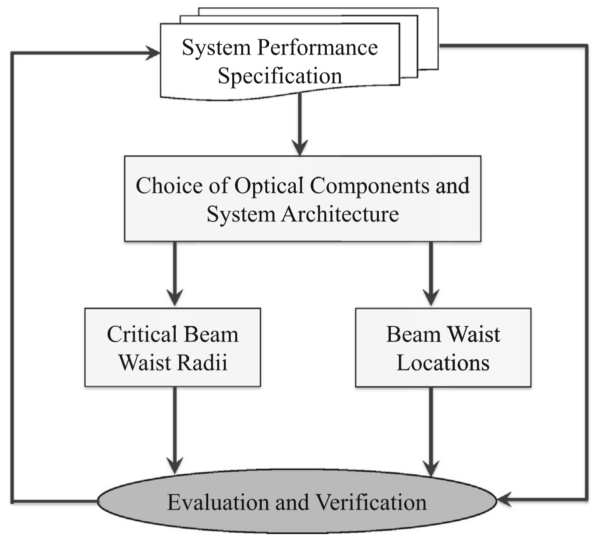

Fig. 8. Quasi-optical system design flow [Gol98]

Table 2. Beam waist criticality of different quasi-optical components [Gol98]

Nature of criticality	Optical component
Non-critical	Polarization
Moderately critical	Diffraction grating, Plate filters, Dielectric filled Fabry Perot resonator
Highly critical	Beam interferometer, Fabry Perot interferometer
Determined by components	Resonators, Feed horn

6. Formal analysis of quasi-optical systems

Given the quasi-optical system design and performance specification, i.e., the information about the size of the overall system, operating frequencies and coupling requirements, we can break the design procedure into four steps as shown in Fig. 8.

- *Determination of the system architecture and quasi-optical components:* The system architecture means the arrangement of optical components (lenses or mirrors), their nature (i.e., reflective or transmissive) and the ability to process frequency bands. In industrial settings, this initial decision is of central importance because the choice of the components can only be considered correct after executing all the steps mentioned in Fig. 8.
- *Beam waist radius:* The beam waist radius provides a suitable measure to evaluate how each component modifies the Gaussian beam. In practice, there are many useful quasi-optical components for which beam waist radius is not important from an application viewpoint (e.g., polarization rotators, which rotate the polarization axis of the light beams [Gol98]). So one of the most important design criteria is the identification of all the components in a system for which the beam waist radius is critical. A summary of such quasi-optical components along with their criticality is described in Table 2.
- *Beam waist location:* The coupling of a Gaussian beam among two optical components is critical to increase the overall performance of systems such as laser resonators [ST07] and feed-horns [Gol98]. This can be done by the identification of the beam waist location along with the beam waist radius of Gaussian beams at the input and output of each quasi-optical component.
- *Evaluation and verification:* Finally, the last step is to evaluate and verify that the selected architecture of the quasi-optical system meets the performance specification, i.e., Gaussian beam waist radius and location are suitable for the correct operation. Moreover, in some practical situations it is compulsory to evaluate the magnification, which is a ratio of the minimum beam waists of input and output Gaussian beams.

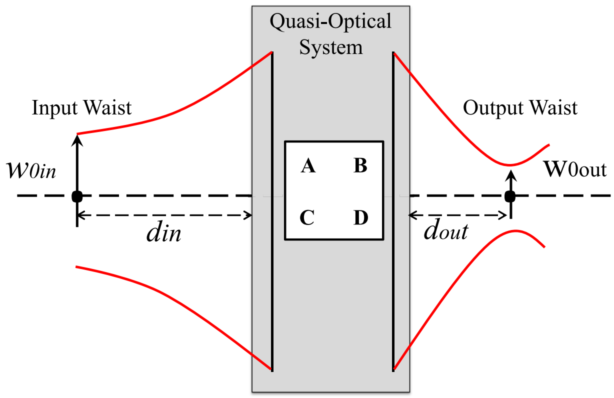

Fig. 9. Generalized quasi-optical system

In order to perform the verification and evaluation (Fig. 8) of quasi-optical systems, we formally derive the expressions for the beam waist radius and location of the output Gaussian beam in the following section.

6.1. Gaussian beam transformation by a generalized quasi-optical system

The generalized properties of beam transformation through a quasi-optical system can be analyzed using the ABCD-law as described in Sect. 5.3. We consider a generic case in which a quasi-optical system is modeled as an arbitrary ABCD-matrix as shown in Fig. 9. The input waist radius $w_{0_{in}}$ of the Gaussian beam is located at distance d_{in} from the input reference plane, and the output waist radius $w_{0_{out}}$, is located at distance d_{out} from the output reference plane. In this situation, the whole system is composed of three subsystems, i.e., a free space (n_i, d_{in}), a quasi-optical system (which we model as an ABCD-matrix), and another free space, i.e., (n_0, d_{out}). Our main goal is to derive the generic expression for the beam waist radius and its location as these are the two critical requirements in the design and analysis of quasi-optical systems. To this aim, we require three steps: (1) building a formal model of the quasi-optical system described in Fig. 9 and then verifying the equivalent matrix; (2) deriving the ABCD-law using the system equivalent matrix; and (3) deriving the general expressions for the output beam width radius and its location, i.e., $w_{0_{out}}$ and d_{out}, respectively. We formally model the quasi-optical system described in Fig. 9 as follows:

Definition 6.1 (*Quasi-Optical System Model*)
$\vdash \forall \mathtt{n_i}\ \mathtt{d_{in}}\ \mathtt{sys}\ \mathtt{n_0}\ \mathtt{d_{out}}.$
$\quad \mathtt{quasi\_optical\_system\ sys\ d_{in}\ d_{out}\ n_i\ n_0} = [[\ ], \mathtt{n_i}, \mathtt{d_{in}};\ \mathtt{sys};\ [\ ], \mathtt{n_0}, \mathtt{d_{out}}]$

where sys represents the quasi-optical system, the parameters $\mathtt{n_i}$ and $\mathtt{n_0}$ represent the refractive index at the input and output, respectively. We next verify the equivalent matrix relation when the system is represented as an arbitrary ABCD-matrix as follows:

Theorem 6.1 (Matrix of Quasi-Optical System)
$\vdash \forall \mathtt{sys}\ \mathtt{d_{in}}\ \mathtt{d_{out}}\ \mathtt{n_i}\ \mathtt{n_0}\ \mathtt{A}\ \mathtt{B}\ \mathtt{C}\ \mathtt{D}.$

$$\mathtt{system\_composition\ sys} = \begin{bmatrix} \mathtt{A} & \mathtt{B} \\ \mathtt{C} & \mathtt{D} \end{bmatrix} \Longrightarrow$$

$$\mathtt{composed\_system\ (quasi\_optical\_system\ sys\ d_{in}\ d_{out}\ n_i\ n_0)} =$$
$$\begin{bmatrix} \mathtt{A + Cd_{out}} & \mathtt{Ad_{in} + B + Cd_{in}d_{out} + Dd_{dout}} \\ \mathtt{C} & \mathtt{Cd_{in} + D} \end{bmatrix}$$

The proof of this theorem involves rewriting the definitions of `quasi_optical_system` and `composed_system` along with the corresponding matrices of the input and output free spaces.

Consequently, we verify the ABCD-law for the quasi-optical system model (Definition 6.1) as follows:

Theorem 6.2 (Quasi-optical System (ABCD))

$$
\begin{array}{l}
\vdash \forall \mathtt{sys}\ \mathtt{d_{in}}\ \mathtt{d_{out}}\ \mathtt{n_i}\ \mathtt{n_0}\ \mathtt{gbeam}\ \mathtt{lam}\ \mathtt{A}\ \mathtt{B}\ \mathtt{C}\ \mathtt{D}.\\
\quad \mathtt{is\_valid\_gbeam\_in\_c\_system\ gbeam\ lam\ (quasi\_optical\_system\ sys\ d_{in}\ d_{out}\ n_i\ n_0)} \wedge\\
\quad \mathtt{is\_valid\_gen\_beam\ gbeam} \wedge\\
\quad \mathtt{is\_valid\_composed\_system\ (quasi\_optical\_system\ sys\ d_{in}\ d_{out}\ n_i\ n_0)} \wedge\ 0 < \mathtt{lam} \wedge\\
\quad \begin{bmatrix} \mathtt{A} & \mathtt{B} \\ \mathtt{C} & \mathtt{D} \end{bmatrix} = \mathtt{system\_composition\ sys} \implies\\
\quad \mathtt{let}\ (\mathtt{z}, \mathtt{w_0}), (\mathtt{z'}, \mathtt{w'_0}) = \mathtt{beam\_origin\ gbeam\ in}\\
\quad \mathtt{let}\ (\mathtt{z_n}, \mathtt{w_{0n}}) = \mathtt{beam\_end\ gbeam\ in}\\
\quad \mathtt{q}\ (\mathtt{z_n}, \mathtt{w_{0n}}, \mathtt{lam}) = \dfrac{\mathtt{Aq}\ (\mathtt{z}, \mathtt{w_0}, \mathtt{lam}) + \mathtt{B}}{\mathtt{Cq}\ (\mathtt{z}, \mathtt{w_0}, \mathtt{lam}) + \mathtt{D}}
\end{array}
$$

where the first assumption ensures the valid behavior of the beam when it propagates through the quasi-optical system. The proof of this theorem is a direct consequence of Theorem 5.7.

Our next step is to verify the general expressions for the output beam waist radius. Here, one important point is to ensure that we are only interested in the Gaussian beam waist at the input, which means that the real part of the input q-parameter should be 0. We include this requirement in the verification of the following main theorem:

Theorem 6.3 (Beam Waist Radius and Location)

$$
\begin{array}{ll}
\multicolumn{2}{l}{\vdash \forall \mathtt{sys}\ \mathtt{gbeam}\ \mathtt{d_{in}}\ \mathtt{d_{out}}\ \mathtt{lam}\ \mathtt{n_i}\ \mathtt{n_0}\ \mathtt{w_{0in}}\ \mathtt{w_{0out}}\ \mathtt{z}\ \mathtt{z_n}\ \mathtt{A}\ \mathtt{B}\ \mathtt{C}\ \mathtt{D}.}\\
[\mathtt{H_1}] & \mathtt{sys\_constraints\ (quasi\_optical\_system\ sys\ d_{in}\ d_{out}\ n_i\ n_0)} \wedge\\
[\mathtt{H_2}] & \begin{bmatrix} \mathtt{A} & \mathtt{B} \\ \mathtt{C} & \mathtt{D} \end{bmatrix} = \mathtt{system\_composition\ sys} \wedge\\
[\mathtt{H_3}] & (\mathtt{z}, \mathtt{w_{0in}}) = \mathtt{beam\_origin\ gbeam} \wedge (\mathtt{z_n}, \mathtt{w_{0out}}) = \mathtt{beam\_end\ gbeam} \wedge\\
[\mathtt{H_4}] & \mathrm{Re}(\mathtt{q}\ (\mathtt{z}, \mathtt{w_{0in}}, \mathtt{lam})) = 0 \wedge \mathrm{Re}(\mathtt{q}\ (\mathtt{z_n}, \mathtt{w_{0out}}, \mathtt{lam})) = 0 \wedge\\
& (\mathtt{Cd_{in}} + \mathtt{D})^2 + (\mathtt{C\ rayleigh\_range\ w_{0in}\ lam})^2 \neq 0 \implies\\
& \mathtt{d_{out}} = -\dfrac{(\mathtt{Ad_{in}} + \mathtt{B})(\mathtt{Cd_{in}} + \mathtt{D}) + \mathtt{AC}(\mathtt{rayleigh\_range\ w_{0in}\ lam})^2}{(\mathtt{Cd_{in}} + \mathtt{D})^2 + (\mathtt{C}(\mathtt{rayleigh\_range\ w_{0in}\ lam})^2)} \wedge\\
& \mathtt{w_{0out}}^2 = \dfrac{(\mathtt{AD} - \mathtt{BC})\ \mathtt{w_{0in}}^2}{(\mathtt{Cd_{in}} + \mathtt{D})^2 + (\mathtt{Cw_{0in}}^2 \dfrac{\pi}{\mathtt{lam}})^2}
\end{array}
$$

where the first assumption $[\mathtt{H_1}]$ packages three conditions as system constraints, i.e., the validity of the composed optical system architecture, the validity of a general beam and the valid behavior of a general beam in the composed system. The second assumption $[\mathtt{H_2}]$ ensures that the composed system can be described by an arbitrary matrix. Finally, the third and fourth assumptions (i.e., $[\mathtt{H_3}]$ and $[\mathtt{H_4}]$) ensure that the real part of q-parameters are zero and the values $\mathtt{d_{out}}$ and $\mathtt{w_{0out}}$ are finite. The proof of Theorem 6.3 is mainly based on Theorem 6.2 and involves the properties of complex numbers (mainly, equating the real and imaginary parts of the input and output q-parameters).

Note that the expressions obtained in Theorem 6.3 can be applied to any quasi-optical system and Gaussian beam parameters. The given system itself can be arbitrarily complicated, and the analysis is reduced to the problem of obtaining its overall ABCD-matrix from a cascaded representation of its constituent optical components. We apply these results to verify a real-world optical system in the next section.

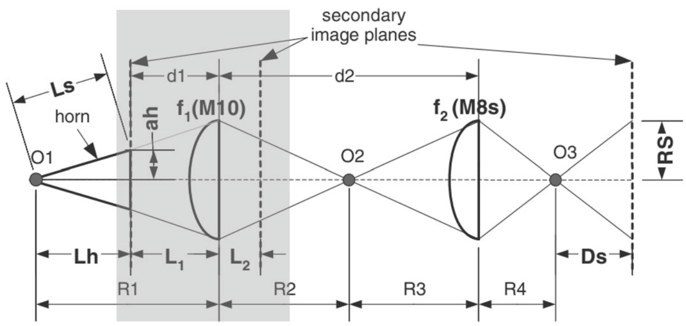

Fig. 10. Optical layout of the APEX telescope facility receiver [NLD+09]

6.2. Optics verification of Atacama pathfinder experiment (APEX) telescope receiver

The Atacama Pathfinder EXperiment (APEX)[3] is a single dish (12-m diameter) telescope for millimeter and sub-millimeter astronomy, which has been in operation since its inauguration in 2005 [NLD+09]. The main mission of the APEX is to conduct the astronomical study of cold dust and gas in our own milky way and in distant galaxies. Recent observations based on APEX reveal the cradles of massive star-formation throughout our galaxy [APE15]. In addition to these interesting aspects of the APEX telescope, the other main function is radiometry which helps to provide reliable weather forecasts and environmental dynamics. One of the main modules of the APEX is the Swedish Heterodyne Facility Instrument (SHeFI) receiver which was installed in 2008. In [NLD+09], the authors used a Quasi-optics-based model for the SHeFI receiver to derive the conditions in terms of beam parameters using a paper-and-pencil based proof approach. Furthermore, these constraints are used to optimize (i.e., minimization of dimensions and distortions) the telescope design for all optical components. In this thesis, we propose to formally analyze the SHeFI receiver within the sound core of HOL Light by using our formalization of Gaussian beams and quasi-optical systems. The main component of the SHeFI receiver is the optical system, which is designed to provide the coupling of the SHeFI channels and other instruments within the telescope. The optical layout of the receiving cabin is shown in Fig. 10. The points O_1, O_2, and O_3 represent focal points, traced from the original Cassegrain focal point [NLD+09]. Here, M_{8s} and M_{10} are ellipsoidal mirrors with focal distances f_2 and f_1 [NLD+09], respectively. In this situation, the Gaussian beams' transformation is the best possible way to understand the processing of light in the receiver module of the APEX telescope [NLD+09].

Our main goal is to verify a generic expression for the system magnification, which is represented by the ratio of output and input beam waist radii, i.e., $\frac{w0_{out}}{w0_{in}}$. This can be done by using already verified theorems in our framework. We analyze one module of the receiving system, i.e., the gray shaded region in Fig. 10. Indeed, this can be considered as the quasi-optical system with the input and output distances L_1 and L_2 and a thin lens inside. Our problem is mainly reduced to the derivation of the equivalent matrix relation for the thin lens and then utilize Theorems 6.2 and 6.3. A thin lens is represented as the composition of two transmitting spherical interfaces such that any variation of the beam parameters is neglected between both interfaces. So, at the end, a thin lens is the composition of two spherical interfaces with a null width free space in between. We formalize a thin lens as follows:

Definition 6.2 (*Thin Lens*)

```
⊢ ∀R1 R2 n0 n1. thin_lens R1 R2 n0 n1 =
  ([(n0,0),spherical R1,transmitted; (n1,0),spherical R2,transmitted],(n0,0))
```

where $\mathtt{R}_1$, $\mathtt{R}_2$, $\mathtt{n}_1$, $\mathtt{n}_2$, represent the radius of curvatures of two interfaces and the refractive indices of the input and output planes, respectively. We prove that a thin lens is indeed a valid optical system if the corresponding parameters satisfy some constraints:

[3] http://www.apex-telescope.org/.

Theorem 6.4 (Valid Thin Lens)
$\vdash \forall \mathtt{R_1\ R_2\ n_0\ n_1}.\ \mathtt{R_1} \neq 0 \wedge \mathtt{R_2} \neq 0 \wedge 0 < \mathtt{n_0} \wedge 0 < \mathtt{n_1} \Longrightarrow$
$\quad \mathtt{is\_valid\_optical\_system\ (thin\_lens\ R_1\ R_2\ n_0\ n_1)}$

The proof of this theorem is done automatically by our developed tactic, called `VALID_OPTICAL_SYSTEM_TAC`. Next, we verify the matrix relation of the thin lens as follows:

Theorem 6.5 (Thin Lens Matrix)
$\vdash \forall \mathtt{R_1\ R_2\ n_0\ n_1}.\ \mathtt{R_1} \neq 0 \wedge \mathtt{R_2} \neq 0 \wedge 0 < \mathtt{n_0} \wedge 0 < \mathtt{n_1} \Longrightarrow$

$$\mathtt{system\_composition\ (thin\_lens\ R_1\ R_2\ n_0\ n_1)} = \begin{bmatrix} 1 & 0 \\ \frac{\mathtt{n_1} - \mathtt{n_0}}{\mathtt{n_0}}\left(\frac{1}{\mathtt{R_2}} - \frac{1}{\mathtt{R_1}}\right) & 1 \end{bmatrix}$$

At this point, we have all the necessary ingredients to analyze the module of interest of the SHeFI receiver as shown in Fig. 10. We reuse the definition of the generalized quasi-optical system (Definition 6.1) to define the module as follows:

Definition 6.3 (*SheFI Receiver Module*)
$\vdash \forall \mathtt{R_1\ R_2\ L_1\ L_2\ n_1\ n_2}.$
$\quad \mathtt{SHeFI\_receiver\_model\ L_1\ L_2\ n_1\ n_2\ R_1\ R_2} =$
$\quad \mathtt{quasi\_optical\_system\ (thin\_lens\ R_1\ R_2\ n_1\ n_2)\ L_1\ L_2\ n_1\ n_2}$

Finally, we verify the system magnification of the SheFI receiver module as follows:

Theorem 6.6 (APEX Beam Waist)
$\vdash \forall \mathtt{gbeam\ L_1\ L_2\ lam\ n_1\ n_2\ R_1\ R_2}.$
$\quad \mathtt{SHeFI\_constraints\ gbeam\ L_1\ L_2\ lam\ n_1\ n_2\ R_1\ R_2} \Longrightarrow$

$$\mathtt{let}\ \mathtt{f} = -\frac{1}{\left(\frac{\mathtt{n_2} - \mathtt{n_1}}{\mathtt{n_1}} \frac{1}{\mathtt{R_2}} - \frac{1}{\mathtt{R_1}}\right)}\ \mathtt{and}$$
$$\quad (\mathtt{z}, \mathtt{w_0}) = \mathtt{beam\_origin\ gbeam\ and}$$
$$\quad (\mathtt{z_n}, \mathtt{w_{0n}}) = \mathtt{beam\_end\ gbeam\ in}$$
$$(1 - \mathtt{L_1}\frac{1}{\mathtt{f}})^2 + (\frac{1}{\mathtt{f}}(\mathtt{rayleigh\_range\ w_{0in}\ lam}))^2 \neq 0 \Longrightarrow$$
$$\frac{\mathtt{w_{0out}}^2}{\mathtt{w_{0in}}^2} = \frac{1}{\left(1 - \mathtt{L_1}\frac{1}{\mathtt{f}}\right)^2 + \left(\frac{1}{\mathtt{f}}(\mathtt{rayleigh\_range\ w_{0in}\ lam})\right)^2}$$

We verify the above expression using Theorems 6.5 and 6.3. Note that Theorem 6.6 is in a general form and can further be utilized to reason about different cases such as the input and output distances (L_1 and L_2) are equal to f, or $2f$, in order to maximize or minimize the magnification depending upon the practical requirements. We can easily evaluate the real values of the parameters provided by physicists and optical engineers. Indeed, the only requirement is to check `SHeFI_constraints`, which ensures the valid behavior of `gbeam` in thin lens-based quasi-optical systems.

This completes the formal analysis of the quasi-optical systems based on the Gaussian beam transformation. Due to the generic nature of our models and verified theorems, we have been able to analyze a cost and safety critical application, i.e., the receiver module of the APEX telescope within the sound core of HOL Light theorem prover. This improved accuracy comes at the cost of the time and effort spent, formalizing the underlying theories of ray optics and Gaussian beams. But, the availability of such a formalized infrastructure significantly reduces the time required to analyze the quasi-optical systems and the APEX telescope application. For example, the analysis of the application, i.e., the modeling and verification of the system magnification of the receiver module took less than 100 lines of HOL light code and a couple of hours by an expert user of HOL Light. Apart from the formalization of a number of concepts of ray optics and Gaussian optics, another contribution of our work is to bring out all the hidden assumptions about the physical models of ray, beams, lenses and mirrors, which otherwise are not mentioned in the optics literature (e.g., [Trä07]). Moreover, we automatized parts of the verification task by introducing new tactics. Some of these tactics are specialized to verify (or simplify) the proofs related to our formalization of optical systems (e.g., `VALID_OPTICAL_SYSTEM_TAC`). However, some tactics are general and can be used in different verification tasks involving matrix/vector operations. An example of such a tactic is `common_prove`, which allowed us to verify the ray-transfer matrices in our development. The automation of proof tasks is mainly done using the derived rules and tactics of HOL Light, so that the application to a particular system does not involve the painful manual proofs often required with interactive (higher-order logic) theorem proving.

7. Conclusion and future work

In this paper, we reported a new engineering application of formal methods in the area of Gaussian optical systems. In particular, we proposed to use higher-order-logic theorem proving as a complementary technique to verify some important properties of real-world optical systems. We also presented the formalization of Gaussian beams based on the notion of q-parameters along with the verification of ABCD-law for arbitrarily composed optical systems. Building on top of this infrastructure, we analyzed a cost and safety critical application, i.e., the receiving unit of the Atacama Pathfinder Experiment (APEX) telescope. The analysis accuracy of theorem proving allowed us to unveil all the assumptions required to verify the magnification of the SHeFI receiver.

The formal analysis of Gaussian optical systems along with the real-world critical applications (e.g., APEX telescope) provides some thoughtful indications: (1) Theorem proving systems have reached maturity, where complex physical models can be expressed with less effort than ever before; and (2) Formal methods can assist in the verification of futuristic optical systems, which are largely becoming parts of critical applications such as military setups, biomedical surgeries and space missions. However, the question of the utilization of higher-order-logic theorem proving in industrial settings (particularly, physical systems) still persists due to the huge amount of time required to formalize the underlying theories. We believe that an important factor is the gap between theorem proving and engineering communities, which limits its usage in industrial settings. For example, it is hard to find engineers (or physicists) with a theorem proving background and vice-versa. One of the several solutions to tackle this issue is the continuous formal development of optics theories including the libraries of the most frequently used optical components and devices, which can ultimately reduce the cost of using formal methods (particularly theorem proving) as an integral part of the design and verification of physical systems. The work reported in this paper can be considered as a step towards this goal with more efforts to follow in the same or closely related disciplines such as quantum optics, photonic signal processing and optoelectronics. Our plan is to extend this work in order to obtain an extensive library of verified optical components such as phase-conjugate mirrors and resonators with active sources inside [HW05a], which would allow a practical use of our formalization in industry. We also intend to formalize and verify the correctness and soundness of the ray and Gaussian beam tracing algorithms [Trä07], which are included in almost all optical systems design tools.

The formal analysis of real-world systems involves mathematical models which usually represent an approximated behavior of physical phenomena. In order to formally treat the approximations introduced in physical models, we need to consider non-standard analysis and asymptotic notations. For example, small angle approximation (or paraxial approximation) entails that $\sin(\theta) \approx \theta$, and it can be treated using asymptotic notations. Interestingly, both non-standard analysis [Fle01] and asymptotic notations [AD04] are available in Isabelle/HOL.[4] Our work can be extended by using these concepts which will bring more rigor to the formal models of optical systems. Similar concepts can also be used to formally prove that ray optics models are approximations of wave and electromagnetic optics models. Recently, a learning-assisted automated reasoning

[4] http://isabelle.in.tum.de.

support [KU14] has been developed for HOL Light, which can also be applied in our formalization to see the future of automation tools for optics and the formalization of Physics [KUS+15]. Moreover, our formalization can be adapted to other widely used theorem provers (e.g., PVS[5] and Isabelle/HOL), which provide efficient automation support. In the future, it might be possible to automatically port our formalization to other theorem provers with the maturity of some ongoing projects like the Open Theory Project[6] and ProofPeer.[7]

References

[AD04] Avigad J, Donnelly K (2004) Formalizing O notation in Isabelle/HOL. In: Automated reasoning, Lecture Notes in Computer Science, vol 3097. Springer, Berlin, Heidelberg, pp 357–371

[APE15] Atacama Pathfinder EXperiment (APEX) (2015) http://www.apex-telescope.org/

[ASM+14] Afshar SK, Siddique U, Mahmoud MY, Aravantinos V, Seddiki O, Hasan O, Tahar S (2014) Formal analysis of optical systems. Math Comput Sci 8(1):39–70

[CSBE10] Chabory A, Sokoloff J, Bolioli S, Elis K (2010) Application of gaussian beam based techniques to the quasi-optical systems of radiofrequency radiometers. In: European Conference on Antennas and Propagation, vol 2010, pp 12–16

[Dam05] Damask JN (2005) Polarization optics in telecommunications. Springer Series in Optical Sciences. Springer

[FAGP13] Franke-Arnold S, Gay SJ, Puthoor IV (2013) Quantum process calculus for linear optical quantum computing. In: Reversible Computation, Lecture Notes in Computer Science, vol 7948. Springer, pp 234–246

[Fle01] Fleuriot JD (2001) Nonstandard geometric proofs. In: Automated deduction in geometry, Lecture Notes in Computer Science, vol 2061. Springer, pp 246–267

[Gol98] Goldsmith PF (1998) Quasioptical systems: gaussian beam quasioptical propogation and applications. IEEE Press Series on RF and Microwave Technology. Wiley

[Gri05] Griffiths DJ (2005) Introduction to quantum mechanics. Pearson Prentice Hall

[Har09a] Harrison J (2009) Handbook of practical logic and automated reasoning. Cambridge University Press

[Har09b] Harrison J (2009) HOL light: an overview. In: Theorem Proving in Higher Order Logics, Lecture Notes in Computer Science, vol 5674. Springer, pages 60–66

[Har13] Harrison J (2013) The light theory of Euclidean space. J Autom Reason. 50(2):173–190

[HW05a] Hodgson N, Weber H (2005) Optical resonators: fundamentals, advanced concepts, applications. Springer Series in Optical Sciences. Springer

[HW05b] Hodgson N, Weber H (2005) Optical resonators: fundamentals, advanced concepts, applications. Springer

[KAHT14] Khan-Afshar S, Hasan O, Tahar S (2014) Formal analysis of electromagnetic optics. In: Novel optical systems design and optimization, SPIE, vol 9193, pp 91930A–91930A–14

[KL66] Kogelnik H, Li T (1966) Laser beams and resonators. Appl Opt. 5(10):1550–1567

[KU14] Kaliszyk C, Urban J (2014) Learning-assisted automated reasoning with flyspeck. J Autom Reason. 53(2):173–213

[KUS+15] Kaliszyk C, Urban J, Siddique U, Khan-Afshar S, Dunchev C, Tahar S (2015) Formalizing physics: automation, presentation and foundation issues. In: Intelligent computer mathematics, Lecture Notes in Computer Science, vol 9150. Springer, pp 288–295

[LAS15] LASCAD (2015) http://www.las-cad.com/

[Moo01] Mookherjea S(2001) Analysis of optical pulse propagation with two-by-two (ABCD) matrices. Phys Rev E. 64(016611):1–10

[MPM+11] Malak M, Pavy N, Marty F, Peter Y, Liu AQ, Bourouina T (2011) Stable, high-Q fabry-perot resonators with long cavity based on curved, all-silicon, high reflectance mirrors. In: IEEE international conference on micro electro mechanical systems, pp 720–723

[MT14] Mahmoud MY, Tahar S (2014) On the quantum formalization of coherent light in HOL. In: NASA formal methods, LNCS, vol 8430. Springer, pp 128–142

[NKST98] Nakazawa M, Kubota H, Sahara A, Tamura K (1998) Time-domain ABCD Matrix Formalism for Laser Mode-Locking and Optical Pulse Transmission. IEEE J Quantum Electron 34(7):1075–1081

[NLD+09] Nystrm O, Lapkin I, Desmaris V, Dochev D, Ferm S-E, Fredrixon M, Henke D, Meledin D, Monje R, Strandberg M, Sundin E, Vassilev V, Belitsky V (2009) Optics design and verification for the APEX Swedish heterodyne facility instrument (SHeFI). J Infrared Millim Terahertz Waves. 30(7):746–761

[Opt15] Optica (2015) http://www.opticasoftware.com/

[reZ15] reZonator (2015) http://www.rezonator.orion-project.org/

[SAT13a] Siddique U, Aravantinos V, Tahar S (2013) Formal stability analysis of optical resonators. In: NASA formal methods, Lecture Notes in Computer Science, vol 7871, pp 368–382

[SAT13b] Siddique U, Aravantinos V, Tahar S (2013) On the formal analysis of geometrical optics in HOL. In: Automated deduction in geometry, Lecture Notes in Computer Science, vol 7993, pp 161–180

[Sid15] Siddique U (2015) Formal analysis of gaussian optical systems: source code. http://hvg.ece.concordia.ca/projects/optics/gaussian.html

[5] http://pvs.csl.sri.com.

[6] http://www.gilith.com/research/opentheory/.

[7] http://www.proofpeer.net.

[ST07] Saleh BEA, Teich MC (2007) Fundamentals of photonics. Wiley

[SXS+11] Su B, Xue J, Sun L, Zhao H, Pei X (2011) Generalised ABCD matrix treatment for laser resonators and beam propagation. Opt Laser Technol 43(7):1318–1320

[SZL+06] Song WZ, Zhang XM, Liu AQ, Lim CS, Yap PH, Hosseini HMM (2006) Refractive index measurement of single living cells using on-chip Fabry-Perot cavity. Appl Phys Lett 89(20):203901

[Trä07] Träger F (2007) Handbook of lasers and optics. Springer.

[WA05] Wilson WC, Atkinson GM (2005) MOEMS modeling using the geometrical matrix toolbox. Technical report, NASA, Langley Research Center

[Wel91] Wellner M (1991) Wave optics. In: Elements of physics, pp 543–575. Springer

Received 28 May 2015

Accepted in revised form 15 February 2016 by Stephan Merz, Jun Pang, and Jin Song Dong

Published online 15 March 2016